

Redes Fabric

Projete sua rede para o futuro — de
10G até 400G, e além



Conteúdos

Resumo executivo	3
Novas arquitecturas para data centers	3
Design da topología fabric – capacidade	4
Network fabric (estrutura de redes) – a rede física	6
Design da topología fabric – densidade de portas do switch	7
Design da topología fabric – alcance	8
Design da topología fabric – suporte de aplicações	12
Resumo das solucións de migración	16

Resumo Executivo

A ampla adoção da virtualização e do cloud computing derivou na necessidade de novas arquiteturas de equipamentos para data centers que forneçam uma latência mais baixa e taxas de transferência mais altas. Estas novas arquiteturas se baseiam em switches de redes do tipo fabric (estruturais) e são diferentes das topologias de comunicação tradicional de três camadas.

Estes switches fabric podem tomar várias formas —de extensões fabric em uma implementação na parte superior do rack, a soluções fabric centralizadas no HDA (Área de Distribuição Horizontal) ou IDA (Área de Distribuição Intermediária), até uma arquitetura de rede completa. Em tal sentido, deve-se considerar especialmente a forma em que se projeta e implementa a infraestrutura da camada física para garantir que o switch de estrutura ou o fabric possa crescer de forma fácil e eficiente.

O presente documento técnico oferece uma perspectiva geral da tecnologia fabric, junto com as considerações sobre design e um olhar prático para implementar a solução de conectividade de fibra que puder se adaptar às frequentes mudanças na arquitetura, assim como a mais altas velocidades de linha à medida que a rede cresce.

Também se oferecem exemplos práticos de designs de redes fabric com a solução de fibra pré-terminada SYSTIMAX® InstaPATCH® 360, destacando a importância de projetar uma infraestrutura que suporte velocidades mais altas e o crescimento da rede.

Novas arquiteturas para data centers

Os designs e as arquiteturas para data centers evoluíram para se adequar ao crescimento dos serviços informáticos e de armazenamento baseados em cloud. Os data centers corporativos privados tradicionais estão adequando suas atuais arquiteturas para estar preparados para designs novos e ágeis baseados no cloud. Estas novas arquiteturas empresariais se parecem às instalações do “tamanho de um armazém”, mas foram projetadas para suportar as mais variadas aplicações empresariais.

Para estar preparado para as arquiteturas de cloud, estabelece-se uma rota direta otimizada para o estabelecimento das comunicações entre servidores através de uma arquitetura do tipo “leaf-spine” (ver Figura 1). Este design permite que as aplicações em qualquer computador ou dispositivo de armazenamento trabalhem juntas de forma previsível e com capacidade de crescimento, independentemente de sua localização física dentro do data center.

As “redes na cloud” (ou cloud networks) se baseiam em uma arquitetura que consiste em uma rede de conexões entre switches do tipo leaf e spine. A rede de links de rede com frequência é mencionada como “network fabric” (estrutura de redes). O desempenho da tecnologia fabric é ideal para estabelecer “serviços no cloud” universais: isto permite a conectividade multidirecional, com capacidade previsível e baixa latência. A tecnologia fabric oferece redundância de forma inerente, já que se espalham múltiplos recursos de comutação ou switches ao longo do data center, o que ajuda a garantir uma maior disponibilidade das aplicações. Estes designs de redes distribuídas podem ser muito mais rentáveis de implementar e estender, se forem comparados com as grandes plataformas de comutação centralizadas tradicionais.

