

FCoE/DCЕ: объединенный транспорт для LAN и SAN

Обзор особенностей применения нового протокола — FCoE на базе расширенного транспорта Ethernet, призванного упростить и в дальнейшем удешевить построение инфраструктур LAN- и SAN-сетей за счет их консолидации в рамках одной сетевой инфраструктуры. Появление решений на базе нового FCoE-интерфейса планируется с 4 кв. 2008 г.

Введение

Протокол Fibre Channel over Ethernet (FCoE) — это стандарт, который стал продвигаться с конца 2007 г. и позволяющий инкапсулировать структуру Fibre Channel на транспорт Ethernet-сети, обеспечивая создание смешанных сетей, или одной сети вместо двух — LAN и SAN, в центрах обработки данных. Спецификация стандарта была предложена комитету T11 Национального института стандартизации США (American National Standards Institute, ANSI) сообществом ведущих IT-компаний, в число которых вошли IBM, Intel, Brocade, Cisco, EMC, Emulex, Nuova, QLogic и Sun Microsystems.

На конец сентября 2008 г. полностью согласованы все спецификации (по заявлениям ряда вендоров, прим. ред.), необходимые для производства аппаратных компонент решений на основе FCoE — HBA, сетевые коммутаторы, контроллеры в системах хранения. Первое FCoE-решение было продемонстрировано еще в мае 2008 г. В настоящее время на рынке доступны FCoE-HBA от QLogic и Emulex, а также FCoE-коммутаторы производства Cisco и Nuova. Первые поставки FCoE-решений потребителям запланированы уже с конца этого года, более массовое использование FCoE-инфраструктур — с середины 2009 г.

В ближайшей перспективе (год-полтора), по заявлениям поставщиков сетевого оборудования, стоимость FCoE-порта будет оставаться выше стоимости порта FC, но в дальнейшем она должна снизиться. Еще одна особенность сетевых FCoE-инфраструктур — продажа сетевого оборудования через всех сертифицированных партнеров соответствующих вендоров, а не только через OEM-партнеров — поставщиков FC-СХД, т.е. рынок FCoE-инфраструктур будет более открытым.

В настоящее время поддерживаемая полоса пропускания IP-сетей для FCoE-инфраструктур — 10 Гбит/с. В дальнейшем планируется поддержка скорости 40 и 100 Гбит/с (рис. 1).

Чем хороша унификация сетей?

Стандарт протокола FCoE разрабатывается комитетом INCITS T11. Спецификация FCoE-протокола позволяет осуществить прозрачную имплементацию Fibre Channel по Ethernet, сохраняя при этом также на одном физическом линке одновременную поддержку протоколов TCP/IP, UDP/IP и др. Это дает возможность эволюционной консолидации FC-доступа к данным на основе Ethernet, при той же самой задержке, защите, атрибутах управления трафиком FC при сохранении всех инвестиций в FC-инструментарий, обучение и SAN-инфраструктуру.

В современных датацентрах компании используют и Ethernet для TCP/IP-сетей, и Fibre Channel (FC) — для сетей хранения (SANs). При этом каждая из сетей решает свои задачи. Сети Ethernet преимущественно используются в случаях, когда конечным пользователям/приложениям необходимо передать относительно небольшой объем информации по локальным/глобальным сетям. SANs востребованы компаниями, когда необходимо поддерживать консолидированный блочный доступ к данным, например, для приложений, осуществляющих сетевую загрузку серверов

или работающих с большими базами данных; почтовых серверов; файловых серверов и др. Основные преимущества от развертывания SAN в следующем:

- централизованное управление, защита и администрирование ресурсов хранения;
- централизованное унифицированное предоставление файловых сервисов, например таких, как резервное копирование/восстановление, архивирование, дедупликация и др.;
- повышение эффективности использования ресурсов.

