

WHITE PAPER

#14

01/2024

НАПРАВЛЕНИЯ РАЗВИТИЯ СКС для ЦОД



Eurolan 2002—2024

Направления развития СКС для ЦОД

Жизнь подавляющего большинства современных людей буквально пронизана информационными технологиями. Высокая практическая востребованность той поддержки, которую обеспечивают средства вычислительной техники, обусловлена как значимым увеличением эффективности профессиональной деятельности, так и существенным упрощением решения ряда бытовых вопросов. Физически услуга данной разновидности оказывается соответствующей вычислительной инфраструктурой, функции опорных элементов которой в реалиях сегодняшнего дня берут на себя центры обработки данных (ЦОД).

Вычислительная инфраструктура машинного зала как ключевого компонента самого ЦОД, в свою очередь, представляет собой типичную сложную техническую систему, основной функцией которой является максимально качественная обработка поступающего пользовательского запроса. Выполнение соответствующих вычислительных процедур осуществляют многочисленные виртуальные и физические серверы. Они находятся в пространстве машинного зала и связаны между собой каналами внутриобъектовой связи. Сама инфраструктура машинного зала ЦОД при этом реализуется на основе проверенной временем модели взаимодействия открытых систем – OSI.

Отметим, что отдельные компоненты ЦОД как системного объекта в процессе выполнения своих основных функций не взаимодействуют с человеком как с потребителем его ресурсов, ограничивая это взаимодействие только отправкой ответа на поступающий запрос. Данная особенность позволяет при дальнейшем анализе отказаться от известных предельно жестких с точки зрения современной электроники скоростных ограничений системы «человек – машина», определяемых первым звеном этой пары. Благодаря этому открываются перспективы значимого наращивания качества функционирования ЦОД привлечением большего объема данных и увеличением глубины их переработки. Все это в комплексе с удешевлением современной электроники предполагает высокую информационную пропускную способность тех многочисленных каналов внутриобъектовой связи, которые организуются в машинном зале ЦОД.

Особенности информационной проводки машинного зала ЦОД

Физический уровень каналов связи машинного зала ЦОД формируется с привлечением ресурсов структурированной кабельной системы (СКС). Продолжительность эксплуатации СКС при условии правильного исходного выбора проектных решений в разы превышает аналогичный параметр для активного сетевого оборудования. Поэтому, несмотря на некоторое непринципиальное замедление темпов роста скоростей передачи данных в машинном зале ЦОД, которые демонстрирует отрасль на текущем этапе своего развития и которые представлены на Рис. 1, в информационную проводку целесообразно изначально закладывать 3–4-кратный резерв продолжительности ее нормальной эксплуатации.

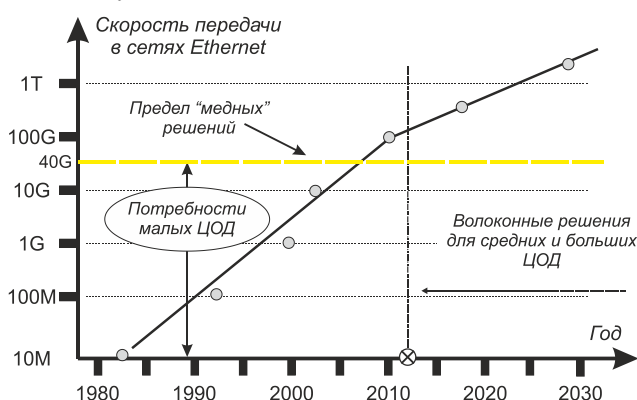


Рис. 1. Динамика роста скоростей передачи данных в сетях Ethernet как основной технологии передачи данных каналов внутриобъектовой связи машинного зала ЦОД

Ситуация с выбором максимальных скоростей передачи, поддерживаемых информационной проводкой на момент ее сдачи в текущую эксплуатацию, осложняется тем обстоятельством, что проводка крайне плохо поддается модернизации. Любой кабельный тракт, обеспечивающий доставку информации от разъема до разъема активного сетевого оборудования, представляет собой последовательное соединение отдельных компонентов и, соответственно, тем или иным образом попадает под действие известного ограничения слабого звена.

В отношении разъемов устранение проблемы скоростного «бутылочного горлышка» не представляет больших проблем из-за точечного характера этого компонента и легко решается применением модульно-кассетных решений. Шнуры из-за своей небольшой длины и нахождения за пределами стационарной линии также легко меняются при возникновении такой необходимости. Ситуация, однако, радикально меняется на противоположную в случае линейных кабелей. Это обусловлено тем, что любые действия, связанные с их заменой, серьезно нарушают нормальное функционирование системы воздушного охлаждения как основного средства утилизации тех больших объемов тепла, которые выделяются в процессе функционирования серверов.

Обозначенную проблему дополнительно усугубляют также следующие два обстоятельства:

- серверы эксплуатируются в режиме высокого уровня разгона, а заметное ограничение продолжительности нормальной эксплуатации данной техники из-за связанного с этим повышенного износа не считается критически важным по причине ее быстрого морального устаревания;
- в отличие от телефонной сети в ЦОД отсутствуют явно выраженные суточные пики нагрузок и простоев, а сам объект штатно эксплуатируется в режиме $24 \times 7 \times 365$, что не позволяет выделить «окна» для выполнения довольно продолжительных по времени работ по модернизации линейной части информационной кабельной системы.

Косвенным подтверждением справедливости последнего положения является высокая популярность решений, обеспечивающих горячую замену отдельных блоков различного активного оборудования.

В таких условиях резко возрастает важность тех проектных решений, которые должны снизить практически до нуля вероятность морального устаревания СКС машинного зала ЦОД в горизонте по меньшей мере ближайших 10...15 лет, т. е. до 2035 года.

При этом базовые требования к СКС относительно ее перспектив остаются довольно простыми. Они заключаются в ее готовности к поддержке скоростей передачи вплоть до 1,6 Тбит/с, которые будут массово востребованы к середине 30-х гг. текущего столетия. Данное значение получено линейной экстраполяцией имеющейся динамики роста этого параметра на указанный период, что демонстрирует Рис. 1.

Приведенная оценка достаточно достоверна. Это связано с тем, что наиболее скоростные сетевые интерфейсы эксплуатируются именно в ЦОД, т. е. прорыва за счет заимствования технических решений из смежных областей ожидать не приходится.

Кроме того, необходимо принимать во внимание топологически статический характер аппаратной части, на основе которой строится канальный уровень информационной инфраструктуры машинного зала ЦОД. Это определяется:

- уже отмеченным ранее довольно небольшим сроком штатной эксплуатации активного сетевого оборудования;
- широким использованием технологии Port Breakout как средства бесшовного перехода на следующую скоростную ступень и наращивания эксплуатационной надежности информационной системы в целом, Рис. 2.