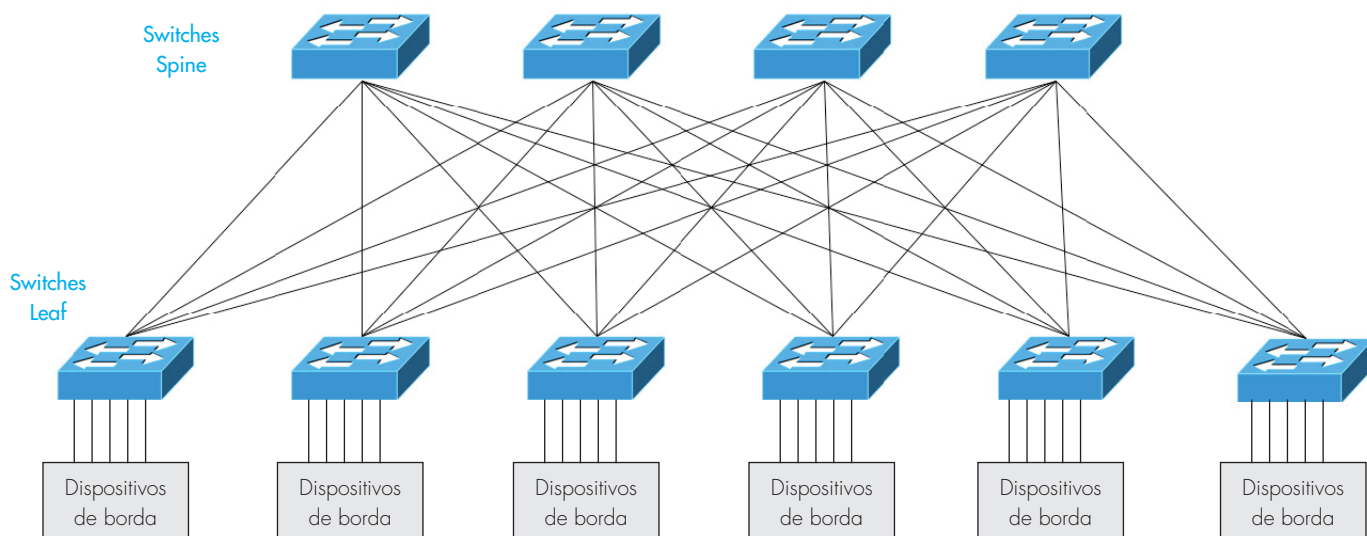


Figura 1. Uma arquitetura de network fabric “com uma trama”, com conexões entre cada switch leaf e spine, que fornece conectividade multidirecional

Design da topologia fabric — capacidade

A forma tradicional de projetar redes é concentrar o tráfego através da agregação (quer dizer, se uma rede necessita suportar 10 fluxos de dados de 1 Gb cada um, para agregar tais fluxos seriam necessários links de rede de 10 Gb). As redes do tipo leaf-and-spine, no entanto, funcionam de forma diferente. Para fazer crescer as redes do tipo fabric, os designers precisam considerar os seguintes fatores:

- A velocidade (ou largura de banda) dos links fabric
- A quantidade de servidores/portas de dispositivos de storage (também conhecidos como portas de borda ou “edge ports”)
- A largura de banda total da rede fabric necessária para oferecer serviços a todas as aplicações do data center

A velocidade da rede fabric não é a capacidade de transporte total entre cada par de switch leaf dentro da estrutura; é o total da largura de banda necessário entre cada switch leaf e todos os switches spine.

No exemplo a seguir há quatro switch spine. Se cada switch leaf tem um link de 40 Gb para cada switch spine, o resultado é uma estrutura ou fabric de 160 Gb. Adverta-se que cada switch leaf deve ter a mesma velocidade de link para cada switch spine. Também é preciso advertir que não há conexões de dispositivos diretamente aos switches spine.

A velocidade da estrutura ou fabric precisa ser dimensionada para suportar a maior quantidade de tráfego que qualquer switch leaf individual puder transmitir. Por exemplo, se existirem 48 portas de 10 Gb conectados a servidores de alta velocidade, a rede fabric necessitaria suportar $48 \times 10 \text{ Gb}$ —ou 480 Gb de largura de banda.

O número total de portas de borda é a próxima consideração de importância. Esta é uma função do número de switches leaf na rede fabric. Por exemplo, se um switch leaf fornece 24 portas de 10 Gb, cada switch leaf adicional soma outras 24 portas à rede fabric total. Só é possível adicionar um novo switch leaf se cada switch spine tiver uma porta adicional disponível para o novo switch leaf.