Компании, которые развивают серверную виртуализацию (как UNIX так и стандартных) для развертывания критичных приложений в своих центрах данных обычно конфигурируют серверы парой FC HBA и двумя или более сетевых интерфейсных карт (network interface cards — NICs). В ряде случаев число NIC может составлять 8 и более. Серверные FCoE-адаптеры позволяют консолидировать SAN- и Ethernet-трафик на один общий сетевой адаптер, сокращая тем самым требуемое их количество (рис. 2), а также затраты на кабели, администрирование и энергопотребление. В соответствии со спецификацией FCoE на одном физическом FCoE-порту может поддерживаться до 8 виртуальных линков.

Последний фактор в условиях дефицита энергоресурсов может быть весьма существенным. Так, при использовании

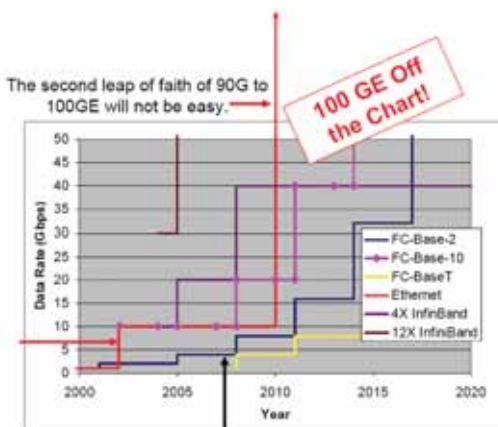


Рис. 1. Планируемое развитие по производительности различных протоколов (ист.: Brocade, 2007 г.).

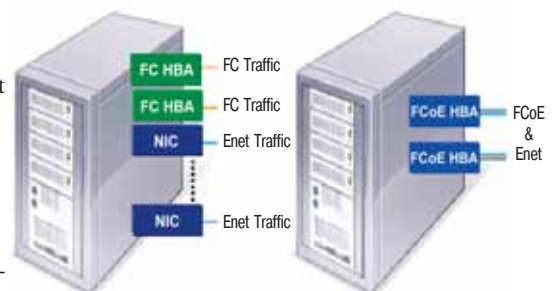


Рис. 2. FCoE-адаптеры позволяют объединить на одном порту сервера несколько FC HBA и NIC.

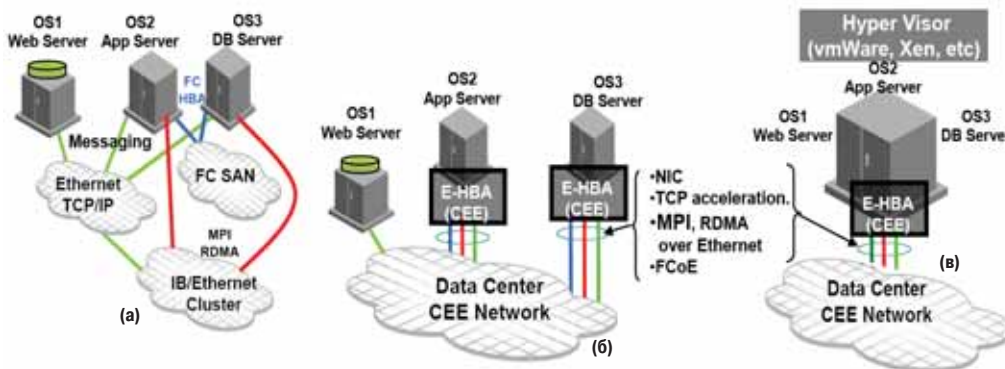


Рис. 3. Поэтапное развитие современной топологии IT-инфраструктуры (а) к унифицированной сетевой топологии (б) и далее – к полностью виртуализированной IT-инфраструктуре на базе унифицированной сетевой топологии (в) (см.: John L Hufferd, Brocade, "FCoE: Fibre Channel Over Ethernet", www.snia.org).

стандартного PCI Express разъема каждый адаптер может в среднем потреблять до 25 Вт мощности. Если число адаптеров в рэк-сервере снижается с 6 до 2, то экономия электроэнергии при этом составит порядка 100 Вт на сервер.

