Fibre Channel over Ethernet: a Pragmatic Approach to Data Center Network Convergence

Business white paper

Introduction

As traffic volumes continue to mount and data storage requirements increase, enterprises must cope with the multiple network silos they have built over the years to address their evolving data center needs. They often must support three distinct networks—a local area network (LAN) for core business application Ethernet traffic, a Fibre Channel storage area network (SAN) for storage data traveling between servers and corresponding storage devices, and high-performance computing (HPC) environments that rely on low-latency InfiniBand networking. Significant challenges are created by the distinctive technologies, protocols, and management requirements of each of these networking environments, frequently resulting in dramatic cost increases and operational complexity.

Clearly, enterprises must find a way to simplify data center infrastructures and reduce operational expenses. Many have decided that the best approach is to chart a measured route toward a unified infrastructure. Gradual migration is the logical approach—one characterized by low risk and high reward. Leaping directly to a unified network architecture can be too risky, not to mention costly and disruptive.

As enterprises consider options for enabling a sensible migration to a converged network scenario, emerging Fibre Channel over Ethernet (FCoE) technology quickly comes to the fore. Developed by a group of storage and data network vendors working within the INCITS T11 committee, FCoE is designed to let native Fibre Channel and standard network traffic run alongside each other over a 10 Gigabit Ethernet infrastructure. The Fibre Channel frames are encapsulated over the Ethernet network, allowing the consolidation of computing and storage (LAN and SAN) I/O environments at the server network edge in a high-performance, low-cost, and standards-based manner.

The FCoE Standard

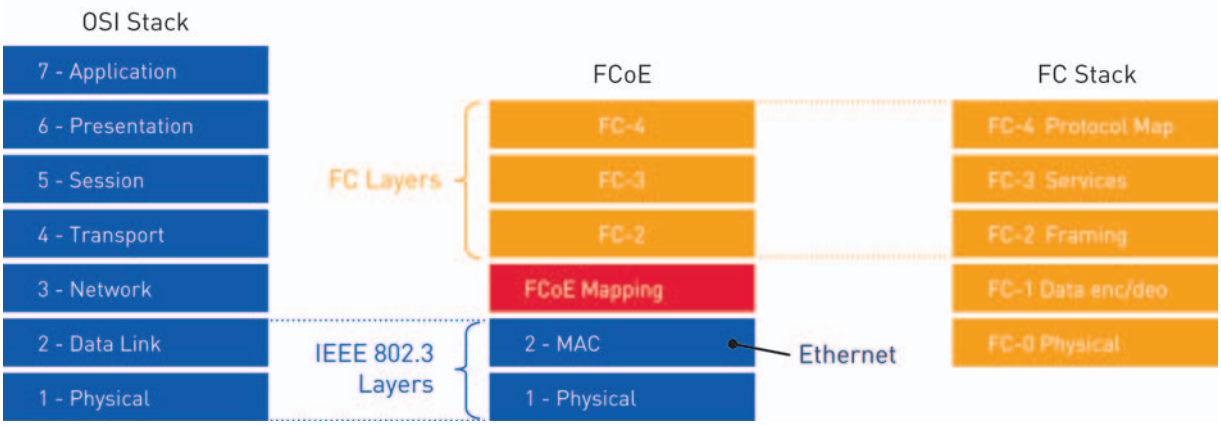

Figure 1. FCoE, a proposed standard within the INCITS T11 FC-BB-5 committee is a less cumbersome alternative to iSCSI, an IP-based standard for linking data storage devices. iSCSI requires a gateway function to terminate/originate the TCP/IP connections; FCoE operates at the same Open System Interconnection (OSI) layer as IP, but can handle FC fabrics and devices as local area extensions of Ethernet without the overhead of TCP.

Bringing storage and data networks together with FCoE

Most of today's enterprise data centers complement their server infrastructure with a large, centrally managed pool of storage, or SAN. Fibre Channel is the established SAN interconnect, spanning from redundant Fibre Channel host bus adapters (HBAs) in each server through SAN switches, and eventually to enterprise storage systems. Data centers supporting mission-critical operations also use redundant Ethernet network interface cards (NICs) in their servers and redundant pairs of switches at each layer in their network architecture so that any one component can fail without impacting service availability.

FCoE provides a straightforward solution targeted at consolidating these multiple redundant infrastructures by encapsulating Fibre Channel frames into Ethernet frames using a lossless transmission model (Figure 1). FCoE maintains the Fibre Channel operation model, which provides seamless connectivity between the two networks. For example, the encapsulation occurs with no fragmentation, eliminating the need for higher-level protocols for packet reassembly.

Access layer simplification

With this technology, IT organizations can incorporate 10 Gigabit Ethernet FCoE-capable switches into the access layer, simplifying network topology and enabling use of only a single pair of new interface cards called converged network adapters (CNAs)—and requisite cabling—to connect each server to both the Ethernet and Fibre Channel networks. The FCoE-aware switches separate LAN and SAN traffic and seamlessly connect to existing storage systems.