Quando os links entre switches leaf e spine (a 40 Gb, por exemplo) têm mais capacidade que os links de porto de borda (a 10G, por exemplo), o design se menciona como um “fat tree”. Se os links forem mantidos na mesma velocidade (por ex., 10G na borda: $4 \times 10\text{G}$ leaf-spine) o design seria mencionado como um “skinny tree”. Os designs do tipo fat tree oferecem óbvios benefícios para o crescimento da rede fabric do data center. Cada switch leaf-and-spine deve ter suficientes portas para permitir a trama de conexões multidirecionais. A quantidade de portas e a capacidade de cada porta predeterminam o tamanho máximo e a largura de banda até o qual a rede fabric pode crescer.

Um fat tree típico poderia usar links fabric de 40 Gb. Em nosso exemplo abaixo temos quatro switches spine e cada um deles suporta seis portas de 40 Gb —com uma largura de banda total de 240 Gb. Supondo que cada switch leaf tem 48 portas de 10G, isto dá como resultado um total de 288 portas de borda de 10 Gb cada um. No entanto, a maioria dos dispositivos terão uma dupla conexão, o que requer duas portas de 10 Gb por cada dispositivo de borda. Esta configuração suportará 144 dispositivos de borda com plena redundância.

Design da topologia fabric - oversubscription

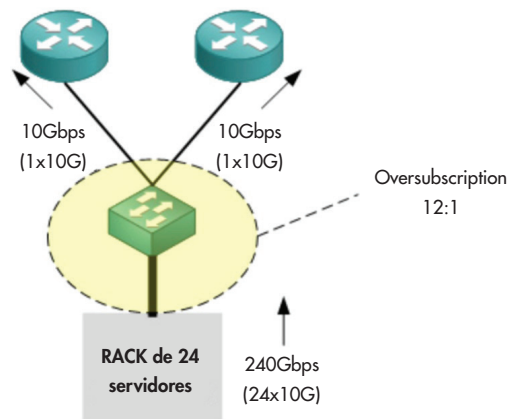


Figura 2. Velocidades do link leaf-spine e oversubscription

A largura de banda total da rede fabric pode ser calculada multiplicando a quantidade de portas de borda pela velocidade de tais portas —ou o número de portas spine pela velocidade das portas spine. Se não existir oversubscription, estes dois números serão iguais. A ideia é que as network fabrics sejam não bloqueadores; quer dizer, nelas todo o tráfego produzido pelos dispositivos de borda pode correr sobre a estrutura ou fabric sem demoras ou “bloqueios”. O que significa que algo do tráfego pode resultar bloqueado ou demorado porque os recursos estão sendo plenamente utilizados por outros. O bloqueio pode impactar de forma severa sobre as aplicações do data center —especialmente as aplicações como FCoE, que dependem de um ambiente não bloqueante. Muitas arquiteturas de redes fabric mantêm redes de armazenamento independentes —algumas com Fibre Channel, outras com armazenamento baseado em IP e algumas com armazenamento distribuído definido pelo software.

Os designers consideram como se comunicam as aplicações e calculam os requisitos de capacidade em geral, que equivale ao tamanho da estrutura (fabric) da rede. Alguns designs de rede incluem um compromisso que se adéqua à qualidade do serviço e ao orçamento apropriado para os serviços que serão oferecidos, o que significa que um nível aceitável de bloqueio ou contenção para os recursos de rede são incluídos no design geral da arquitetura da rede. A taxa de oversubscription descreve o nível de contenção dos recursos que existe para os dispositivos de borda. Um exemplo se mostra na Figura 2, com uma taxa de oversubscription de 12:1.

Se a taxa de oversubscription for muito alta, a performance das aplicações se vê prejudicada. Se a taxa de oversubscription se mantiver baixa, reduz-se a quantidade de servidores —e por conseguinte a quantidade de aplicações que podem ser suportadas pela rede fabric. Este balanço entre custos de capital e capacidade das aplicações é um fator de design crítico. Também é um fator que tem muita probabilidade de mudar rapidamente com a passagem do tempo, à medida que cresce a demanda de aplicações. A capacidade do hardware de servidores tende a aumentar, o que significa que a capacidade dos links da rede fabric se verá super exigida.