Таким образом, в основу нормирования параметров кабельной системы целесообразно положить 2-коннекторную модель кабельного тракта.

Отдельно укажем на то, что обеспечение возможности формирования составных кабельных трактов находится в зоне ответственности производителя кабельной системы. Он обеспечивает реализацию такой возможности соответствующей элементной базой с необходимыми запасами, правилами проектирования и иными технологическими мероприятиями.

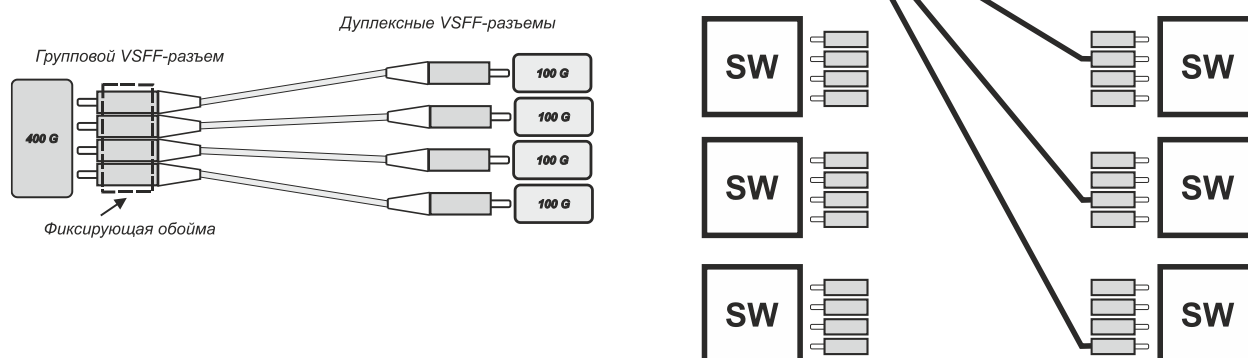


Рис. 2. Некоторые типовые структуры организации каналов связи в машинном зале ЦОД. Слева – схема Port Breakout, справа – отказоустойчивая структура на схеме соединения «каждый с каждым» (условно показаны только соединения одного коммутатора SW с остальными)

Возможности построения информационной проводки на базе кабелей из витых пар

Профильные стандарты СКС (ANSI/TIA-942B, ISO/IEC 11801-5:2017 и ГОСТ Р 59486-2021) допускают реализацию информационной кабельной системы машинного зала ЦОД как на кабелях из витых пар, так и на волоконно-оптических кабелях.

При этом, однако, следует иметь в виду, что современные электропроводные решения изначально не рассчитаны на скорости свыше 40 Гбит/с. Одновременно из данных Рис. 1 непосредственно вытекает, что 40-гигабитное быстродействие перестало соответствовать требованиям практики уже в конце первого десятилетия текущего века. Проблема дополнительно усугубляется отсутствием на рынке массового предложения серийных сетевых интерфейсов указанной разновидности, что косвенно подтверждает малую перспективность электропроводных решений в рассматриваемой области.

Сложившееся положение дел полностью определяет область применения кабелей из витых пар и соответствующего им коммутационного оборудования в машинном зале ЦОД:

- небольшие слабо нагруженные объекты максимум с пятью стойками;
- дополнительные внутрирядные связи;
- вспомогательные цепи управления, в т. ч. инженерным обеспечением, мониторинга параметров окружающей среды и аналогичные им.

Во всех остальных случаях на физическом уровне должна использоваться волоконно-оптическая элементная база.

Особенности кодирования линейного сигнала при передаче на высоких скоростях

Одна из проблем создания каналов связи со скоростями передачи в десятки Гбит/с и выше состоит в том, что тактовые частоты современной электроники ограничены значениями порядка 35...40 ГГц, т. е. обычная бинарная передача не может осуществляться со скоростями свыше 35...40 Гбит/с. Для преодоления этого барьера используется ряд подходов, основанных на глубокой предварительной обработке передаваемого массива данных. Один из приемов этой группы – многоуровневая передача. Ее суть заключается в том, что за один такт линейного сигнала передается несколько бит исходного сообщения. Технически это реализуется простой амплитудно-импульсной модуляцией или же более сложной амплитудно-фазовой модуляцией.

Основную идею такого подхода на примере амплитудно-импульсного кода РАМ4 демонстрирует Рис. 3.

Обращение к сложным линейным кодам дает определенный положительный эффект, но возможности этого приема существенно ограничены. Сказывается то, что в случае, например, амплитудно-импульсной модуляции информационная скорость передачи растет с темпом $\lg(n)/\lg 2$, где n – количество уровней сигнала в канале связи, а помехоустойчивость сигнала падает в n раз, т. е. с

более высоким темпом. Таким образом, реализация многоуровневой передачи возможна только при наличии запаса по соотношению «сигнал / шум» бинарных интерфейсов.

При работе на экономически выгодной длине волны первого окна прозрачности 850...950 нм запасы по отношению сигнала к шуму присутствуют изначально. Сказывается то, что из-за высокого уровня хроматической дисперсии главным фактором, ограничивающим предельную дальность связи, выступает т. н. дисперсионный штраф. Поэтому любые нововведения, снижающие тактовую частоту линейного сигнала, положительно сказываются на дальности действия сетевого интерфейса.

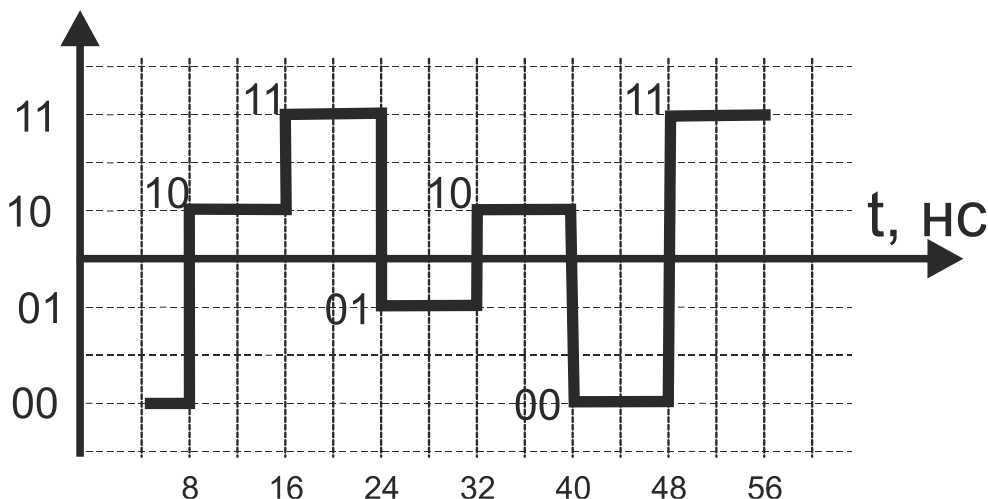


Рис. 3. Схема многоуровневой амплитудно-импульсной модуляции линейного сигнала (на примере линейного кода РМ4)