Such data center I/O consolidation requires 10 Gigabit Ethernet, a key enabling technology with its greater bandwidth and enhanced feature set. Additionally, increased use of multi-core server and blade technology and the explosion of virtualized server and storage workloads in the data center are driving the need for the higher performance and scale enabled by 10 Gigabit Ethernet. Lower performing legacy data center networks cannot handle the requirements placed on them by such infrastructure changes.

As enterprises upgrade their server infrastructure with 10 Gigabit Ethernet to support requirements there, it also behooves them to add the technology in the data network. Top-of-rack switches, which will bring together the data and storage I/O, make a great starting point.

Of course, Ethernet does come with some technical challenges for the storage world, notably that it is a best-effort protocol built with the ability to handle high latencies and packet loss that are intolerable in a storage environment. To address these concerns, the IEEE is working on extending the Ethernet specification to give 10 Gigabit Ethernet the flexibility to seamlessly support FCoE.

Addressing latency and loss

Data traffic is generally capable of handling the occasional dropped packet; one node simply asks the other to retransmit, using the Transmission Control Protocol (TCP), or if the User Datagram Protocol (UDP) is in use, by ignoring the missing data. But storage networks have little tolerance for a dropped packet; retransmits are considered a necessary evil rather than a routine occurrence. As a result, storage protocols such as iSCSI and Fibre Channel Protocol, or FCP (the FC equivalent of TCP), emphasize flow control, and FC engineering leans heavily on over-provisioning bandwidth and switching capacity to ensure against dropped packets.

To accommodate the sensitivity of SAN traffic to packet loss and latency, the IEEE 802.1 Data Center Bridging (DCB) Standards Task Group is working on enhancing the Ethernet protocol with the ability to support end-to-end congestion mechanisms, different classes of service, and flow control (see sidebar). These should achieve lossless Fibre Channel transmissions over 10 Gigabit Ethernet connections, as specified in the FCoE standard.

An enterprise's migration path couldn't be clearer—lossless Ethernet standards are nearing finalization, and low-cost 10 Gigabit Ethernet CNAs and FCoE-ready, cost-effective 10 Gigabit Ethernet switches are becoming available. FCoE-ready 10 Gigabit Ethernet top-of-rack and core data center HP switches are helping pave the way toward this convergence.

The benefits of consolidation

Less is more, much more

Today, by consolidating data and storage I/O in top-of-rack switches, enterprises can reduce cost and complexity, ensure greater and more flexible performance, and lower power consumption—all the while providing seamless connectivity with existing Ethernet and storage networks. With simplified FCoE topology, the previous need for a minimum of four interfaces per server—two NICs and two HBAs (and in many cases as many as six to eight server NICs)—is replaced by the use of only two 10 Gigabit Ethernet CNA adapters per server.

This simplification eliminates a large number of Fibre Channel HBAs and allows IT organizations to more fully enable connection of every server to its central pool of storage. FCoE also helps extend data center SANs to servers with insufficient expansion capability, or where Fibre Channel HBA cost is prohibitive. Now virtually every data center server can leverage the benefits of centrally managed storage.

IEEE Working Group specifications

802.1Qau: Congestion Notification
(http://www.ieee802.org/1/pages/802.1au.html)
The 802.1Qau protocol defines a way to throttle back on frames sent across the network. It relies on congestion points, which are basically switch output queues that have the capability of monitoring the queue length/packet sampling; reaction points, which are basically end station output queues with a flow output rate limiter; and congestion notification messages. In practice, when congestion occurs, the congestion points that monitor queue lengths send a "quantized feedback" value in a congestion-notification-message packet to the reaction point. With this information, the reaction point knows how much frame rate reduction is needed to clear congestion.

802.1Qaz: Enhanced Transmission Selection
(http://www.ieee802.org/1/pages/802.1az.html)
The existing IEEE 802.1Q/p standard's "strict priority" mechanism cannot guarantee minimum bandwidth for a specific traffic class. Without the ability to manage bandwidth allocation properly, high-priority traffic can starve applications with lower priorities. The new 802.1Qaz standard addresses this issue by providing a priority-based processing algorithm to guarantee bandwidth allocation/low latency. In addition, the 802.1Qaz committee is working on Link Layer Discovery Protocol (LLDP) extensions that would allow network devices to exchange capabilities and configuration data using the 802.1AB (LLDP) standard.

802.1Qbb: Priority-based Flow Control
(http://www.ieee802.org/1/pages/802.1bb.html)
After many months of work on the 802.1Qau standard for congestion notification, developers realized that guaranteeing zero frame loss would require a "pause" mechanism (XOFF). Such a mechanism becomes particularly important when multiple high-bandwidth sources burst at the same time. The idea of applying the existing mechanism (PAUSE) on a per-class basis was the genesis of the 802.1Qbb standards effort.