Resulta claro, a partir do debate proposto anteriormente, que uma maior capacidade do link leaf-spine pode melhorar o nível do serviço ao minimizar a taxa de oversubscription e aumentar a quantidade de servidores que podem ser suportados pela network fabric. Idealmente, a capacidade destes links deveria ser tão alta quanto praticamente seja possível.

À medida que a arquitetura fabric se expande, devem ser realizadas conexões a cada dois dispositivos semelhantes. A quantidade de conexões cresce rapidamente, à medida que se agregam os switches leaf. A conectividade da camada física deve se adaptar para suportar estas network fabrics com maior densidade, mais alta velocidade de links de rede e modularidade multifibra —o que, por outro lado, ajuda a acelerar a implementação e a disponibilidade da rede. A seguir, mostram-se os cabos do equipamento MPO que poderiam ser usados para fornecer conexões físicas para o link de rede Ethernet de 40G - QSFP (4 X 10G).

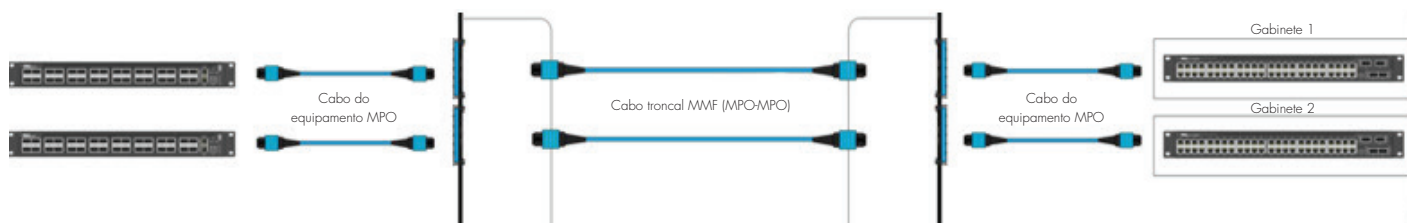


Figura 3. Conectividade de switch de 40G com 40GBASE-SR4 (nota: para simplificar, não se mostram os pins)

Para otimizar a capacidade da rede fabric, os componentes ópticos devem fornecer alta largura de banda e baixa perda —predispondo-se assim para as velocidades de rede de próxima geração. Os 40G, 100G ou inclusive os 400G deveriam ser parte dos requerimentos de design desde o primeiro dia, para prevenir a necessidade de tarefas de redesign da infraestrutura de cabeamento a futuro.

A tecnologia da rede óptica capaz de suportar estes links está evoluindo rapidamente. As velocidades estão aumentando de forma cada vez mais acelerada —em alguns casos, as soluções que são oferecidas nesta área estão muito a frente dos padrões da indústria. Os benefícios em função dos custos destas várias opções são chave para se manter em sintonia com os requisitos quanto à capacidade do data center em geral e constituem um componente fundamental do balanço entre o CapEx e o risco de disponibilidade.

Network fabric — a rede física

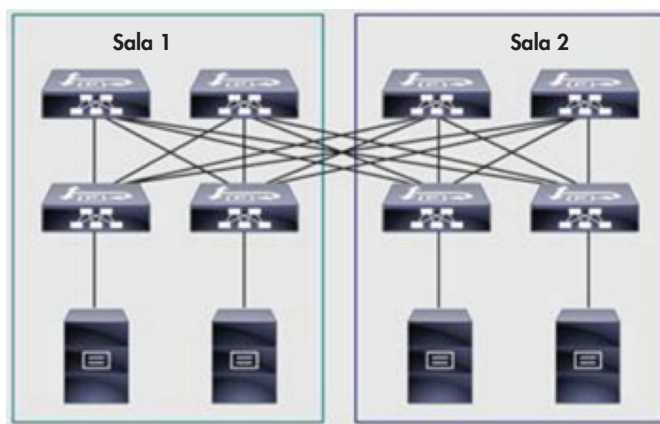


Figura 4. Design de rede do tipo leaf-spine que abrange múltiplos corredores

A implementação das redes fabric é similar à das redes tradicionais de três níveis implementadas no passado de várias formas — devem ter capacidade de expansão, ser facilmente manejáveis e confiáveis. O uso de designs de cabeamento estruturado continua sendo igualmente válido e valioso quando se implementam topologias de redes fabric. Os canais e espaços do cabeamento continuam sendo os mesmos. As redes fabric podem ser estendidas ao longo de múltiplos corredores em um data center. Os componentes da rede fabric também requerem suporte para redes de gestão fora de banda. Estes requerimentos para o design físico são incorporados ao diagrama do andar. A seguir, mostra-se um diagrama típico.

