

Сетевые фабрики

Проектирование сетей
с резервом на будущее:
от 10 Гбит/с до 400G Гбит/с и выше



Введение

Введение	3
Новые типы архитектуры ЦОД	3
Топология сетевой фабрики — пропускная способность	4
Сетевая фабрика — физические сети	5
Топология сетевой фабрики — плотность портов коммутации	7
Топология сетевой фабрики — длина канала	8
Топология сетевой фабрики — поддержка приложений	11
Заключение	16

Введение

Широкое применение виртуализации и облачных вычислений привело к тому, что для центров обработки данных (ЦОД) потребовалась новая коммутационная архитектура, способная обеспечить меньшее время задержки и большую пропускную способность. Такая архитектура основана на коммутаторах сетевых фабрик и заметно отличается от традиционной топологии трехуровневой коммутации.

Сетевые фабрики могут быть сконфигурированы в различных вариантах: от выносов, устанавливаемых в каждой стойке (конфигурация ToR), или централизованных фабрик на уровнях горизонтальной (HDA) или промежуточной (IDA) распределительных зон до архитектуры полной связности (full-mesh). В любом случае необходимо уделить особое внимание проектированию и реализации физического уровня инфраструктуры, чтобы гарантировать легкое и эффективное масштабирование коммутационной фабрики.

В данном документе представлен обзор технологий сетевых фабрик, а также замечания по проектированию и практический взгляд на внедрение волоконно-оптических соединений с расчетом на поддержку изменений в архитектуре и более высокие скорости передачи данных по мере развития сети.

В документе также приведены практические примеры проектирования сетевой фабрики на основе претерминированных волоконно-оптических решений SYSTIMAX® InstaPATCH® 360, чтобы продемонстрировать всю важность проектирования инфраструктуры с расчетом на поддержку более высоких скоростей и развитие сети.

Новые типы архитектуры ЦОД

Технологии и архитектура ЦОД эволюционировали, чтобы не отставать от развития облачных хранилищ и вычислений. Архитектура традиционных корпоративных ЦОД сейчас находится в стадии адаптации к новым, более гибким моделям ИТ, в основе которых лежат облачные технологии. Новая архитектура предусматривает создание огромных «складов» ИТ-ресурсов (warehouse-scale), но рассчитанных на поддержку большого количества разнообразных корпоративных приложений.

Для подготовки к реализации «облачных» технологий необходимо обеспечить оптимизированный канал для прямой передачи данных между серверами на основе сетевой архитектуры «leaf-spine» (см. рис. 1). Такая архитектура позволяет приложениям с любого вычислительного устройства и всем хранилищам данных работать вместе прогнозируемым и масштабируемым образом, вне зависимости от того, где физически они расположены в ЦОД.

«Облачные сети» строятся на архитектуре, основанной на соединениях между коммутаторами доступа (leaf) и ядра (spine). Получающаяся при этом матрица соединений часто называют «сетевой фабрикой» (network fabric). Такая фабрика хорошо подходит для организации универсальных «облачных услуг», обеспечивая соединения «каждый с каждым» с прогнозируемой пропускной способностью и низким временем задержки. Сетевой фабрике присуще внутреннее резервирование благодаря множеству коммутационных ресурсов, распределенных по всему ЦОД, что повышает доступность приложений. Развертывание и масштабирование такой сети может оказаться намного менее затратным по сравнению с развертыванием и масштабированием сетей на базе традиционных огромных централизованных коммутационных платформ.

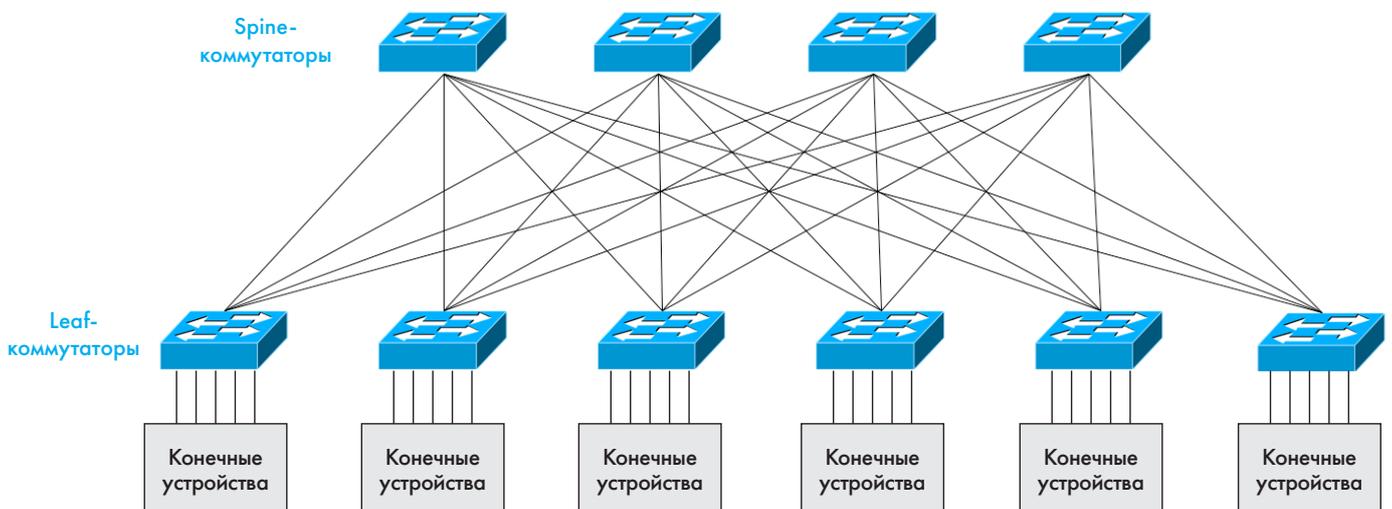


Рис. 1. Архитектура сетевой фабрики: соединения между leaf- и spine-коммутаторами обеспечивают подключение «каждый к каждому»

Топология сетевой фабрики — пропускная способность

Традиционно при проектировании сетей предполагалась агрегация трафика (например, если на уровне доступа сеть должна поддерживать 10 потоков данных по 1 Гбит/с, эти потоки объединяются в 10-гигабитный канал для передачи «наверх»). Сети «leaf-spine» работают иначе. Для правильного масштабирования таких сетевых фабрик необходимо учитывать следующие факторы:

- Скорость (или пропускная способность) соединений в сетевой фабрике;
- Количество портов, предназначенных для подключения вычислительных устройств и хранилищ данных (граничных портов);
- Общая пропускная способность сетевой фабрики, необходимая для обслуживания всех приложений ЦОД.

Скорость передачи данных в сетевой фабрике (далее, просто «скорость сетевой фабрики») определяется не общей пропускной способностью между каждой парой leaf-коммутаторов, а общей пропускной способностью между каждым leaf-коммутатором и всеми spine-коммутаторами.

На приведенной выше иллюстрации изображено четыре spine-коммутатора. Если каждый leaf-коммутатор имеет по одному 40-гигабитному соединению с каждым spine-коммутатором, то в результате получаем на 160 Гбит/с. Важно иметь в виду, что скорость соединений каждого leaf-коммутатора с любым из spine-коммутаторов должна быть одинаковой. Кроме того, ИТ-устройства (серверы и СХД) напрямую к spine-коммутаторам не подключаются.