Помимо сокращения HBA, при переходе на FCoE-топологию уменьшается и количество сетевых коммутаторов/портов, требующихся для поддержания IT-инфраструктуры, а, соответственно, с этим – капитальные и административные затраты.

Поэтапное развитие IT-инфраструктуры на базе унифицированной сетевой топологии в представлении SNIA дано на рис. 3.

Архитектурные особенности DCE/FCoE

Концепция создания унифицированной сетевой топологии, или объединенного транспорта ЦОД одновременно для LAN и SAN требует решения следующих основных задач:

- расширение Ethernet, позволяющее использовать его в качестве транспорта без потерь пакетов данных. В стандартном Ethernet с инкапсулированным TCP/IP при перегрузках в сети используется механизм уменьшения пакета (packet drop flow control mechanism), который допускает потерю данных (для сетей хранения это является неприемлемым);
- реализация прозрачной инкапсуляции Fibre Channel фрейма в Ethernet фрейм с целью сохранения полной преемственности всего управляющего ПО для FC в среде протокола FCoE. При этом FC сервисы на FCoE фабриках также должны управляться полностью идентично, как и на FC фабриках;
- замена FC-линков в Ethernet-фабрике (без потерь пакетов) MAC-адресами. Данная функциональность позволяет выполнять манипулирование с серверами (добавление, перемещение, замена) без влияния на дизайн LAN-, SAN-сети и, соответственно, не требуя вмешательства LAN-, SAN-администраторов и перекоммутации, что снижает задержки и эксплуатационные затраты, повышает доступность и т.д. Эта технология представляет собой промышленную имплементацию виртуализации ввода-вывода серверов,

которая создает дополнительный слой между серверами и внешними сетями, такой, что LAN и SAN "видят" пул серверов, а не индивидуальные серверы. Это осуществляется за счет установления т.н. Media Access Control (MAC) адресов для всех Network Interface Controllers (NICs), World Wide Names (WWNs) для всех Host Bus Adapters (HBAs) (и, соответственно, SAN-boot-параметров). Далее эти параметры уже остаются неизменными при всех манипуляциях с серверами. Эта технология позволяет "разъединить" серверы и сети так, что при любых изменениях серверной инфраструктуры не требуется сложной хореографии изменений. Например, если сервер заменяется или добавляется новый, LAN и SAN уже "знают", как связываться с ним. Никакое ручное вмешательство не требуется. В частности, данная технология активно развивается компанией HP с 2007 г. (прим. ред.).

Расширения Ethernet

Расширения спецификации Ethernet, связанные с реализацией FCoE (или улучшения работы Ethernet в ЦОД и управления им), получили термин Data Center Ethernet (DCE), или CEE – Converged Enhanced Ethernet. Эти расширения регламентируются стандартом IEEE 802.1 и содержат 3 основные опции:

- Priority-Based Flow Control (PFC);
- Enhanced Transmission Selection (ETS);



Рис. 4. Опция PFC дает возможность приоритетного управления потоками данных (без потерь пакетов и до 8 виртуальных линий на одном физическом линке).

- DCB Capability eXchange Protocol (DCB CXP).

Опция PFC расширяет возможности Ethernet в качестве транспорта без потерь пакетов данных с возможностью поддержки на одном физическом соединении до 8 виртуальных линий (рис. 4). Предотвращение потери данных в DCE реализуется за счет отправки портом-получателем данных управляющего фрейма PAUSE для приостановки передачи данных в случае переполнения его буфера. Помимо этого, PFC обеспечивает:

- выделение ресурсов для каждой виртуальной линии (VL);
- дифференцированное поведение для

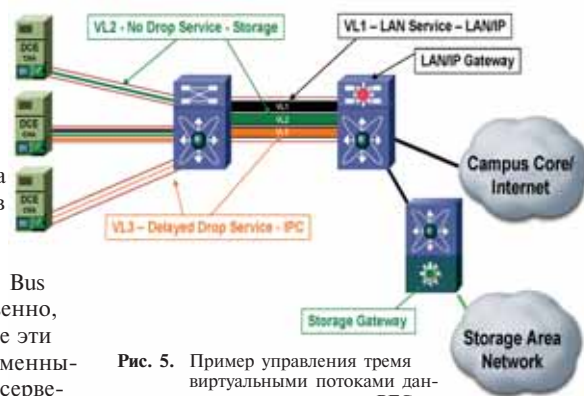


Рис. 5. Пример управления тремя виртуальными потоками данных на основе опции PFC.