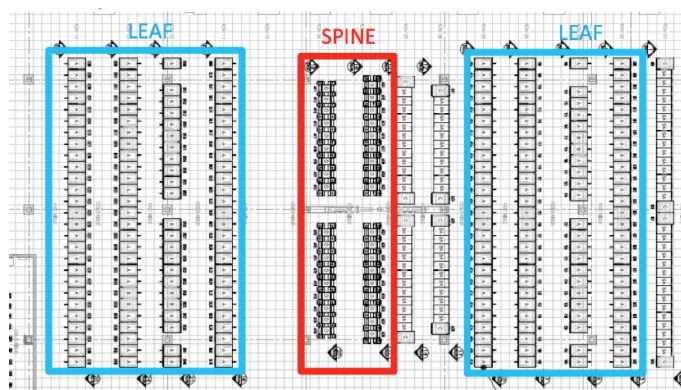


Figura 5. Vista em planta do diagrama com gabinetes leaf-spine

A Figura 6 ilustra uma topologia física de data center com uma conexão cruzada à área de distribuição intermediária (IDA): os switches spine são mostrados na Área de Distribuição Principal (MDA) e os switches leaf são mostrados na Área de Distribuição Horizontal (HDA).

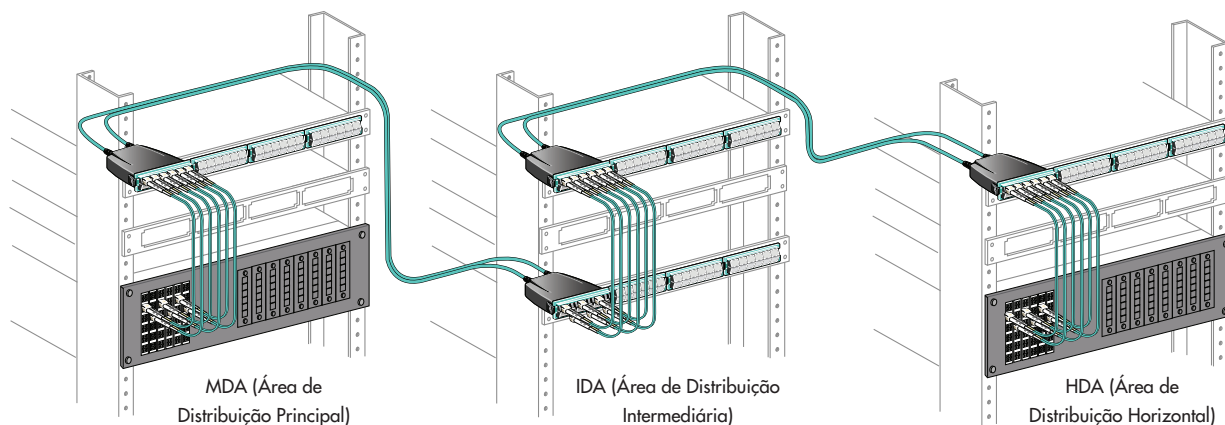


Figura 6. Design de rede leaf-spine com InstaPATCH 360, com conexão cruzada no IDA (Área de Distribuição Intermediária)

O design de uma rede fabric com conectividade cruzada aumenta consideravelmente a flexibilidade e a agilidade e simplifica o que pode ser um canal muito complexo. O uso das conexões cruzadas nos data centers já se tornou obrigatório conforme a norma CENELEC EN 50600-X na Europa. Baseado na capacidade para fornecer conectividade multidirecional entre portas, recomenda-se a arquitetura de conexões cruzadas ou cross-connect.