Необходимость параллельной передачи и ее предпочтительная схема

Фактически единственным эффективным средством достижения в канале связи скорости 100 Гбит/с и выше становится разбиение на передающем конце потока данных на n обычно одинаковых составляющих с последующей передачей каждой из них по отдельному независимому субканалу. При таком подходе быстродействие каждого субканала можно снизить в n раз. На приемном конце осуществляется обратная операция, т. е. восстановление исходного информационного потока. Решения данной разновидности обозначаются как параллельная передача.

Количество субканалов из соображений упрощения схемотехнических решений электроники приемника и передатчика выбирается кратным степени 2, т. е. равно 2, 4 и т. д.

Параллельная передача представлена в серийной технике двумя основными разновидностями: физической и спектральной. Отличаются они друг от друга тем, что в случае физической параллельной передачи под каждый субканал отводится отдельное волокно (т. н. пространственное уплотнение), а сам обмен данными между интерфейсами осуществляется на одной длине волны. В случае спектрального мультиплексирования каждый субканал организуется с привлечением отдельной оптической несущей со своей уникальной длиной волны. Совокупность канальных оптических несущих объединяется в оптическом мультиплексоре в групповой сигнал и передается по одному волокну с последующим разделением на приемном конце с помощью оптического демультиплексора.

При пространственном мультиплексировании число задействованных для этого волокон может достигать 32 (по 16 для направлений приема и передачи), хотя на практике наибольшее распространение получила схема с 8 или 12 волокнами, которая может быть реализована на разъемах МРО/МТР.

Спектральное мультиплексирование в экономически выгодном спектральном диапазоне 850 нм представлено 4-волновой дуплексной схемой SWDM (одно волокно для передачи, второе – для приема) с расстоянием между каналами 30 нм. Известно также промежуточное решение BiDi, которое также предполагает 2-волновое спектральное мультиплексирование, но с увеличенным вдвое расстоянием между оптическими несущими 850 и 910 нм.

Идею указанных разновидностей параллельной передачи в «чистой» форме их реализации схематически демонстрирует Рис. 4.

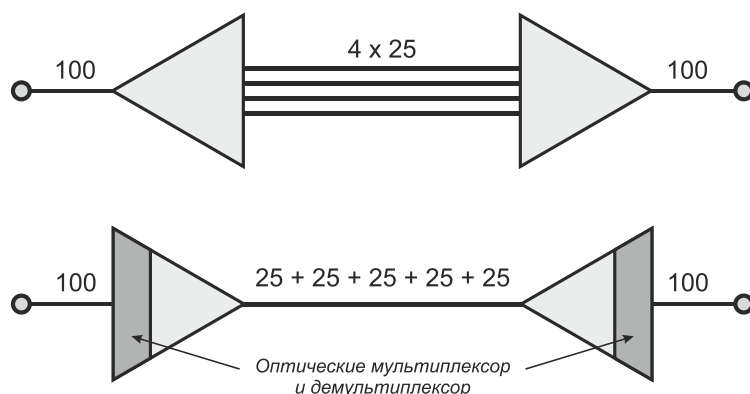


Рис. 4. Схемы реализации параллельной передачи на примере скорости 100 Гбит/с. Вверху – физическая параллельная передача (пространственное мультиплексирование), внизу – спектральное мультиплексирование

На Рис. 5 приводятся задаваемые нормативными документами рабочие длины волн при выборе той или иной схемы реализации интерфейса.

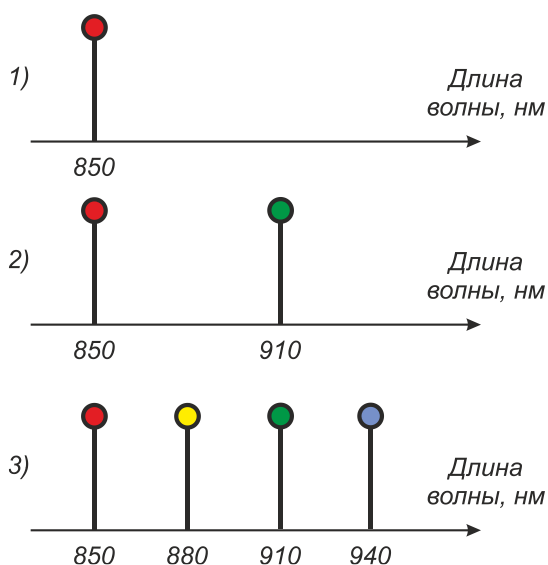


Рис. 5. Сетки длин волн оптических несущих при различных схемах параллельной передачи: 1) традиционное пространственное мультиплексирование; 2) схема BiDi; 3) схема 4-канального коротковолнового спектрального мультиплексирования

При этом многоуровневые методы кодирования и параллельная передача в любой форме ее реализации взаимно независимы друг от друга. Обеспечиваемая этим свойством инвариантность позволяет применять оба приема совместно для наращивания скорости передачи в канале связи. Выбор того или иного решения определяется конкретными местными условиями.

В качестве иллюстрации указанной особенности техники сошлемся на спектрально-волоконную диаграмму (Рис. 6), на которой показаны основные возможности достижения заданной скорости в канале связи. Как видим, известная с середины второго десятилетия текущего века технология BiDi занимает промежуточное положение между спектральным и пространственным мультиплексированием. Ее довольно высокая практическая популярность есть прямое следствие характерного для современных условий некоторого замедления темпов роста скоростей в каналах межмашинной связи. BiDi как технология выгодна тем, что при переходе на следующую скоростную ступень внедрение соответствующих интерфейсов в ряде случаев дает возможность не менять конфигурацию кабельных трактов.