Скорость Пропускная способность сетевой фабрики должна быть такова, чтобы обеспечивать поддержку наибольшего объема трафика, исходящего от любого leaf-коммутатора. Например, если высокоскоростные серверы подключены к 48 портам 10 Гбит/с такого коммутатора, то фабрика должна иметь возможность поддерживать 480 Гбит/с (48×10 Гбит/с).

Следующий важный момент — это общее число граничных портов, которое зависит от числа leaf-коммутаторов в сетевой фабрике. Например, если один leaf-коммутатор «несет» 24 порта по 10 Гбит/с, то каждый дополнительный подобный коммутатор увеличивает общее число портов фабрики еще на 24 порта. Добавить новый leaf-коммутатор в сеть можно лишь в том случае, если свободный порт для подключения такого коммутатора присутствует на каждом spine-коммутаторе.

Когда пропускная способность каналов между leaf- и spine-коммутаторами (например, 40 Гбит/с) больше скорости граничных портов (например, 10 Гбит/с), то соответствующая топология называется «толстым деревом» (fat tree). Если же значения одинаковы (скажем, по 10 Гбит/с), то это «тонкое дерево» (skinny tree). «Толстое дерево» имеет очевидные преимущества в плане масштабирования фабрики ЦОД. Каждый spine- и leaf-коммутатор должен иметь достаточно портов для обеспечения mesh-соединений «каждый с каждым». Количество портов и пропускная способность каждого порта предопределяют максимальный размер и пропускную способность, до которых возможно масштабирование сетевой фабрики.

В типичной сетевой фабрике, построенной по топологии «толстого дерева», могут использоваться соединения 40G Гбит/с. В нашем примере выше обозначено четыре spine-коммутатора, каждый из которых поддерживает по шесть портов 40G Гбит/с (всего 240 Гбит/с). Если предположить, что каждый leaf-коммутатор имеет 48 портов 10G, получается 288 граничных портов по 10 Гбит/сG. Однако большинство конечных устройств будет использовать двойное подключение, то есть на каждое устройство потребуется по два 10-гигабитных порта. Значит, такая конфигурация сможет поддерживать 144 конечных устройств в резервированной конфигурации.

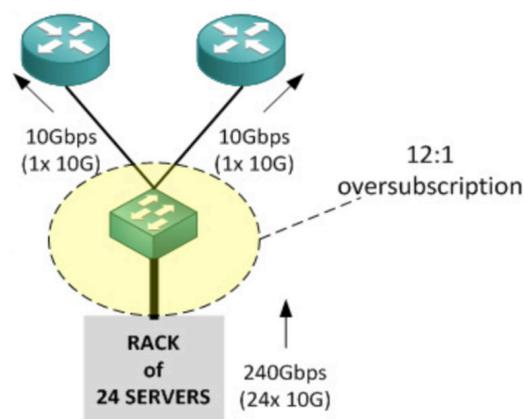


Рис. 2. Подключение с переподпиской 12:1. Суммарная скорость каналов доступа ($24 \times 10G$ Гбит/с = 240 Гбит/сG) в 12 раз превосходит суммарную пропускную способность каналов между leaf-коммутатором и spine-коммутаторами (2×10 Гбит/сG = 20 Гбит/сG)

Общую пропускную способность сетевой фабрики можно подсчитать, умножив количество граничных портов на их скорость либо количество портов spine-коммутаторов на их скорость. При отсутствии переподписки (oversubscription) полученные числа будут одинаковыми. Сетевые фабрики рассчитаны на неблокирующий режим, когда весь трафик от граничных устройств может пройти по сети без задержек и блокировок. Наличие переподписки означает, что часть трафика может быть задержана по причине нехватки ресурсов. Блокировки могут негативно влиять на работу приложений ЦОД, особенно тех, что

рассчитаны на неблокирующую среду, например, использующих протокол FCoE. Архитектура многих сетевых фабрик поддерживает отдельные сети хранения данных (на базе Fibre Channel или IP), другие же ориентированы на работу с распределенными программно-определяемыми хранилищами.

Принимая во внимание особенности работы приложений и сетевых сервисов, сетевые архитекторы оценивают суммарные требования к пропускной способности, которые и определяют характеристики сетевой фабрики. Некоторые сети спроектированы с учетом компромиссов между стоимостью и обеспечением необходимого качества обслуживания, при этом допустимый уровень блокировок или конкуренции за сетевые ресурсы закладывается уже при проектировании сети.

Коэффициент переподписки указывает на уровень конкуренции за сетевые ресурсы для оконечных устройств. На рис. 2 приведен пример, где этот коэффициент составляет 12:1.

Если коэффициент переподписки слишком велик, то страдает производительность приложений. А если его держать на слишком низком уровне, то уменьшается количество поддерживаемых сетью серверов, а значит, и количество поддерживаемых приложений. Ключевым фактором при проектировании сети является баланс между капитальными затратами и числом/производительностью приложений. Но именно этот коэффициент с большой вероятностью будет меняться с течением времени, поскольку потребности приложений растут. Производительность серверного оборудования имеет тенденцию к росту, что ведет к увеличению нагрузки на сеть.

Из описанного выше видно, что более высокая пропускная способность каналов leaf–spine поможет улучшить уровень обслуживания путем минимизации коэффициента переподписки и увеличения числа поддерживаемых фабрикой серверов. Пропускная способность этих соединений в идеале должна быть как можно выше.

При добавлении в фабрику leaf-коммутаторов количество соединений растет очень быстро. Физический уровень должен быть адаптирован для поддержки такого роста: обеспечения более высокой плотности портов и, скорости соединений, для чего необходима модульная многоволоконная структура. На рисунке 3 изображены аппаратные шнуры MPO, которые можно использовать для формирования канала Ethernet 40G при использовании трансиверов адаптеров QSFP (4 x 10 Гбит/с).



Рис. 3. Соединение 40-гигабитных коммутаторов по стандарту 40GBASE-SR4 (для простоты отдельные контакты одноволоконные коннекторы на рисунке не показаны).

Для достижения оптимальной пропускной способности сетевой фабрики оптические компоненты должны обеспечивать высокую скорость передачи данных и малые потери — для возможности перехода на большие скорости сети. Чтобы избежать последующей модернизации кабельной инфраструктуры, поддержку 40G, 100G и даже 400G следует закладывать в техническое задание с самого начала.

Оптические сетевые технологии, способные обеспечивать такие соединения, развиваются семимильными шагами. Скорости растут чрезвычайно быстро, и в некоторых случаях характеристики предлагаемых решений намного опережают те, что заложены в отраслевых стандартах. Рентабельность разнообразных решений играет главную роль в попытке угнаться за ростом требований к производительности ЦОД, а также является ключевым фактором в стремлении соблюсти баланс между капиталовложениями и рисками, связанными со снижением доступности сервисов.