¹ Para obter mais informação sobre topologias de rede, consulte a Norma BICSI para Data Centers.

Canais e espaços no data center que suportam as arquiteturas de redes fabric

O design dos links da rede física dependerá em grande medida da topologia da rede em geral e do provedor de networking (conexões de rede). Alguns provedores de switches oferecem ópticas patenteadas com preferência pela ativo monomodo ou multimodo. Embora outros favoreçam os switches maiores, baseados no chassis, e o cabeamento de área horizontal baseado na zona. A visão a futuro das redes de próxima geração também mudará. Em muitos casos, a equipe de design de cabeamento é a última a saber que hardware de rede particular deverá ser suportado desde o primeiro dia. O pacote de ferramentas ideal permitirá aproveitar a grande variedade de opções que possam ser apresentadas e tornará mais fácil avaliar as futuras opções de rede — respaldando as novas iniciativas enquanto se evita comprometer com um único provedor.

O sistema de cabeamento de fibra pré-terminada InstaPATCH 360 que se mostra na Figura 6 se adapta perfeitamente a uma plataforma de cabeamento estruturado de alta performance produzida em fábrica, que maneja facilmente o amplo conjunto de aplicações de fibra necessárias para suportar as redes fabric. Os cabos troncais InstaPATCH 360, assim como os módulos de distribuição e os cabos de interconexão ou patch cords, configuram-se para se adequar aos requerimentos dos equipamentos informáticos, troncais e do switch desde o primeiro dia, assim como para fornecer uma rota de migração que cumpra com os requerimentos subsequentes.

No exemplo da rede fabric, discutimos como a conectividade leaf-spine deveria ser vista como se ilustra abaixo — um troncal de fibras multimodo paralelas. Este design usa ópticas multimodo de custo mais baixo, mantém retrocompatibilidade com relação às tecnologias de rede anteriores e pode fornecer um plano de melhorias para atingir uma capacidade de 100G no futuro, desde que o design do link seja válido.

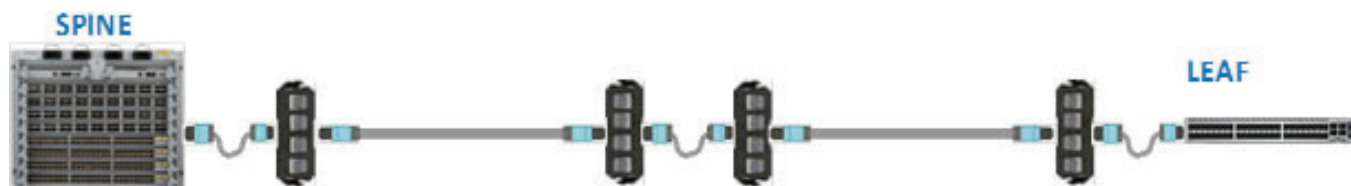


Figura 7. Design de rede leaf-spine com troncais de fibras paralelas e Área de Distribuição Intermediária (IDA)