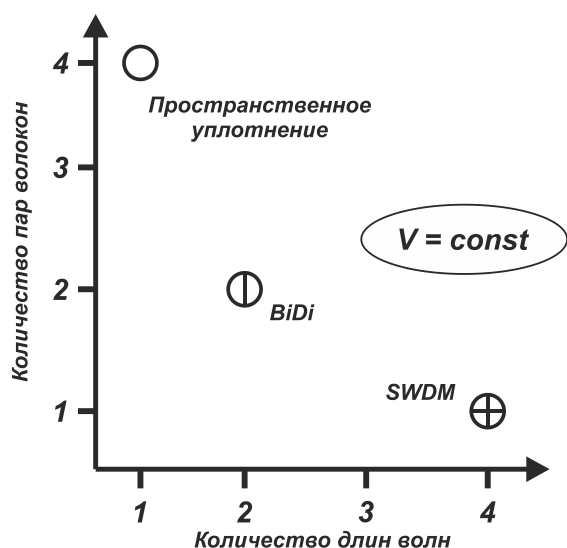


Рис. 6. Спектрально-волоконная диаграмма современных высокоскоростных оптических сетевых интерфейсов

При прочих равных условиях применение пространственного мультиплексирования оказывается экономически более выгодным по сравнению со спектральным. Увеличение расстояния между спектральными каналами для упрощения и, соответственно, удешевления решения в целом, в частности примененное в BiDi, не в состоянии радикально изменить существующее положение дел, которое определяется принципиально небольшой протяженностью кабельных трактов в рассматриваемой области. По этой причине при создании СКС машинного зала ЦОД целесообразно всегда закладывать в нее физическую параллельную оптику.

Base8 как основная схема построения стационарных оптических линий СКС машинного зала ЦОД

Базовые требования к информационной проводке машинного зала ЦОД предельно просты. Она должна нормально функционировать в горизонте нормативного срока службы объекта и при этом быть максимально дешевой.

Нормальное функционирование определяется поддержкой максимальных скоростей передачи, ожидаемых в среднесрочной перспективе. Принимая во внимание отнесение ЦОД к объектам массового применения, оценку значения этого параметра целесообразно проводить с учетом применения известных и уже внедренных в широкую инженерную практику методов при достигнутом уровне техники.

Для определенности принимается, что кабельные тракты реализованы по схеме Base8, а для передачи используется комбинированное амплитудно-фазовое кодирование при тактовой частоте сигнала 25 ГГц. Соответствующие вычисления дают:

$$W = 25 \text{ ГГц} \times 4 \text{ волокна} \times 4 \text{ бит} / \text{Гц} \times \\ \times 4 \text{ длины волны SWDM} = 1,6 \text{ Тбит} / \text{с}$$

Полученное значение, с одной стороны, отвечает текущим потребностям отрасли, а с другой, не потеряет актуальности в среднесрочной перспективе по меньшей мере до 2035 года.

Перспективы одномодовой техники в СКС машинного зала ЦОД

Указанные выше профильные стандарты допускают на равных правах построение волоконно-оптической подсистемы СКС машинного зала ЦОД на одномодовой и многомодовой элементной базе, но не содержат никаких рекомендаций в отношении выбора конкретной базы. Решение данного вопроса остается в компетенции исключительно автора проекта кабельной системы.

За основу дальнейших рассуждений допустимо принять, что одномодовое и многомодовое решения функционально идентичны. При таких условиях вопрос выбора одного из них переходит главным образом в финансовую сферу.

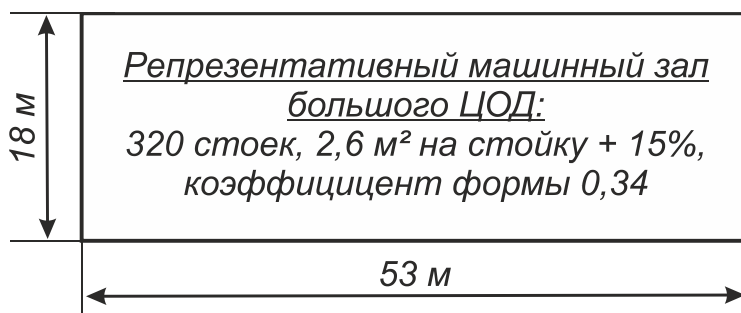


Рис. 7. Топологическая модель машинного зала большого ЦОД

В качестве прототипа СКС для ЦОД использовались офисные кабельные системы со 300-метровыми максимальными длинами трактов волоконно-оптической подсистемы. Столь большие протяженности трактов избыточны для типового ЦОД, машинный зал которого при типовой системе воздушного охлаждения всегда будет отличаться геометрической компактностью.

Методом экспертной оценки установлено, что максимальное количество стоек, которые устанавливаются в машинном зале, может быть принято равным 320. При превышении этого значения целесообразно переходить на многозальные конфигурации.

Расчет по модели, показанной на Рис. 7 и соответствующей проекту отечественных Правил проектирования зданий и сооружений центров обработки данных, показывает, что ожидаемая максимальная протяженность тракта не превысит 71 м (получена как полупериметр машинного зала прямоугольной в плане формы в виде оценки сверху), т. е. обеспечение работоспособности кабельного тракта протяженностью свыше 70–150 м лишено практического смысла при любой скорости передачи. При таких дальностях передачи многомодовая техника оказывается предпочтительнее с финансовой точки зрения.

В качестве фокусной области применения одномодовой техники в составе именно информационной инфраструктуры машинных залов ЦОД целесообразно рассматривать относительно немногочисленные межзальные соединительные линии.

Особо подчеркнем, что посыл о необходимости реализации внутренних линий машинного зала на многомодовой технике не действует в отношении общедоступных ЦОД. В этой разновидности объектов заказчику фактически предоставляются стойко-места с гарантированными температурными режимами и качеством электроснабжения. В результате целесообразно применять прямое подключение активного оборудования к соединительной линии и переход на многомодовую оптику лишен практического смысла.

Перспективы многомодового волокна категории OM5

Многомодовое волокно категории OM5 было целенаправленно разработано для поддержки функционирования перспективных быстродействующих сетевых интерфейсов, использующих схему 4-волнового спектрального мультиплексирования SWDM в диапазоне 850...940 нм с расстоянием между каналами 30 нм. Обратно совместимо со своим предшественником – волокном OM4 по коэффициенту широкополосности на длине волны 850 нм и отличается от него нормированием коэффициента широкополосности на длине волны 953 нм при условии лазерного возбуждения.

Несмотря на зафиксированное в стандарте ISO/IEC 11801 формальное равенство по коэффициенту широкополосности с волокном OM4, OM5 реально значимо превосходит предшественника благодаря более совершенной технологии формирования заготовок.