Сетевая фабрика — физическая инфраструктура

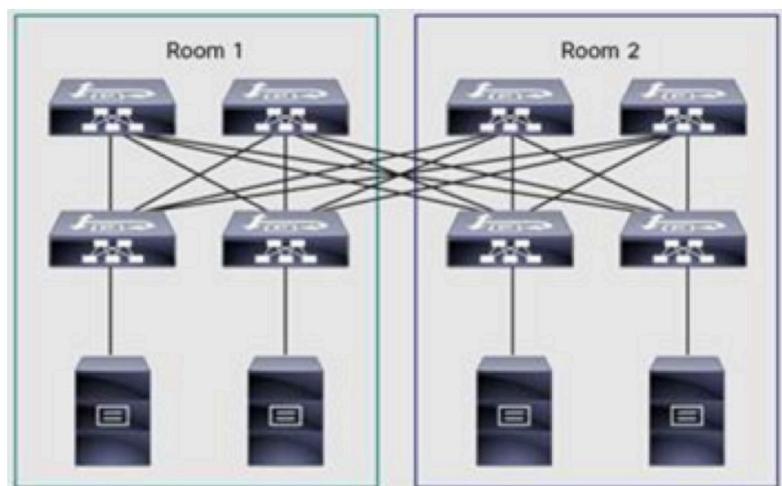


Рис. 4. Пример сети leaf–spine, охватывающей нескольких машинных залов

Сетевые фабрики схожи с традиционными трехуровневыми иерархическими сетями в том, что они должны быть масштабируемыми, управляемыми и надежными. Использование структурированных кабельных решений продолжает быть актуальным и значимым при реализации новой топологии. Кабелепроводы и технические пространства для кабельных систем остаются прежними. Сетевые фабрики могут охватывать множество машинных залов в ЦОД (см. рис. 4). Элементы сетевых фабрик также требуют обеспечения управления по выделенному каналу. И все эти требования к физической подсистеме необходимо предусмотреть в проекте (пример одного из планов размещения оборудования приведен на рис. 5).

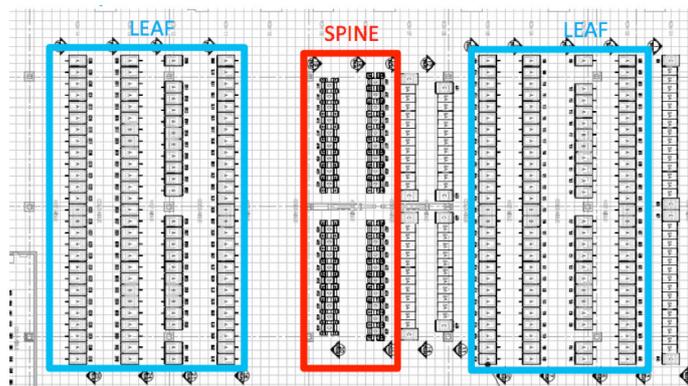


Рис. 5. План размещения стоек в сети leaf-spine

На рис. 6 изображена стандартная топология кабельной системы ЦОД* с коммутационным полем кросс-коннект в промежуточной распределительной зоне (IDA); spine-коммутаторы размещены в главной распределительной зоне (MDA), а leaf-коммутаторы — в горизонтальной распределительной зоне (HDA).

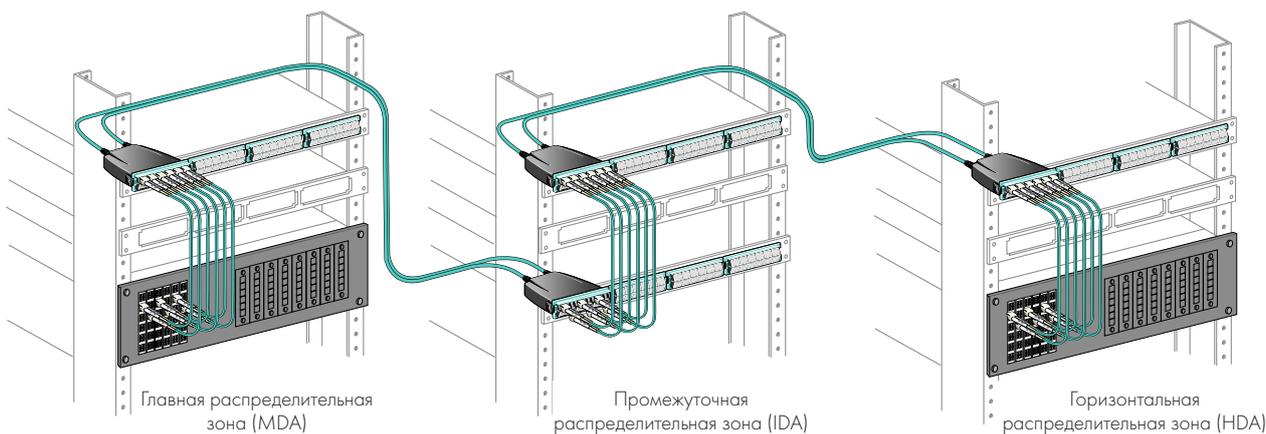


Рис. 6. Схема кабельной системы сети leaf-spine с использованием панели InstaPATCH 360 для кросс-коннекта в промежуточной распределительной зоне.

Организация сетевой фабрики на основе кроссовой коммутации заметно увеличивает гибкость и свободу маневра, упрощая потенциально сложный канал. Использование архитектуры «кросс-коннект» в ЦОД уже стало обязательным требованием в Европе, согласно стандарту CENELEC EN50600-X. Такая архитектура является предпочтительной благодаря возможности обеспечивать соединение портов по принципу «любой к любому».

* Для получения дополнительной информации по сетевым топологиям см. Стандарты для ЦОД BICSI (BICSI Data Center Standard).

Кабелепроводы и технологические пространства

Проектирование кабельной системы во многом зависит от общей топологии сети и специфики активного сетевого оборудования конкретного производителя. Некоторые производители коммутаторов предлагают собственные оптические решения, отдавая предпочтение одномодовой или многомодовой технике. Другие предпочитают крупные коммутаторы на основе шасси и зонную горизонтальную разводку. Взгляды на будущее сетей следующего поколения тоже разнятся. В большинстве случаев группа, ответственная за проектирование СКС, последней узнает о том, какое конкретно сетевое оборудование необходимо будет поддерживать к моменту ввода сети в эксплуатацию. Идеальный проект должен поддерживать все разнообразие возможных решений и вариантов развития сети, быть совместимым с новыми отраслевыми инициативами и не быть привязанным к решениям конкретного вендора.

Претерминированное оптическое решение InstaPATCH 360 (см. рис. 6) идеально подходит для создания высокоскоростной платформы СКС заводской готовности и имеет широкую область применения для обеспечения работы сетевых фабрик. Магистральные кабели (транки), распределительные модули

и коммутационные шнуры InstaPATCH 360 сконфигурированы таким образом, чтобы соответствовать требованиям к магистральному, коммутационному и вычислительному оборудованию на момент ввода в эксплуатацию, а также обеспечивать задел на будущее для соответствия требованиям завтрашнего дня.

В примере с сетевой фабрикой были рассмотрены соединения leaf-spine, которые могут быть реализованы посредством изображенного ниже многомодового транка. Решение выполнено на основе недорогого многомодового волокна, сохраняет обратную совместимость с прежними сетевыми технологиями и в будущем сможет обеспечить наращивание пропускной способности до 100 Гбит/с при условии соответствующего проектирования соединений.



Рис. 7. Соединение leaf-spine, реализованное на основе многомодового транка с промежуточной распределительной зоной.