каждой VL, например, возможность приостановки передачи данных для приоритетной виртуальной линии при перегрузках (VL2 – на рис. 5).

Опция ETS осуществляет поддержку гарантированной полосы пропускания для заданной виртуальной линии(ий) в соот-

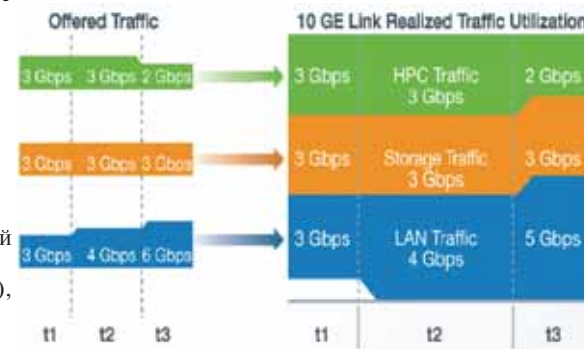


Рис. 6. Опция ETS осуществляет поддержание гарантированной полосы пропускания для заданной виртуальной линии/линий в соответствии с заданными приоритетами.

ветствии с заданными приоритетами, а также разделение полосы пропускания между виртуальными линиями (рис. 6).

Наконец, последняя из опций (ее называют еще также Data Center Bridging eXchange, прим. ред.) обеспечивает согласование настроек между сетевыми компонентами для:

- управления полосой по классам;
- управления потоком по классам;
- управления перегрузками (BCN/QCN).

Инкапсуляция Fibre Channel фрейма в Ethernet фрейм

Необходимо сразу отметить архитектурные отличия FCoE от других типов про-

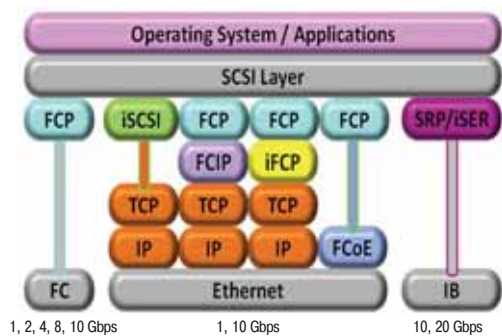


Рис. 7. Реализация FCoE в сравнении с другими типами протоколов на семиуровневой модели

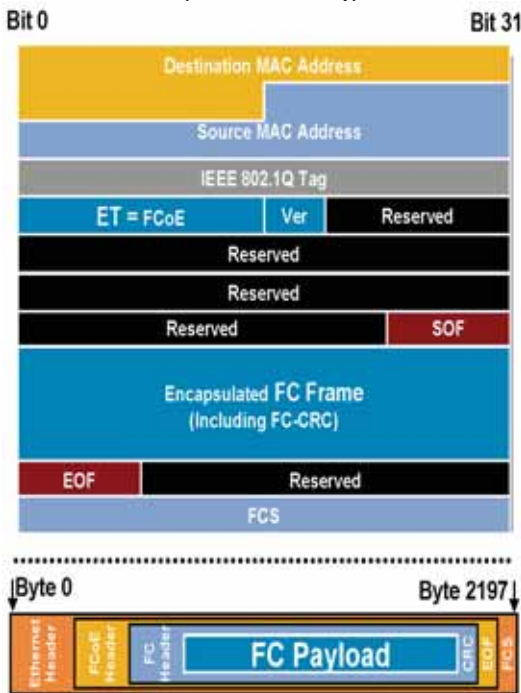


Рис. 8. Формат фрейма FCoE.