Таблица 1. Максимальная протяженность кабельного тракта для параллельной передачи по ISO/IEC 11801

Скорость, Гбит/с	Тип интерфейса	Максимальная протяженность кабельного тракта, м			Количество пар волокон
		OM3	OM4	OM5	
100	100G Base-SR2	70	100	100	2
	100G Base-SR4	70	100	100	4
	100G Base-SR10	100	150	150	10
200	200G Base-SR4	70	100	100	4
400	400G Base-SR16	70	100	100	16
	400G Base-SR4.2	70	100	150	4
	400G Base-SR8	70	100	100	8

При этом для наиболее предпочтительная на текущем этапе развития техники схемы Base8 организации связи гарантированная дальность действия примерно в 1,5 раза превышает определенную ранее потребность в 70 м, что закрывает потребности сегодняшнего дня и ближайшей перспективы. Данный факт непосредственно вытекает из Таблица 1.

Что касается среднесрочных прогнозов до 2035 года, то волокно категории OM5 обеспечивает надежные запасы с точки зрения отсутствия морального устаревания. Здесь сказывается то, что согласно соглашениям MSA перспективные сетевые интерфейсы для скоростей 800 Гбит/с и 1,6 Тбит/с предполагается реализовывать на основе спаренных (в первом случае) или счетверенных (во втором) хорошо отработанных 400-гигабитных интерфейсов. Спецификация оптической части у таких устройств с минимальными изменениями заимствуется у прототипа. В качестве иллюстрации сошлемся на Рис. 8, на котором показан один из вариантов реализации трансивера 800G Ethernet как раз в виде спарки двух 400-гигабитных элементов с дополнительной связью между ними для синхронизации их функционирования.

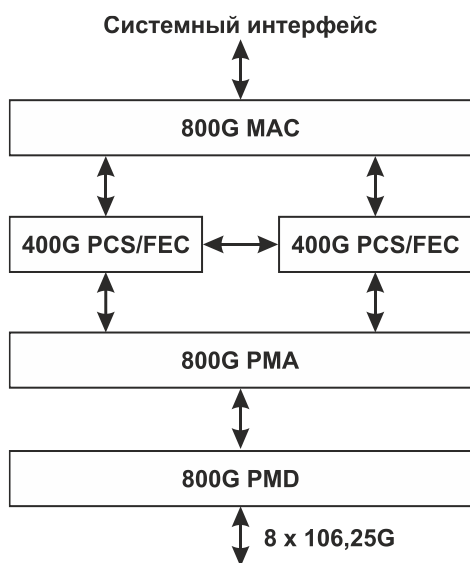


Рис. 8. Увеличенная структурная схема одного из вариантов трансивера 800G Ethernet

Архитектура Spine-Leaf

Архитектура Spine-Leaf представляет собой адаптацию для сетевой компьютерной техники известных из телефонии т. н. сетей Клоза. Наиболее привлекательная сторона подобных сетей заключается в том, что при выполнении определенных условий они позволяют построить строго не-блокируемую или неблокируемую при условии дополнительных перекоммутаций сетевую структуру.

Традиционная сетевая архитектура в общем случае представляет собой 3-слойную структуру, в которой четко выделяются уровни доступа, агрегации и ядра, Рис. 9. За счет известных преимуществ получила широкое распространение и в настоящее время перешла в разряд классических.

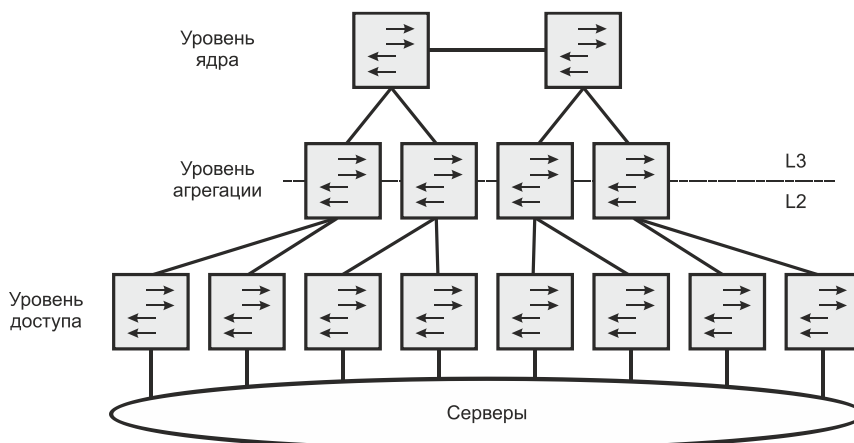


Рис. 9. Традиционная 3-уровневая структура построения ЛВС машинного зала ЦОД

Архитектура Spine-Leaf отличается от традиционной сокращением числа уровней до двух и отсутствием горизонтальных связей на высшем уровне. Данная особенность дополнительно подчеркивается терминологически: уровням присвоены специальные обозначения: Spine и Leaf, Рис. 10.

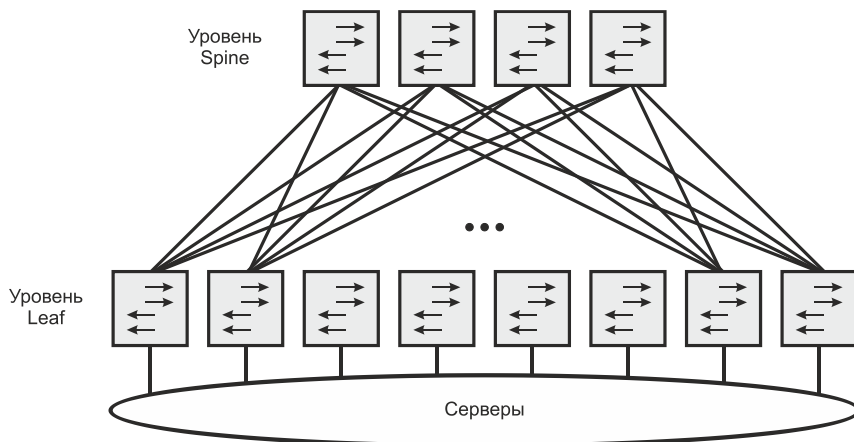


Рис. 10. 2-уровневая архитектура Spine-Leaf построения ЛВС машинного зала ЦОД

Основное преимущество структур подобной разновидности заключается в гарантированном сокращении до трех количества коммутаторов, через которые проходит пакет данных между двумя произвольно взятыми серверами. Это сопровождается значимым уменьшением задержки сигнала при его передаче между двумя конечными точками канала и улучшает эффективность функционирования аппаратной части вычислительной инфраструктуры машинного зала ЦОД.

Немаловажное значение имеет возможность реализации в сетях Spine-Leaf схемы Top of Rack. Для этого достаточно разместить коммутаторы уровня Leaf непосредственно в серверных шкафах.