Топология сетевой фабрики — плотность портов коммутации

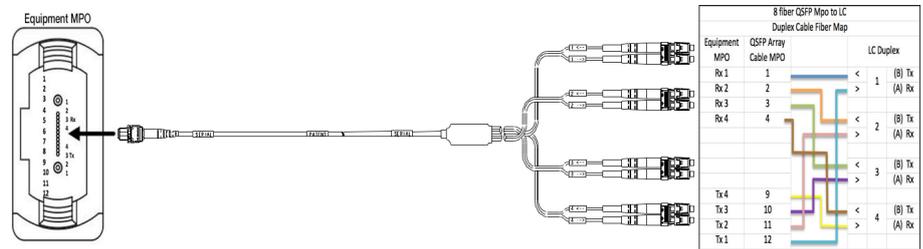


Рис. 8. Кабельная сборка QSFP MPO-LC

Несколько портов 10 Гбит/с могут быть объединены для получения соединения с большей скоростью. Стандарты IEEE предусматривают группирование четырех портов 10 Гбит/с (с помощью 12-волоконного разъема MPO) для образования соединения 40 Гбит/с. Соответствующий стандарт QSFP используется для создания соединений с большей пропускной способностью (40 Гбит/с), а также для подключения к одному порту leaf-коммутатора четырех серверов, что повышает эффективность использования ресурсов сети. Объединение четырех портов LC в одном модуле QSFP ведет к повышению плотности подключений к leaf-коммутатору примерно в четыре раза по сравнению с использованием отдельных портов, предназначенных для интерфейса SFP+.



Рис. 9. Подключение с использованием кабельных сборок MPO-LC

Производители оборудования используют MPO-соединители для обеспечения более высокой плотности подключений. Например, некоторые производители размещают на одном 24-волоконном MPO-соединителе 12 портов по 10 Гбит/сG. Волокна могут быть сгруппированы для формирования 3 соединений по 40 Гбит/сG или разделены для формирования 12 портов 10 Гбит/сG для подключения устройств. Группирование волокон в соединители более высокой плотности позволяет рационально использовать коммутационное поле на сетевых устройствах и обеспечивает удобное масштабирование.

Топология сетевой фабрики — длина канала

Центры обработки данных, как правило, занимают большие площади, где размещено множество серверов и систем хранения данных. Существуют различные стандарты, описывающие наилучший подход к проектированию кабельных сетей и помещений для кабельных систем, в том числе: ANSI/TIA-942-B, ISO/IEC 11801-5, CENELEC EN50173-5 и EN50600-X.

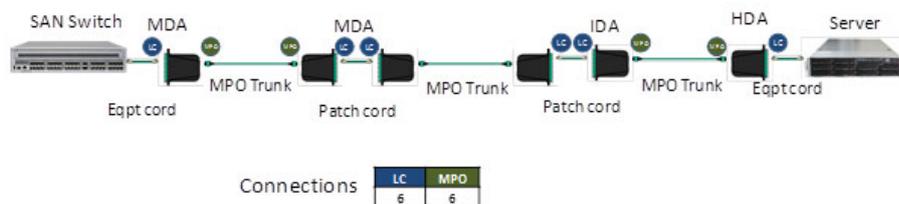
Использование СКС позволяет обеспечить необходимое масштабирование и улучшить как показатели среднего времени ремонта (MTTR), так и общий уровень доступности систем. Поэтому крайне целесообразно сохранить принципы СКС и для топологии сетевой фабрики. В сетях leaf-spine коммуникационные трассы должны быть размещены в специально отведенных коммуникационных пространствах, так же как и в сетях других топологий.

Ключевой элемент проектирования сетевых фабрик — обеспечение высокопроизводительных линий связи при разумных затратах. Многомодовая техника обычно обходится дешевле, чем одномодовая, что особенно заметно с ростом сетевых скоростей. В настоящий момент сетевому архитектору доступен большой выбор как стандартизированных, так и фирменных (нестандартных) решений с разными соотношениями производительности, цены и операционных рисков/преимуществ. Продолжают появляться новые интерфейсы передачи данных, что ведет к дальнейшему увеличению числа вариантов. Кабельные технологии должны обеспечивать необходимую в перспективе производительность сетей и проложить дорогу для проектов сетевых фабрик, обеспечивающих большую емкость и пропускную способность.

При построении сетевых фабрик очень важно уделить должное внимание проектированию сетевых подключений. Коммутационные зоны имеет смысл размещать в каждой из распределительных зон (как в приведенной ниже конфигурации). При всех различиях сетевых технологий (приложений), как правило, чем выше скорость, тем меньше допустимая длина канала СКС. Увеличение числа соединений также ослабляет сигнал и, стало быть, уменьшает допустимую длину. Довольно часто производители сетевого оборудования указывают в своих спецификациях максимальную длину канала «точка-точка». При проектировании СКС важно понимать, как эти спецификации соотносятся с практическими задачами.

Предположим, планируется реализовать подключения, показанные на рис. 10. Изначально необходимо будет обеспечить поддержку 10-гигабитного Ethernet и 8-гигабитного Fibre Channel (FC). ЦОД поделен на отдельные машинные залы, что обеспечивает его хорошую управляемость. Возникнут ли проблемы с длинами линий связи, которые требуются по проекту?

Сценарий 1: канал «сервер-SAN» — 8G FC по волокну OM4



Сценарий 2: канал «сервер-LAN» — 10G Ethernet по волокну OM4

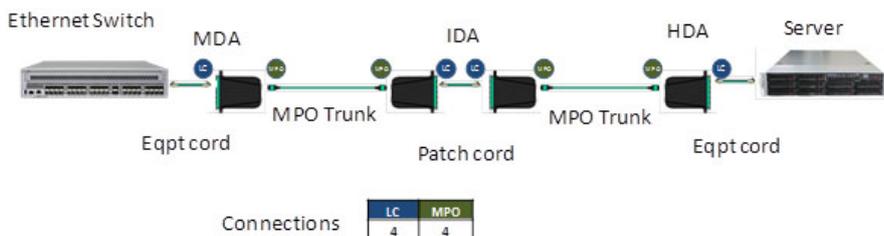


Рис. 10: Примеры подключений в ЦОД

В исходных требованиях проекта указано, что инфраструктура сети должна иметь запас по скорости каналов связи как минимум «на один шаг вперед» в целях обеспечения будущего роста пропускной способности. Производители предлагают разные варианты: на основе одномодовой и многомодовой техники, стандартизированные и на основе новых фирменных решений. Какие из будущих сетевых приложений проект будет поддерживать изначально?

Процесс поиска лучшего решения начинается с изучения различных вариантов. Будет ли обеспечена надежная работа сетевых устройств в предложенной топологии? Если доступны разные варианты, то какая стратегия поможет добиться лучшей цены и максимальной надежности? Для ответа на эти вопросы в первую очередь рассмотрим стандарты, которые описывают возможные варианты для линий передачи в сети Ethernet. Приведенная ниже таблица включает в себя как уже принятые стандарты, так и находящиеся на рассмотрении. Существуют и дополнительные варианты (50G и 200G), которые комитет IEEE802.3 сейчас изучает.