Design da topologia fabric — densidade de portas do switch

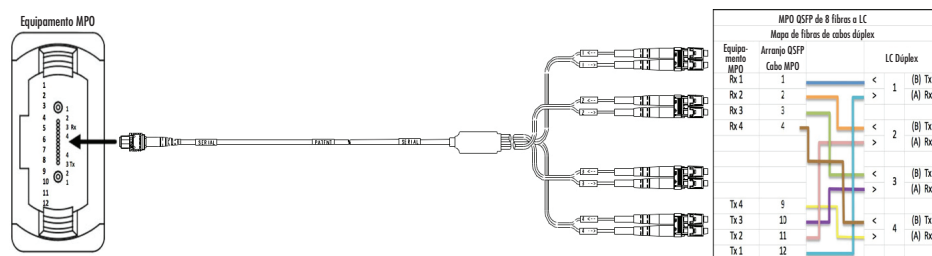


Figura 8. Encaixe de portas LC com QSFP

Múltiplas portas de 10G podem se reunir para suportar links de capacidade mais alta. As normas de IEEE indicam a formação de um grupo de quatro portas de 10G, que se combinam em um único conector MPO de 12 fibras para formar um link de 40G. Este padrão de QSFP é usado para criar links de maior capacidade (40G), mas também com frequência se usa para conectar uma única porta em um switch leaf com quatro servidores — aumentando a densidade do cabo troncal e a capacidade do painel do switch leaf. A combinação de quatro portas LC em uma mesma aplicação QSFP gera aproximadamente um aumento da densidade por painel em uma proporção de 4:1 sobre um switch leaf se for comparado com o uso de portas seriais independentes, projetadas para interfaces SFP+.

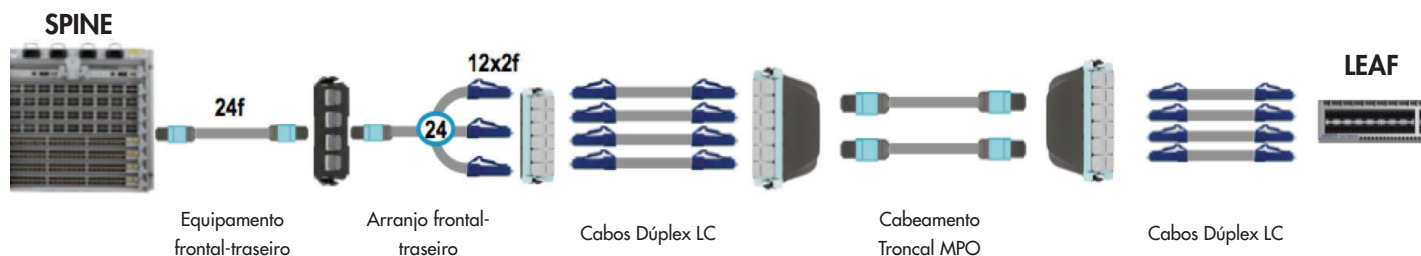


Figura 9. Conectividade de servidor a switch leaf com InstaPATCH 360

Os provedores de equipamento fazem uso dos conectores MPO para oferecer maior densidade de fibras. Por exemplo, alguns provedores oferecem 12 portas de 10 Gb em um conector MPO de 24 fibras. Estes podem se agrupar de três em três (de 40G cada um) ou talvez também se subdividir em até 12 portas de 10G para as conexões de dispositivos. Ao agrupar as fibras em sistemas de conexão de mais alta densidade se economiza espaço nos painéis onde se conectam os dispositivos de rede e isto permite uma cômoda gestão através dos arranjos de troncais de fibras paralelas.

Design da topologia fabric —alcance

Os data centers com frequência são grandes em termos de espaço físico e da quantidade de computadores e dispositivos de armazenamento que contêm. Existem várias normas sobre data centers que detalham as melhores práticas quanto a cabeamento para a extensão de redes e o design do espaço. Entre os exemplos, incluem-se ANSI/TIA-942-B, ISO/IEC 11801-5, CENELEC EN50173-5 e EN50600-X.

Os designs de cabeamento estruturado permitem fazer a rede crescer, melhorar o tempo médio de reparação (MTTR) e a disponibilidade em geral. Portanto, é altamente recomendável manter esta estrutura de cabeamento com topologias de rede fabric. Os links leaf-spine também devem usar canais de comunicação adequados em espaços de comunicação designados —tal como acontecia com as topologias de redes anteriores.

Fornecer links de alta capacidade por um custo razoável é um componente de design chave nas redes fabric. Os dispositivos óticos multimodo são normalmente menos caros do que os dispositivos óticos monomodo equivalentes —especialmente à medida que aumenta a velocidade da rede. Existe atualmente uma ampla variedade de opções disponíveis para o designer de rede: tanto soluções baseadas nos padrões quanto patenteadas, que oferecem diferentes combinações quanto à capacidade, custos e riscos operativos/benefícios. Estão surgindo novas interfaces de links de dados que oferecerão ainda mais opções para o design dos links. A tecnologia de cabeamento deve habilitar a capacidade de rede a curto prazo e possibilitar designs do tipo fabric com aumento de tamanho e capacidades.