токолов, используемых в сетях хранения. FCoE существенно отличается от реализации iSCSI, и, если смотреть с потребительской точки зрения, то за счет отсутствия в FCoE (в отличие от iSCSI) уровня TCP/IP (рис. 7), производительность FCoE в максимальной степени приближена к FC. Также iSCSI не подразумевает какую-либо многопоточность на одном линке и архитектурно гораздо проще.

FCoE не заменяет FCIP или iFCP. Так же, как и iSCSI, два последние протокола используют TCP/IP, а FCoE – нет. Основное назначение FCIP и iFCP – связь между дата-центрами, в то время как FCoE ориентирован для сетевой инфраструктуры самого датацентра.

Спецификация фрейма FCoE была разработана комитетом ANSI INCITS FC T11.3 и согласована в августе 2007 г. Схема адресации ратифицирована в феврале 2008 г. и в настоящее время поддерживается 24 компаниями (рис. 8).

Стандартный Fibre Channel фрейм содержит 2112 байт передаваемых данных плюс заголовок и FCS. “Классический” Ethernet фрейм содержит 1,5 Кбайт передаваемых

данных или меньше. Соответственно, чтобы поддерживать хорошую производительность FCoE фрейм должен использовать jumbo фреймы или 2,5 Кбайт “baby jumbo”, чтобы предотвратить разбиение FC фрейма на два Ethernet фрейма. По эффективности упаковки данных FCoE фрейм всего на несколько процентов уступает фрейму FC.

FCoE фрейм составляется из слов по 32 бита. Первые 48 битов в фрейме используются для задания MAC адрес получателя, следующие 48 битов определяют MAC адрес источника. 32-битный IEEE 802.1Q тэг обеспечивает функцию, требующуюся для поддержания виртуальных LANs, позволяя множеству виртуальных сетей разделять одну физическую инфраструктуру. FCoE имеет собственный Ethernet тип, который определяется следующими 16 битами, за которыми следует поле версии из 4 бит.

Следующие 100 битов зарезервированы, за ними следуют 8-битный Start of Frame (SOF) и сам FC-фрейм (минимум 28 байт – 7 слов, максимум – 2180 байт – 545 слов), включая Fibre Channel CRC. За 8-битным концом фрейма (End-of Frame – EOF) следуют зарезервированные 24 бита.

Фрейм заканчивается 32 битами, выделенными для FCS-функции, которая обеспечивает обнаружение ошибок в составе фрейма Ethernet.

Обзор рынка

Первая реализация инфраструктуры с использованием FCoE была продемонстрирована на конференции Storage Networking World в мае 2008 г. (рис. 9). На

ней компания NetApp – первая из storage-вендоров – продемонстрировала поддержку протокола FCoE в составе своей системы хранения (см. рис. 9).

В настоящее время на рынке уже доступны 3 коммутатора с поддержкой DCE/FCoE от: Cisco – Nexus 5000, Nuova и Blade Network Technologies. Первые поставки коммутаторов Cisco серии Nexus 5000 в Россию уже начались.

Активно работают над развитием FCoE-технологии: Cisco, CA, Data Direct, Dot Hill, EMC, Emulex, FalconStor, Finisar, HP, Infotrend, Intel, LSI, MiraLink, Microsoft, NetApp, Overland Data, Promise, QLogic, Symantec и др.

В апреле с.г. Emulex стала поставлять LightPulse LP21000 семейство конвергентных сетевых адаптеров (converged network adapters – CNAs) на базе 10GbE, позволяя прозрачно интегрировать их в существующую инфраструктуру SAN с имеющимся для нее ПО. В сентябре с.г. Emulex также анонсировала, что ее семейство LightPulse Fibre Channel HBAs и FCoE CNAs адаптеры поддерживают Windows Server 2008 Hyper-V и Microsoft System Center Virtual Machine Manager 2008, а FCoE CNAs адаптеры уже сертифицированы для работы с VMware ESX с использованием N-Port ID Virtualization (NPIV) технологии.