Приложение	Стандарт	Индекс IEEE	Тип волокна	Скорость	Макс. расстояние
10G Ethernet	10GBASE-SR	802.3ae	Многомодовое волокно (MMF)	10 Гбит/с	от 33 м (OM1) до 550 м (OM4)
	10GBASE-LR		Одномодовое волокно (SMF)		10 км
	10GBASE-LX4		MMF		300 м
	10GBASE-ER		SMF		40 км
	10GBASE-LRM	802.3aq	MMF		от 220 м (OM1/OM2) до 300 м (OM3)
25G Ethernet	25GBASE-SR	P802.3by	MMF	25 Гбит/с	70 м (OM3), 100 м (OM4)
40G Ethernet	40GBASE-SR4	802.3bm	MMF	40 Гбит/с	100 м (OM3), 150 м (OM4)
	40GBASE-LR4		SMF		10 км
	40GBASE-FR		SMF		2 км
	40GBASE-ER4		SMF		40 км
100G Ethernet	100GBASE-SR10	802.3bm	MMF	100 Гбит/с	100 м (OM3), 150 м (OM4)
	100GBASE-LR4		SMF		10 км
	100GBASE-SR4		SMF		70 м (OM3), 100 м (OM4)
	100GBASE-ER4		SMF		40 км
400G Ethernet	400GBASE-SR16	P802.3bs	MMF	400 Гбит/с	70 м (OM3), 100 м (OM4)
	400GBASE-DR4		SMF		500 м
	400GBASE-FR8		SMF		2 км
	400GBASE-LR8		SMF		10 км

Рис. 11: Стандарты Ethernet на базе оптоволоконна (завершенные стандарты, находящиеся в разработке на момент публикации, выделены красным цветом).

Отраслевые стандарты содержат правила проектирования, которые позволяют определить, будет ли выбранная топология СКС соответствовать требованиям архитектуры приложений, обеспечивая необходимую надежность. Следует оценить каждую из возможных топологий каналов, чтобы определить максимальную длину линии связи и максимальные потери сигнала. Каковы общие потери от всех соединений? Как комбинация длины и потерь соотносится с ограничениями, заданными стандартом приложения? Сравнивая каждый вариант с требованиями стандартов, можно определиться с решением, которое будет использовано в проекте.

Оценка потерь в линии требует понимания специфики используемых компонентов. Характеристики варьируются от производителя к производителю и даже в рамках одной партии изделий. Естественно, нас интересует наихудший показатель вносимых потерь, чтобы гарантировать, что допустимые сетевым оборудованием значения превышены не будут. Волокна с высокой пропускной способностью поддерживают намного большие длины каналов, тогда как волокно низкого качества требует коротких соединений для надежной работы. Если проект основан на стандартах и предоставленной вендором информации о характеристиках компонентов, то все расчеты соединений остаются за инженером-проектировщиком СКС.

Чтобы гарантировать соответствие проекта требованиям завтрашнего дня, необходимо заложить поддержку как минимум следующего уровня скорости в сети относительно изначально указанного. При этом нужно принять во внимание ряд факторов.

Рассмотрим максимальный (не средний и не стандартный) уровень потерь, которые каждый элемент кабельной системы внесет в проектируемый канал. Необходимо учитывать пропускную способность конкретного типа волокна — например, у OM3 она меньше, чем у OM4. Следует принимать во внимание возможность будущего использования параллельных многомодовых соединений. Наконец, надо предусмотреть влияние масштаба и размера ЦОД — какие ограничения на скорости систем следующего поколения накладывает поддержка требуемой длины каналов?

Для анализа двух приведенных выше сценариев с использованием стандартных компонентов необходимо рассматривать наихудший показатель вносимых потерь для всех компонентов в канале. В этом примере модули LC/MPO имеют показатель вносимых потерь 0,50 дБ, а у магистральных волоконно-оптических кабелей (транков) он равен 3,5 дБ/км. Длина дуплексных коммутационных шнуров всего лишь несколько метров, поэтому они не будут значительно влиять на общую величину вносимых потерь.

Исходя из этих данных, общий показатель вносимых потерь для Сценария 1 равен 3,34 дБ (см. рис. 12), что превышает максимально допустимый показатель для 8-гигабитного Fibre Channel — 2,19 дБ. Спроектированный таким образом канал, скорее всего, либо не сможет обеспечить работу оборудования, либо будет причиной большого числа битовых ошибок.

Сценарий 1: канал «сервер–SAN» — 8G FC по волокну OM4

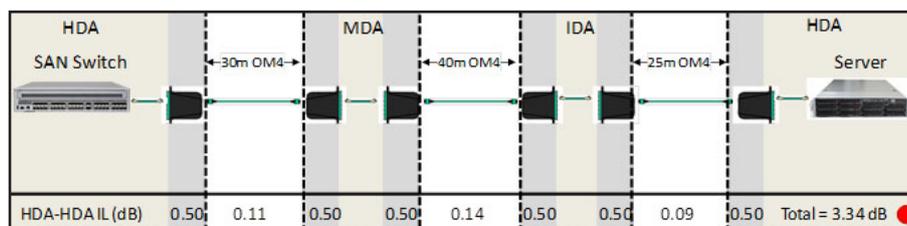


Рис. 12. Расчет бюджета потерь канала «сервер–SAN» по волокну OM4

Похожим образом анализируется и второй сценарий: 10-гигабитное подключение Ethernet «сервер–LAN» на расстоянии 130 м по волокну OM4. В этом примере потери составляет 2,39 дБ, что меньше допустимого значения для данного применения волокна OM4. Значит, такое соединение будет работать должным образом.

Стандартной практикой является добавление соединений по мере роста сети и ввода в эксплуатацию новых машинных залов ЦОД. В указанном примере организовано дополнительное соединение в зоне IDA, что увеличивает общую длину канала до 150 м и добавляет еще два модуля LC/MPO. Как показано на рис. 13, новый показатель общих вносимых потерь теперь составляет 3,53 дБ, что превышает максимально допустимое значение. Такое соединение, в том виде, в котором оно запроектировано, либо откажет, либо будет причиной большого числа битовых ошибок.

Сценарий 2: канал «сервер–LAN» — 10G Ethernet по волокну OM4

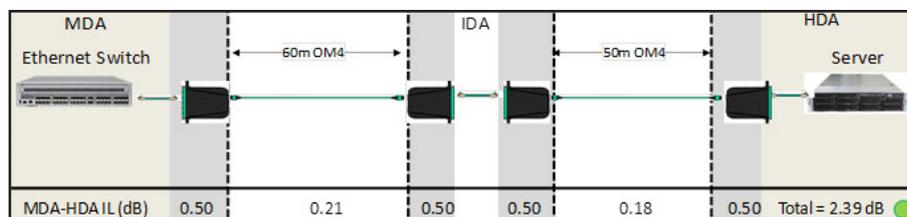


Рис. 13. Расчет бюджета потерь канала «сервер–LAN» по волокну OM4

Сценарий 2а: канал «сервер–LAN» с использованием дополнительного участка

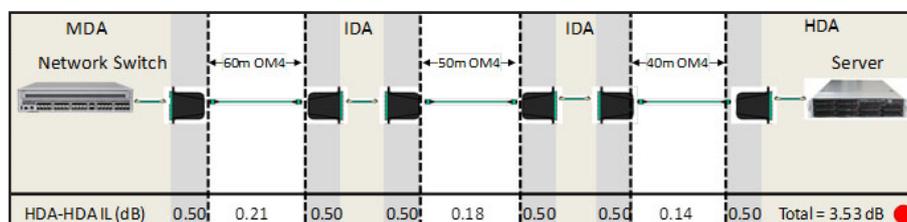


Рис. 14. Расчет бюджета потерь канала «сервер–LAN» по волокну OM4 с дополнительными соединениями