Прямое сравнение традиционной и Spine-Leaf архитектур прямо указывает на то, что при равном количестве m коммутаторов уровней агрегации и Spine количество линий связи увеличивается в те же m раз. Эта особенность требует заметно больших капитальных затрат на реализацию Spine-Leaf. Поэтому переход на подобные структуры:

- возможен из-за относительно небольшого объема средств, направляемых на реализацию СКС, на фоне общего объема финансирования создания ЦОД;
- из-за неизбежно повышенных затрат имеет смысл только при значительном объеме трафика «восток – запад», т. е. при массовом использовании параллельных вычислений.

Еще одним недостатком архитектуры Spine-Leaf при ее классической реализации является плохая масштабируемость. Для его преодоления целесообразно структурно разделить уровни Spine и Leaf коммутационным полем так, как это показано на схеме Рис. 11. При переходе на такие структуры добавление в нее любого коммутатора с полноценным включением в созданную ранее систему сводится к прокладке двух транковых кабелей (если таковые не были предусмотрены проектом заранее), обычным образом выводимых на панели СКС. Затем коммутаторы подключаются к этим панелям коммутационными шнурами.

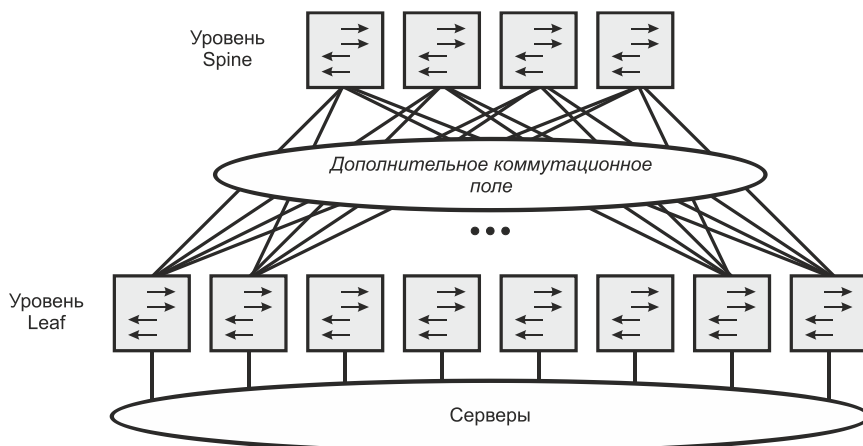


Рис. 11. Архитектура Spine-Leaf с выделенным промежуточным коммутационным полем

Необходимость перехода на новые типы оптических разъемов

Стандартами СКС для ЦОД определено, что в качестве оптического разъема дуплексных трактов должен использоваться LC, а в многоволоконных трактах устанавливается MPO/MTP. Последний создан в еще в 90-х гг. прошлого столетия на основе наконечников MT одноименного разъемного соединителя для внутриблочного монтажа. Как переделка унаследовал ряд недостатков прототипа, среди которых отдельного упоминания достойны:

- несоответствие схеме Base8 и сложности формирования структур Port Breakout;
- функциональная избыточность (прототип рассчитан на 72 волокна, тогда как в ЦОД не требуется более 32);
- сложность обеспечения правильной полярности формируемого тракта;
- большие проблемы с обеспечением нормального физического контакта сращиваемых волокон.

В частности, стремление к решению проблемы полярности и достижению эксплуатационной простоты формирования структур Port Breakout в немалой степени способствовало внедрению в широкую инженерную практику модульно-кассетного построения стационарной линии, схема которого изображена в верхней части Рис. 12. При этом кассета фактически выполняла роль корпусного адаптера, а ее использование представляет собой скорее вынужденный шаг: при необходимости изменения структуры стационарной линии старая кассета просто и, самое главное, быстро отключалась от транкового кабеля и на ее место столь же просто и быстро устанавливалась новая.

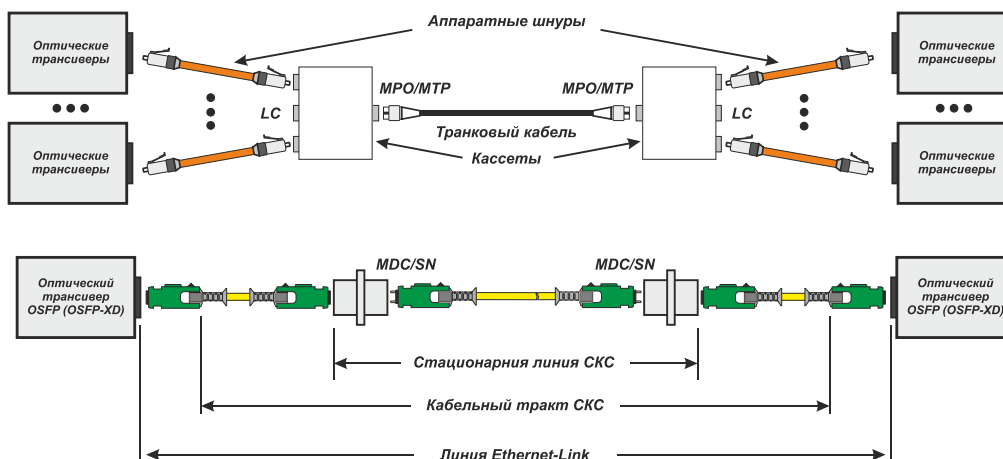


Рис. 12. Структуры кабельных трактов оптической СКС в машинном зале ЦОД при их реализации на разных типах оптических разъемов

При всех достоинствах модульно-кассетной схемы она обладает недостатком: в кабельном тракте появляется дополнительный разъем, который неизбежно вносит потери в передаваемый сигнал. Данный недостаток особенно критичен для высокоскоростных интерфейсов, которые обладают чрезвычайно низким энергетическим потенциалом. Согласно данным Таблица 2 на скоростях 40 Гбит/с и выше величина этого параметра не превосходит 1,9 дБ. В технике последнего поколения нормируются близкие характеристики. Выход из сложившейся ситуации видится в переходе на новый тип разъема, который свободен от указанных недостатков.

Таблица 2. Базовые параметры высокоскоростных волоконно-оптических интерфейсов Ethernet

Тип интерфейса	Год утверждения	Спецификация	Длина тракта, м	Допустимые потери, дБ
1G Ethernet-1000Base-SX	1998	IEEE 802.3z	550	3,56
10G Ethernet-10GBase-SR	2003	IEEE 802.3ae	300 (OM3) 490 (OM4)	2,6 (OM3) 2,9 (OM4)
40G Ethernet-10GBase-SR4	2010	IEEE 802.3ba	100 (OM3) 150 (OM4)	1,9 (OM3) 1,5 (OM4)
100G Ethernet-10GBase-SR10	2010	IEEE 802.3ba	100 (OM3) 150 (OM4)	1,8 (OM3) 1,5 (OM4)

Внедренные в широкую инженерную практику разъемы следующего поколения представляют собой качественно иную технику, что подчеркивается отнесением их к группе VSFF (англ. very small form factor). Они серийно выпускаются американской компанией US Cones (изделие MDC) и японской Senko (изделие SN). О перспективности техники данной разновидности свидетельствуют попытки вывода на рынок своих разработок некоторыми производителями из Юго-Восточной Азии.