Better overall network performance

Using virtual LANs for security isolation and 802.1p Priority Flow Control settings, enterprises can create high-speed FCoE lanes within each 10 Gigabit Ethernet link to provide high-capacity, lossless carrying capacity for storage traffic. Indeed, with 10 Gigabit Ethernet bandwidth to work with—and 40 and 100 Gigabit Ethernet in the future—they can create multiple FCoE lanes on a single network with room to spare.

Simplified cabling

The ability to carry FCoE and IP traffic over the same physical cables allows IT organizations to cut in half the number of network cables within each rack. This reduces 40 Ethernet and 40 Fibre Channel connections to only 40 FCoE connections in a typical 20-server rack with fully redundant connectivity. Fewer cables can reduce cabling errors and the number of required in-rack fiber runs, as well as increase the speed of configuring and deploying new server racks for greater business agility. In addition, fewer cables lessen restriction in front-to-back airflow, increasing cooling efficiency.

Reduced power consumption

The reduced power consumption that comes from using fewer NICs and switches provides relief to organizations

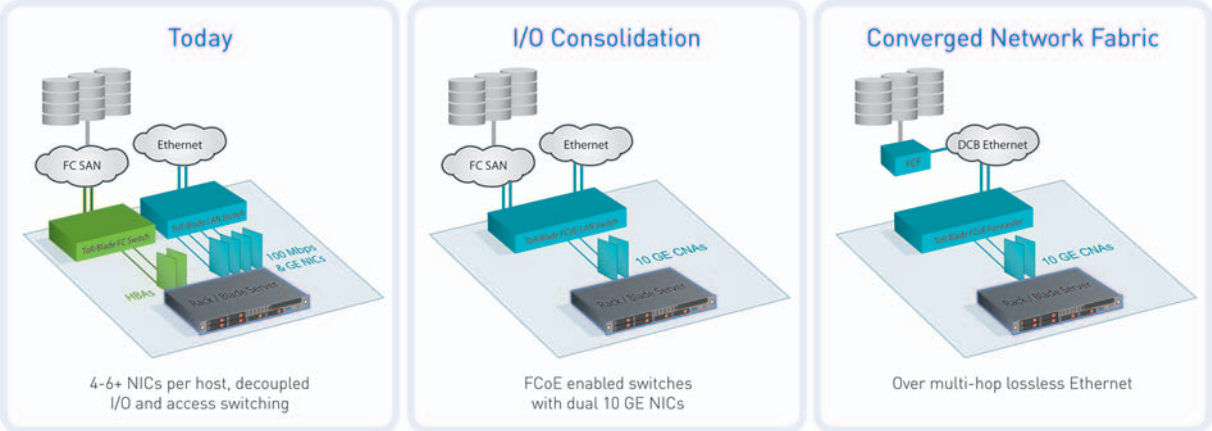

Figure 2. FCoE extends rather than replaces Fibre Channel, allowing seamless integration with Ethernet and Fibre Channel networks at the pace and to the extent that matches business plans. Deploying FCoE at the server-to-network access layer edge enables meaningful I/O consolidation which reduces "in rack" cost and complexity, cabling, and power consumption while delivering up to 10X greater performance.

that are reaching the limits of their data center power and cooling envelopes. With a typical rack of 20 servers and a typical Fibre Channel HBA consuming roughly 12.5 watts, the savings can be as much as 500 watts per rack by eliminating two interface cards per server. If the energy needed to cool the rack is also computed, savings can be doubled to 1,000 watts per rack.

Unified management with CNAs

This new class of converged network adapters enables enterprises to incorporate FCoE without disrupting current data center management practices, software, or the roles of network and storage administrators. CNAs, which can be implemented in hardware or software, present both an Ethernet interface and a Fibre Channel initiator to the server, allowing the operating system to see two physical devices and storage and network administrators to manage their respective domains just as they do today. Consistent management helps ease FCoE deployment while reducing operating expenses.

Simplified migration with investment protection

FCoE extends rather than replaces Fibre Channel, allowing organizations to seamlessly integrate their Ethernet and Fibre Channel networks at their own pace, enabling reduced spending on NICs, cables, and switches, as well as the ability to extend SAN connectivity to servers that could not be reached otherwise. Since FCoE and Fibre Channel can coexist throughout the network, pre-existing management tools and techniques can be used. In addition, because FCoE networks connect directly to Fibre Channel networks, organizations can preserve the investments they have made in their storage infrastructure.

Looking ahead

As enterprise data centers converge Ethernet and Fibre Channel networks, they can improve performance and reduce power consumption, infrastructure complexity, and cost.

Enterprises that are interested in access layer (top-of-rack) LAN/SAN I/O consolidation and 10 Gigabit Ethernet migration will find that HP A5820 series switches provide a flexible combination of Gigabit, 10 Gigabit Ethernet, and FCoE capabilities for cost-effective migration to 10 Gigabit Ethernet while simplifying I/O complexity. They can support copper, fiber, and SFP+ connectivity to a core switch, or be used with up to as many as nine other switches as a small consolidated 10 Gigabit Ethernet core in a "virtual" stacked configuration to support up to 192 10 Gigabit Ethernet connections.

4AA1-7424ENW, Created May 2010