Рассмотрим переход на данной линии с 10-гигабитного Ethernet на 40-гигабитный (40GBASE-SR4). Расчет вносимых потерь приведен на рис. 15. Отметим, что общие вносимые потери уменьшаются при переходе с последовательной схемы 10GbE на 40GbE с параллельной оптикой (при этом модули LC/MPO заменяются на простые адаптерные панели MPO). Но, несмотря на снижение общих потерь, уровень потерь в линии превышает допустимый для 40GBASE-SR4 (при применении волокна OM4 он равен 1,5 дБ). Следовательно, при переходе на 40-гигабитный Ethernet такое соединение с большой вероятностью либо откажет, либо будет причиной большого числа ошибок.

Сценарий второй: канал «сервер–сеть» — переход на 40-гигабитный Ethernet по волокну OM4

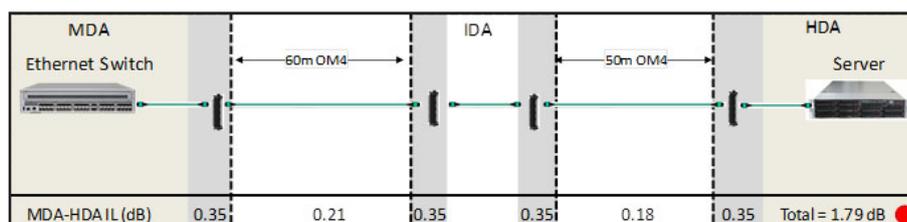
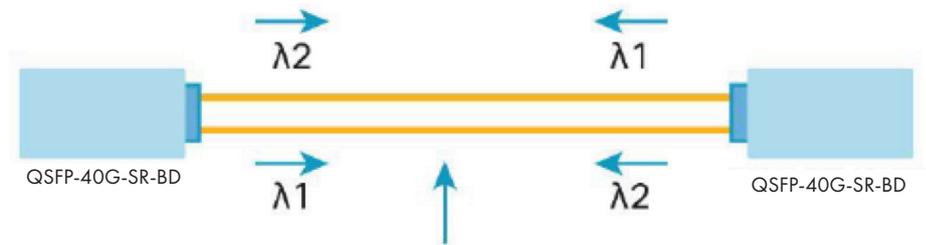


Рис. 15. Расчет бюджета потерь канала «сервер–сеть» по волокну OM4 — переход на 40 Гбит

На иллюстрациях выше приведены простые расчеты, сделанные с учетом потерь на соединителях, в кабелях и установленных в спецификациях предельных значений. Реальный проект сетевой инфраструктуры ЦОД включает в себя множество различных линий, которые должны быть просчитаны подобным образом.

К решению данной задачи можно подойти просто. Сложим общие потери, сравним их с требованиями стандартов, и получим ответ «подойдет/не подойдет». Если затем данный процесс повторить для каждого типа линии и каждого приложения (сетевой технологии), то получим общую оценку всей инфраструктуры ЦОД. Но такой процесс требует больших временных затрат. Например, при смене производителя могут измениться максимально допустимые показатели потерь каждого компонента. Более того, некоторые производители не указывают максимальные значения, а лишь стандартные показатели потерь, которые для данной задачи неприменимы.



Дуплексное многомодовое волокно с LC-соединителями на обоих концах

Рис. 16. Двусторонняя передача данных

Существуют фирменные технологии со своими предельными значениями для длин линий, но они не определены отраслевыми стандартами. Двухнаправленная схема Cisco BiDi на основе трансиверов QSFP-40G-SR-BD (рис. 16) — пример линии с использованием двух длин волн в каждом из двух волокон, общая пропускная способность такой системы — 40 Гбит/с. Для такого случая нет стандартизованных предельных значений, которые можно использовать при проектировании кабельной системы. Проект системы зависит от заявлений производителя и корректируется в зависимости от информации, предоставляемой для разных кабельных топологий.

В связи большим числом различных архитектур сетей и типов сред передачи, существует множество вариантов построения сети. Проектирование сетевых фабрик на основе параметров линий — непростая задача. Несмотря на присущую сетевым фабрикам отказоустойчивость, физические линии не должны вносить дополнительных рисков. Решение работать со стандартизованными компонентами требует от конечного пользователя оценки всех типов линий с целью определения их пригодности для выполнения конкретной задачи. Производители обычно сертифицируют лишь производительность отдельных компонентов, а не общую работу линий.

Топология сетевой фабрики — поддержка приложений

Поддержка конкретного сетевого приложения (сетевой технологии) зависит от рассмотренных ранее факторов: пропускной способности, кабельной топологии, плотности портов, длины линий, требований сетевого оборудования. Открытость решения означает готовность к поддержке различных сочетаний и комбинаций, которые имеют смысл для конкретного ЦОД. Накладывает ли привязка к решению определенного производителя существенные ограничения в будущем?

Компания CommScope разработала модульное, развертываемое по принципу plug-and-play решение InstaPATCH 360, которое поддерживает все возможные комбинации различных оптических волокон, каналов и топологий, используемых при создании сетевых фабрик. Более того, CommScope предлагает расчеты линий для поддержки конкретных приложений (технологий): внеобходимое ее еструктурой. Сократите количество ошибок при проектировании и предугадайте будущие планы развития с помощью Руководства по программе поддержки приложений CommScope (CommScope Application Assurance Guide). Компания предоставляет данные по поддерживаемой длине линий с учетом типа волокна, количества соединителей и особенностей приложения.

В приведенных ниже таблицах приведены длины линий, обеспечивающих гарантированную поддержку конкретных приложений (технологий), в том числе для описанных выше сценариев 1, 2 и 2а. Согласно данной таблице, сценарий 1 (линия 8G Fibre Channel протяженностью 95 м, выполненная на волокне IazrSPEED® 550 [OM4] с шестью MPO- и шестью LC-соединителями) поддерживается полностью. Как указано, максимально поддерживаемая длина линии составляет 150 м.

8G Fibre Channel, «последовательный приемник», 850 нм (FC-PI-4 800-MX-SN)

Поддерживаемая длина в футах (метрах)

LazrSPEED550 с LC-соединителями

Кол-во LC соединений с:	1 MPO	2 MPO	3 MPO	4 MPO	5 MPO	6 MPO
0	790 (240)	740 (225)	740 (225)	690 (210)	690 (210)	640 (195)
1	740 (225)	740 (225)	690 (210)	690 (210)	640 (195)	640 (195)
2	740 (225)	740 (225)	690 (210)	640 (195)	640 (195)	590 (180)
3	740 (225)	690 (210)	690 (210)	640 (195)	640 (195)	590 (180)
4	690 (210)	690 (210)	640 (195)	640 (195)	590 (180)	540 (165)
5	690 (210)	640 (195)	640 (195)	590 (180)	590 (180)	540 (165)
6	690 (210)	640 (195)	590 (180)	590 (180)	540 (165)	490 (150)

Табл. 17. Характеристики линии 8G Fibre Channel на основе волокна LazrSPEED550

Стандарт 40GBASE-SR4 предполагает использование MPO. Как следует из таблицы, максимальная длина линии с шестью MPO-соединениями, проложенной с использованием кабеля LazrSPEED550 OM4, составляет 140 м. Сравним эти показатели с данными для 100GBASE-SR4 — для этой технологии максимальная длина линии составляет 115 м. Значит, заложив при проектировании первой очереди максимальную длину 115 м, можно обеспечить в будущем поддержку 100 Гбит/с без изменения кабельной инфраструктуры.