Разъемы VSFF имеют следующие отличительные особенности:

- по формфактору сборка из четырех вилок соответствует дуплексному LC, т. е. изначально адаптирована к схеме Base8;
- для оконцевания волокон используются индивидуальные наконечники LC, которые монтируются в корпусе по плавающей схеме, что обеспечивает не только уменьшение, но и стабильность вносимых потерь;
- переход от дуплексной вилки к многоволоконной может производиться прямо на объекте путем установки отдельных вилок в общую пластиковую крепежную обойму;
- при необходимости изменения полярности как всей групповой вилки, так и отдельной дуплексной они просто переворачиваются в корпусе защелки на 180°.

Обе последние операции не только не требуют обращения к специальным инструментам, но и выполняются даже без привлечения отвертки.

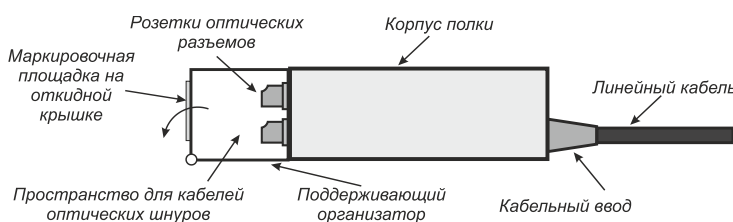
Отдельно укажем на то, что разъемы группы VSFF за счет изначально заложенной в них простоты изменения конфигурации позволяют легко реализовывать структуры, изображенные на Рис. 2, и поэтому потенциально хорошо востребованы в практике текущей эксплуатации кабельной системы.

За счет предельно минимизированных габаритов конструкции MDC и SN позволяют довести конструктивную плотность типовой оптической полки СКС до значения 216 дуплексных розеток на 1U монтажной высоты.

Наращивание эффективности обслуживания кабельной системы

Количество точек коммутации портов волоконно-оптической подсистемы СКС в машинном зале ЦОД постоянно увеличивается и для сохранения на должном уровне удобства ее текущей эксплуатации необходимо принимать специальные меры. Кроме собственно применения сверхмалогогабаритных оптических разъемов группы VSFF в широкую инженерную практику внедряется ряд других компонентов различного уровня сложности, целенаправленно разработанных с учетом этой задачи.

Рассмотрим некоторые из таких решений.

**Рис. 13. Оптическая полка с организатором поддерживающего типа**

В области производства оптических панелей (полок) наметилась четкая тенденция перехода на конструкции, содержащие интегральный поддерживающий организатор с дверцей, Рис. 13. Последняя откидывается вниз на петлях и содержит площадку для нанесения пользовательской маркировки. Сама площадка может располагаться как на внутренней, так и на внешней стороне дверцы. Внутреннее расположение облегчает коммутацию за счет свойства зрительной опоры, внешнее – поиск нужного порта.

Откидываемая вниз горизонтальная дверца закрывает кабели коммутационных шнуров и устраняет известный эффект «кабельного салата». Выход кабелей из организатора происходит с левой и правой сторон, откуда они сразу же попадают в вертикальный организатор.

Увеличение количества линейных и шнуровых кабелей в связи с переходом на структуры Spine–Leaf стимулирует отрасль на улучшение массогабаритных показателей кабелей. Необходимый положительный эффект достигается несколькими основными конструктивными приемами.

Первым из таких приемов становится отказ от классической конструкции типа Zip-cord в пользу кабеля с общей круглой оболочкой. Выигрыш от ее применения демонстрирует Рис. 14.

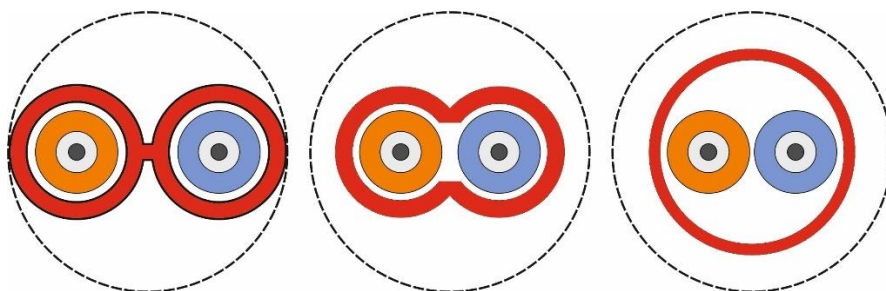


Рис. 14. Сравнение поперечных размеров дуплексных кабелей: слева – классический Zip-cord, в центре – квази-Zip-cord, справа – с круглой оболочкой

Еще одна группа решений основана на уменьшении диаметра волоконного световода. При этом сокращается как диаметр первичного защитного покрытия, так и (пока реже) диаметр оболочки самого волокна.

Тренд на сокращение диаметра оболочки пока явно не обозначился, и это направление относится, скорее, к ближайшей перспективе. Техническая возможность реализации таких волокон определяется совершенствованием технологии изготовления и некоторой избыточностью диаметра стандартной 125-микронной оболочки даже для одномодового световода, значительная часть излучения в котором распространяется по оболочке. Выдвигаются предложения снизить диаметр до 100 и 85 мкм, общее решение пока не выработано.

Кроме улучшения массогабаритных показателей кабелей, волокна рассматриваемой разновидности позволяют увеличить плотность конструкции групповых оптических соединителей за счет сокращения расстояния между осями направляющих каналов. Это уменьшает длину ряда и облегчает тем самым достижение физического контакта оптически активных поверхностей.

В области выбора диаметра первичного защитного покрытия как технологически более простой процедуры работы идут более активными темпами. Ряд производящих компаний уже сейчас предлагают серийные волокна со значением этого параметра в диапазоне 180–200 мкм. Представление об эффекте от перехода на «тонкие» волокна дает Рис. 15.