Корпорация Intel в начале апреля с.г. объявила, что семейство ее серверных адаптеров 10 GbE будет поддерживать технологию FCoE. Intel разработала инициатор, который будет устанавливаться в серверные адаптеры Intel 10GbE на базе шины PCI Express для поддержки функций FCoE. Intel уже начала с мая с.г. массовое производство двухпортовых серверных адаптеров Intel 10 Gigabit AF DA, цена которых порядка 799 долларов США. Этот недорогой, энергосберегающий, низкопрофильный,

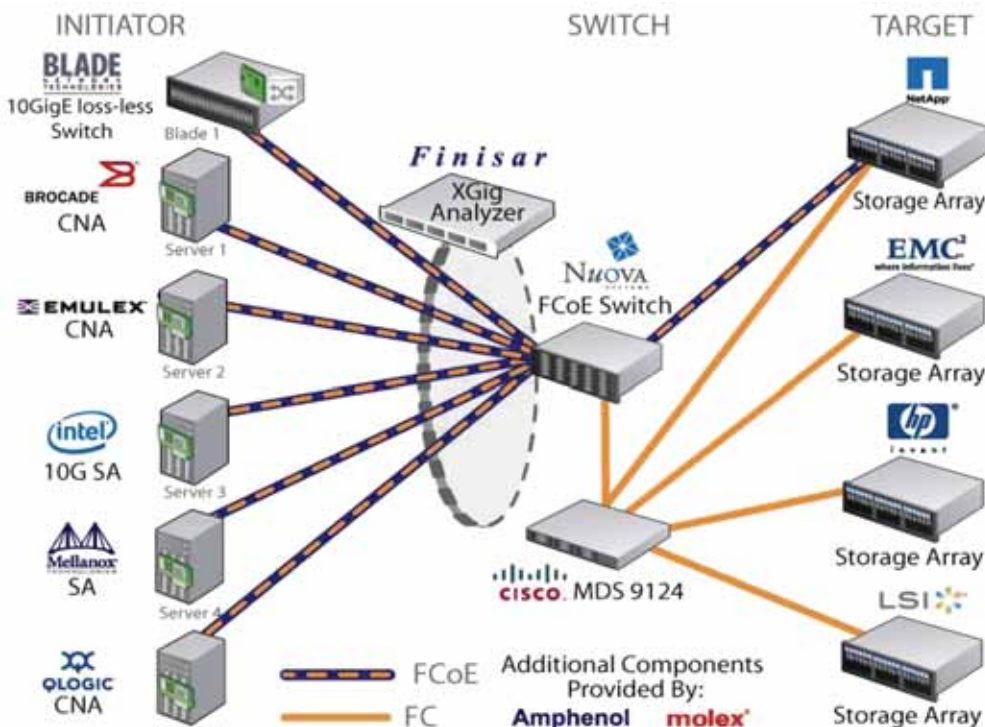


Рис. 9. Конфигурация IT-инфраструктуры, впервые продемонстрированная на конференции Storage Networking World в мае 2008 г.

монтируемый в стойку адаптер на базе медного кабеля с поддержкой протокола FCoE предназначен для центров обработки данных.

Поддержка технологии FCoE всем семейством продукции Intel 10GbE под ОС Red Hat Enterprise Linux будет реализована к июлю, под ОС Windows — до конца 2008 г.

Одним из главных препятствий для широкого распространения решений 10GbE, применение которых стимулируется необходимостью расширения полосы пропускания подсистемы ввода/вывода из-за растущего использования многоядерных процессоров и виртуальных серверов, была их высокая стоимость в расчете на один порт. Двухпортовый серверный адаптер Intel® 10 Gigabit AF DA Dual Port представляет собой модуль SFP+ Direct Attach, который четверо дешевле, чем современные двухпортовые оптоволоконные адаптеры 10GbE SR. Это новое экономичное решение обеспечивает обмен данными на расстоянии до 10 метров и может подключаться к коммутирующему порту 10GbE SFP+, как, например, в платформах Cisco Nexus 5000.