40G Ethernet, параллельный интерфейс, 850 нм (40GBASE-SR4)

Поддерживаемая длина в футах (метрах)

LazrSPEED550

Кол-во MPO соединений	1 MPO	2 MPO	3 MPO	4 MPO	5 MPO	6 MPO
Расстояние в футах (м)	790 (240)	740 (225)	740 (225)	690 (210)	690 (210)	640 (195)

LazrSPEED300

Кол-во MPO соединений	1 MPO	2 MPO	3 MPO	4 MPO	5 MPO	6 MPO
Расстояние в футах (м)	460 (140)	440 (135)	430 (130)	410 (125)	390 (120)	380 (115)

Табл. 18. Характеристики линии 40GBASE-SR4 на основе волокна LazrSPEED

100G Ethernet, параллельный 4-канальный интерфейс, 850 нм (100GBASE-SR4)

Поддерживаемая длина в футах (метрах)

LazrSPEED550 WideBand и LazrSPEED550

Кол-во MPO соединений	1 MPO	2 MPO	3 MPO	4 MPO	5 MPO	6 MPO
Расстояние в футах (м)	390 (120)	390 (120)	370 (114)	370 (114)	350 (108)	350 (108)

LazrSPEED300

Кол-во MPO соединений	1 MPO	2 MPO	3 MPO	4 MPO	5 MPO	6 MPO
Расстояние в футах (м)	280 (85)	280 (85)	280 (85)	260 (80)	260 (80)	250 (75)

Табл. 19. Характеристики линии 100GBASE-SR4 на основе волокна LazrSPEED

Программа поддержки приложений CommScope (CommScope Application Assurance) также распространяется на нестандартные решения. В таблицах ниже показаны данные для систем Cisco BiDi 40G. Также ниже приведены данные для решений CSR4 с увеличенной длиной каналов. Проектировщик может сравнить максимальные длины линий для различных сетевых технологий при использовании волокна LazrSPEED550 OM4: например, 150 м — для Cisco BiDi и 420 м — для стандартной системы 40GBASE-SR4.

40G Ethernet, Cisco BiDi (Трансиверы QSFP-40G-SR-BD)

Поддерживаемая длина в футах (метрах)

LazrSPEED550 WideBand и LazrSPEED550 с LC-соединителями

Кол-во LC соединений с:	1 MPO	2 MPO	3 MPO	4 MPO	5 MPO	6 MPO
0	490 (150)	490 (150)	490 (150)	480 (145)	480 (145)	460 (140)
1	490 (150)	490 (150)	490 (150)	480 (145)	460 (140)	460 (140)
2	490 (150)	490 (150)	480 (145)	480 (145)	460 (140)	440 (135)
3	490 (150)	480 (145)	480 (145)	460 (140)	460 (140)	440 (135)
4	490 (150)	480 (145)	460 (140)	460 (140)	440 (135)	430 (130)
5	480 (145)	460 (140)	460 (140)	440 (135)	440 (135)	430 (130)
6	480 (145)	460 (140)	440 (135)	440 (135)	430 (130)	410 (125)

40G Ethernet, 850-нм параллельный интерфейс Cisco с увеличенной дальностью (Трансиверы QSFP-40G-CSR4)

Поддерживаемая длина в футах (метрах)

LazrSPEED550

Кол-во MPO соединений *	1 MPO	2 MPO	3 MPO	4 MPO	5 MPO	6 MPO
Расстояние в футах (м)	1380 (420)	1310 (400)	1310 (400)	1310 (400)	1310 (400)	1310 (400)

LazrSPEED300

Кол-во MPO соединений *	1 MPO	2 MPO	3 MPO	4 MPO	5 MPO	6 MPO
Расстояние в футах (м)	980 (300)	980 (300)	980 (300)	980 (300)	980 (300)	980 (300)

LazrSPEED150

Кол-во MPO соединений *	1 MPO	2 MPO	3 MPO	4 MPO	5 MPO	6 MPO
Расстояние в футах (м)	490 (150)	490 (150)	490 (150)	490 (150)	490 (150)	490 (150)

*Количество соединений без учета соединений с активным устройством на каждом конце канала

Табл. 20. Характеристики линий для нестандартных систем Cisco

Анализ всех четырех сценариев позволяет сравнить, чем гарантированная поддержка приложений CommScope отличается от стандартного компонентно-ориентированного метода с «ручным» анализом. При компонентно-ориентированном проектировании лишь один из четырех сценариев будет отвечать проектным требованиям и заложенному бюджету потерь. При использовании системы InstaPATCH 360 от CommScope каждый из четырех сценариев будет соответствовать проектным условиям.

Сценарий	Приложение	Общая длина	LC	MPO	Поддержка стандартными компонентами	Поддержка CommScope
1	8G FC	95 м	6	6	Нет	Да
2	10G Ethernet	110 м	4	4	Да	Да
2a	10G Ethernet	150 м	6	6	Нет	Да
2b	40G Ethernet	110 м	0	4	Нет	Да

Табл. 21. Сравнение сценариев

Линии связи для сетевых фабрик — выбор вариантов Ethernet

Одномодовые или многомодовые, параллельные или дуплексные — правильный выбор кабельной системы для ЦОД будет зависеть от его размеров, темпов развития сервисов, бюджета на оборудование, специфики вендорских технологий и многого другого. Проект СКС должен вписываться в архитектуру сети, топологию и планы развития.

Рассмотрим два варианта линий 40G Ethernet (выше):

- Стандарт 40GBASE-SR4, четыре канала 10G по восьми волокнам
- Фирменная дуплексная система BiDi, мультиплексирование двух каналов по 20G в одной паре волокон

Указанные решения поддерживают разную длину линии (см. таблицы). 40GBASE-SR4 работает на расстоянии до 420 м при использовании волокна LaserSPEED550 OM4 в сравнении со 150 м у BiDi. Расстояния и топологии, указанные в таблицах, основаны на использовании претерминированного волоконно-оптического решения Commscope InstaPATCH 360. Этот пример показывает, как можно легко спроектировать и сравнить два варианта: на основе нестандартных и стандартных трансиверов.

Система InstaPATCH 360 поддерживает работу трансиверов 40GBASE-SR4 при использовании шести соединений на 140 м (требования стандарта — 125 м), что обеспечивает большую длину канала и гибкость топологии для проектов сетевых фабрик. Если говорить о 100GBASE-SR4, то система поддерживает длину линии 108 м при шести соединениях.