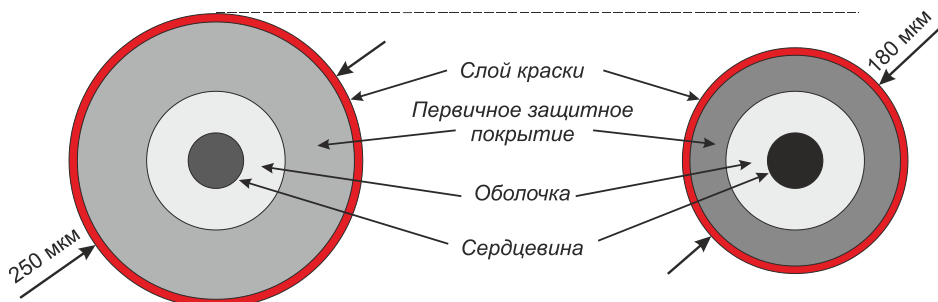


Рис. 15. Волокна с первичным защитным покрытием различного диаметра

Еще одним направлением улучшения массогабаритных показателей линейных кабелей становится широкое внедрение ленточных конструкций. В данном случае выигрыш достигается благодаря более плотному и гарантированному за счет жесткой механической связи взаимному расположению волокон. Возможны также иные решения, которые не получили массового распространения.

Недостаток классических ленточных конструкций в виде плохой гибкости устранен в структурах типа Spider Web. От обычной ленты они отличаются тем, что отдельные волокна связаны друг с другом не непрерывно, а только на определенных участках. Это позволяет растягивать ленту, изгибать ее в поперечном направлении и даже складывать в стопку, что демонстрирует Рис. 16.

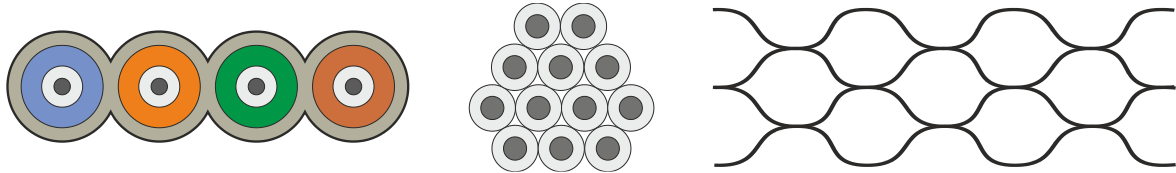


Рис. 16. Разновидности и свойства ленточных волокон. Слева – один из вариантов классической 4-волоконной конструкции, в центре – лента Spider Web в сложенном состоянии, справа – лента Spider Web в растянутом состоянии

Внедрение новых разновидностей оптических кроссов

В составе оптической подсистемы СКС вне зависимости от объекта ее установки в подавляющем большинстве случаев предусматриваются оптические кроссы. Тракты прямого подключения получили крайне ограниченное распространение только в промышленных системах.

На нижних уровнях информационной инфраструктуры в качестве оптического кросса можно применять традиционные конструкции, отличающиеся от офисных в основном только увеличенной плотностью портов. Последнее затрудняет администрирование, но не является критически важным из-за характерной для ЦОД статичности топологической конфигурации кабельной системы.

Ситуация меняется на прямо противоположную в центральных кроссах, в которых из-за массового внедрения структур Spine-Leaf количество одновременно обслуживаемых портов может даже превышать десять тысяч.

Применяемые для построения центрального кросса решения из области телекома не в полной мере отвечают поставленным требованиям. Для устранения имеющихся у них недостатков предложены специализированные конструкции, основанные:

- на применении оптических полок с малогабаритными разъемами группы VSFF;
- использовании для коммутации шнуров одинаковой длины на основе круглых кабелей;
- введении в состав многоволоконного кросса группового интегрального организатора с барабанными поддерживающими элементами.

Основную идею и структуру подобного кросса демонстрирует Рис. 17.

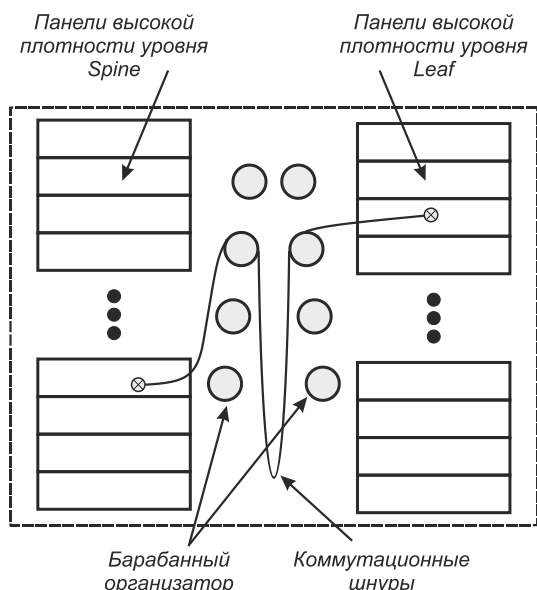


Рис. 17. Конструктивная схема многоволоконного оптического кросса с коммутационными шнурами одинаковой длины

Заключение

1. Физический уровень информационной инфраструктуры машинного зала масштабного ЦОД должен реализовываться исключительно на волоконно-оптической элементной базе.
2. Решения на базе кабелей из витых пар могут применяться исключительно в микроЦОД, а на больших объектах с их помощью можно решать только вспомогательные задачи управления, мониторинга и аналогичные им.
3. В проект СКС машинного зала ЦОД необходимо закладывать по меньшей мере 10...15-летний срок нормальной эксплуатации.
4. Стационарные линии СКС машинного зала ЦОД предпочтительно формировать по схеме Base8 на основе кабелей с многомодовым волокном категории OM5.
5. Необходимость разработки новых типов многомодового волокна отсутствует по крайней мере в среднесрочной перспективе.
6. Массовое применение одномодовой техники в СКС машинного зала в общем случае нецелесообразно, а необходимость обращения к ней должна обосновываться отдельно.
7. В качестве фокусной области применения одномодовых решений в корпоративных ЦОД следует рассматривать соединительные линии отдельных залов масштабных многозальных объектов.
8. При проектировании СКС необходимо изначально ориентироваться на применение физической параллельной оптики в составе стационарных линий.
9. Архитектуры Spine-Leaf, перспективные в смысле наращивания качества формирования ответа на поступающий пользовательский запрос, неизбежно увеличивают затраты на информационную кабельную систему машинного зала ЦОД.
10. Переход на оптические разъемы группы VSFF объективен и значительно расширяет функциональные возможности кабельной системы.
11. Применение более плотных ленточных конструкций кабельных изделий различного назначения с улучшенными массогабаритными параметрами заметно увеличивает эффективность администрирования СКС машинного зала.
12. Рост числа точек коммутации, характерный для архитектуры Spine-Leaf, стимулирует увеличение объемов внедрения панелей высокой плотности и кроссов, позволяющих работать со шнурами одинаковой длины.