Brocade планирует поставлять собственные CEE/FCoE-решения в 2009 г., когда будут официально приняты CEE/FCoE стандарты комитетами IEEE, T11 ANSI и IETF. Среди FCoE решений Brocade:

- коммутатор FCoE Brocade 8000 – Top-of-the-Rack Switch;
- FCoE- и CEE-лезвия для директора Brocade DCX Backbone;
- конвергированные сетевые адаптеры для серверов различных типов – Converged Network Adapters (CNA).

На состоявшейся 27-29 октября 2008 г. во Франкфурте конференции SNW, Bro-



Рис. 9. Внешний вид FCoE-коммутатора – Brocade 8000 FCoE Switch, впервые продемонстрированного на конференции SNW во Франкфурте (27-29 октября 2008 г.).

cade уже продемонстрировала образец коммутатора FCoE Brocade 8000 Top-of-the-Rack Switch (рис. 9), но продукт еще не объявлен в общую доступность.

По мнению Brocade (с учетом прогноза Dell'Oro Group и Gartner Group, прим. ред.), сколько-нибудь заметную долю FCoE-решения начнут занимать на рынке не раньше 2010 г.

В сентябре 2008 г. QLogic объявила о доступности двухпортовых адаптеров QLogic QLE8042 с поддержкой FCoE. Один адаптер CNA может выполнять работу сразу хост-адаптера FC (HBA) и адаптера Ethernet (NIC). QLE8042 сохраняет все привычные принципы FC, такие, как WWN, FC-ID, маскирование LUN и зонирование, устраняя, таким образом, необходимость переподготовки персонала при переходе на новую технологию. Адаптеры QLE8042 и коммутаторы Cisco Nexus 5000 Series сертифицированных как единое решение.

В октябре с.г. компания NetApp объявила, что ее новые системы хранения будут иметь встроенную поддержку FCoE. Как будет реализована эта поддержка – аппаратно или программно – не сообщается. NetApp активно сотрудничает с Cisco, Emulex и QLogic. С ноября в портфеле ее предложений доступен новый Cisco Nexus 5020 FCoE коммутатор.

В октябре с.г. корпорация EMC объявила о сертификации и начале продаж коммутатора EMC Connectrix® NEX-5020 (Cisco Nexus 5020), а также адапте-

ров CNA от Emulex и QLogic. Все продукты сертифицированы для работы в составе решений на базе VMware.

Заключение

Насколько активно FCoE-компоненты будут интегрироваться пользователями в их ИТ-инфраструктуры, покажет время, ценовая политика производителей и ассортимент самих FCoE-решений на рынке. Однако в любом случае это не менее 1–1,5 лет, хотя полноценные FCoE ИТ-инфраструктуры можно будет уже строить, начиная с середины следующего года, и процесс, как говорится, “уже пошел” за счет следующих основных преимуществ FCoE:

- уменьшения числа сетей в ЦОД за счет унификации FC и TCP/IP на одном оборудовании и интерфейсах при полном сохранении управления FC в составе FCoE и совместимости с существующими SAN;
- уменьшения числа адаптеров HBA/NIC не менее чем в 2 раза, что, соответственно, уменьшает число кабелей и размеры системы (меньше слотов);
- снижения электропотребления серверов, повышения эффективности охлаждения (кабели не блокируют потоки воздуха);
- снижения CAPEX и OPEX.

При подготовке публикации были использованы материалы из открытых источников, а также выступления представителей компаний Brocade, Cisco, NetApp на конференции Storage Expo 2008 (г. Москва). Дополнительная информация по протоколу FCoE – www.t11.org/FCoE, www.fibrechannel.org.

Евгений Фохт,
системный инженер,
ДСИ ЛАНИТ