Линии связи для сетевых фабрик — дальнейшие шаги

Выше были рассмотрены преимущества соединений с большой пропускной способностью (от 40G и выше), которые обеспечивают более высокую производительность сети, а значит и поддержку большего числа серверов и устройств хранения данных. И хотя в наши дни 40-гигабитные соединения являются экономичными и эффективными решениями для большинства проектов сетевых фабрик, это лишь вопрос времени, когда скорости портов доступа у серверов вырастут до 25 Гбит/с, а может быть и до 50 Гбит/с. В ближайшие годы соединения в сетевых фабриках необходимо будет нарастить до 100 Гбит/с или даже до 400 Гбит/с.

Если смотреть в будущее, есть разные решения для формирования более скоростных соединений. Одни производители предпочитают одномодовые решения. Другие — многомодовые. Для каждого из этих средств связи потенциально есть дуплексные или параллельные варианты. Показатели относительной стоимости меняется, но некоторые привычные взаимосвязи пока еще в силе: многомодовые системы продолжают требовать меньших капитальных затрат, и, по большей части, продолжают быть более простыми в поддержке и эксплуатации, чем одномодовые системы.

Линии связи Инфраструктура передачи данных для сетевой фабрики — новые виды соединений

Многомодовые кабели OM3 и OM4 поддерживают скорость линии 40G Гбит/с, обеспечивая дальность и широкие топологические возможности, которые соответствуют требованиям практически любых ЦОД, за исключением гигантских мега-ЦОД. Если говорить о стремлении достичь скоростей 100 Гбит/с и выше, то одним из самых многообещающих методов увеличения пропускной способности сети является добавление каналов связи в каждую дуплексную пару многомодовых волокон. Дорогостоящая технология спектрального уплотнения (wavelength division multiplexing, WDM) уже давно доступна для одномодовых вариантов. Новейшие коротковолновые трансиверы WDM (shortwave wavelength division multiplexing, SWDM) объединяют четыре канала в паре волокон, предлагая в четыре раза большую пропускную способность многомодового кабеля. Это более экономичное решение сочетает в себе простоту установки и эксплуатации многомодовых линий с пропускной способностью, масштабируемой для обеспечения требуемого роста сетевых фабрик.



Рис. 22. Спектральное уплотнение с использованием широкополосного многомодового волокна (WBMMF)

Для поддержки технологии SWDM компания CommScope совместно с другими членами Альянса SWDM разработала новое «широкополосное» многомодовое волокно WBMMF (Wideband Multimode Fiber). Оно призвано повысить пропускную способность многомодовых линий благодаря увеличению числа каналов связи на волокно при большей дальности. Волокно WBMMF будет использоваться для более высоких скоростей передачи данных, таких как 100 Гбит/с и 400 Гбит/с, при уменьшении числа волокон, необходимых для поддержки будущих высокопроизводительных сетевых фабрик.

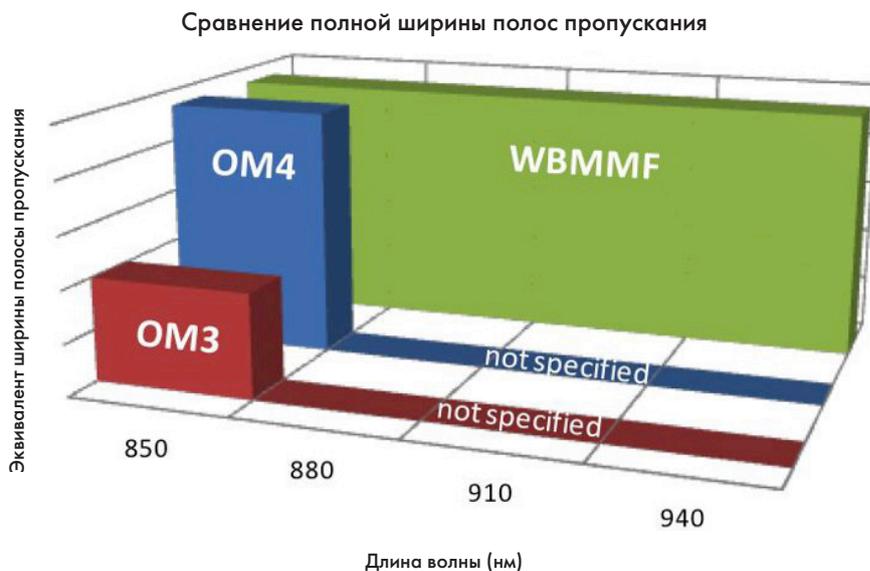


Рис. 23. Сравнение полной ширины полос пропускания волокон OM3, OM4 и WBMMF

Волокно WBMMF обладает обратной совместимостью с волокнами OM3 и OM4. Работа по стандартизации волокна WBMMF ведется в техническом комитете TR-42 Ассоциации TIA, а публикация стандарта ожидается в 2016 г. Компания CommScope выпустила кабельное решение LazrSPEED550 Wideband для всех компонентов платформы InstaPATCH, благодаря которому линии связи нового поколения стали доступны уже сегодня для создания экономичных и высокопроизводительных сетей.

Заключение

В центрах обработки данных активно внедряют новые системы на основе сетевых фабрик, призванные обеспечить эффективную поддержку облачных вычислений и систем хранения данных. Плотность размещения элементов кабельных систем ЦОД растет для поддержания соединений «каждый с каждым» с малым временем задержки, которые, как правило, нужны для распределенных облачных приложений.

Проектирование линий с высокой пропускной способностью — непростая задача, особенно при увеличении числа подключений и скорости передачи. Для обеспечения наибольшей производительности ЦОД необходимо максимально использовать возможности существующих кабельных технологий. Волоконно-оптические системы, в том числе с широкополосным многомодовым волокном WBMMF постоянно развиваются, чтобы выйти на следующую ступень развития пропускной способности и физической плотности, которые будут идеально подходить для архитектуры сетевых фабрик. Одномодовое волокно обеспечивает большую допустимую длину каналов.

Руководства по проектированию CommScope позволяют гарантировать надежность и высокую производительность сетей, а также их соответствие текущим и будущим требованиям к пропускной способности. Системы InstaPATCH 360 обеспечивают большую допустимую длину высокопроизводительных линий, свободу выбора топологии сети, масштабируемость до сверхкрупных и сложных комплексов и гарантированную производительность приложений как для стандартизированных, так и для фирменных систем.

Предлагаемые решения делают сложные сетевые фабрики легкими в проектировании, внедрении и управлении. Претерминированные высокопроизводительные системы обеспечивают поддержку сетевых технологий следующего поколения, а также новейших дуплексных и многоволоконных модульных приложений, тем самым сокращая временные и денежные затраты на управление внедрением.



www.commscope.com

Для получения дополнительной информации посетите наш веб-сайт или свяжитесь с местным представительством CommScope.

© 2016 CommScope, Inc. Все права сохранены.

Все торговые марки, помеченные ® или ™, являются зарегистрированными торговыми знаками или торговыми знаками CommScope, Inc, соответственно.

Настоящий документ предоставлен исключительно в информационных целях и не предназначен для внесения каких-либо изменений или дополнений в любые спецификации или гарантии, относящиеся к продуктам или услугам CommScope. Компания CommScope сертифицирована согласно ISO 9001, TL 9000 и ISO 14001.

TR-110117.1-RU (04/16)