O design dos links de rede, portanto, é uma consideração importante na hora de criar redes fabric. As zonas de patchamento podem ser úteis em cada uma das áreas de distribuição, como se mostra no exemplo de configuração abaixo. O suporte das aplicações de networking varia, mas quanto maior for a velocidade, mais curta é a distância que o link de cabeamento estruturado pode suportar. Aumentar a quantidade dos cabos de interconexão também reduz o sinal no link e, por conseguinte, reduz a distância utilizável do link. Com bastante frequência os provedores de hardware de rede oferecem especificações em termos da distância máxima suportada entre um ponto e outro. É importante entender a relação que estas especificações terão, se observarmos os atuais designs de cabeamento estruturado.

Vamos supor que você esteja planejando a implementação de novos serviços no data center segundo as topologias que se mostram na Figura 10. As aplicações que devem ser suportadas desde o primeiro dia incluem Ethernet 10 Gb e Fibre Channel (FC) 8G. O data center foi organizado em corredores de dados manejáveis dentro do data center. Haverá problemas para suportar os comprimentos de links requeridos por seu design?

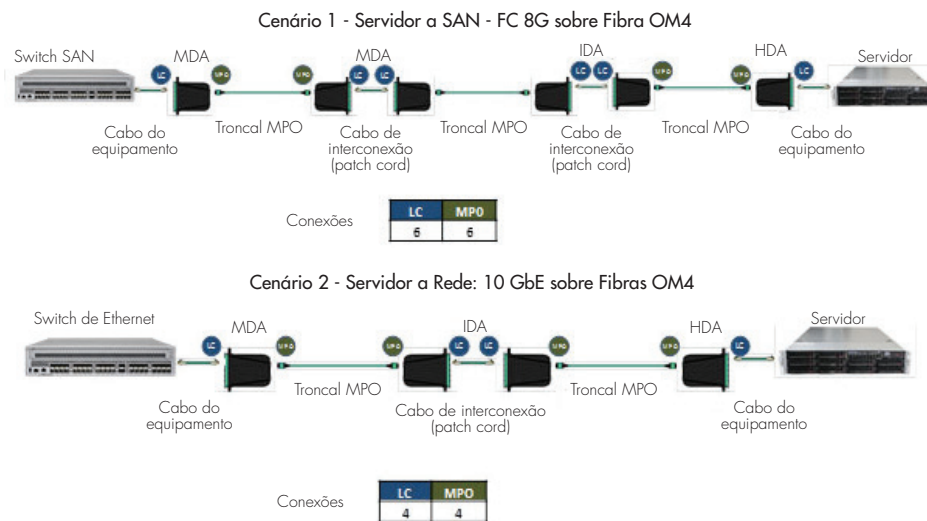


Figura 10. Exemplo de topologias de data centers

Parte da intenção do design requer que, para assegurar futuras melhorias quanto à capacidade, a infraestrutura da rede possa ser capaz de suportar pelo menos as próximas velocidades de links de dados mais altas. A comunidade de provedores oferece várias opções: algumas soluções monomodo, outras multimodo, algumas baseadas nos padrões, e inclusive outras novas e patenteadas. Daqui em diante, qual das potenciais aplicações de rede futuras suportará seu design desde o primeiro dia?

O processo de identificar a melhor solução começa com a compreensão das opções de design que estão sendo consideradas. A topologia proposta funcionará de forma confiável com os equipamentos de rede que estão sendo considerados? Se existirem opções disponíveis, que estratégia parece oferecer o melhor custo comercial e a mais alta confiabilidade? Para responder todas estas perguntas, primeiro observamos os padrões da indústria que detalham as opções existentes para nossos links de dados de Ethernet. Esta tabela inclui os padrões completados, e também aqueles em fase de estúdio. Existem aplicações adicionais, incluindo as de 50G e 200G, que atualmente se encontram na fase de grupo de trabalho em IEEE 802.3.

Aplicação	Padrão	Referência IEEE	Meio	Velocidade	Distancia objetivo
Ethernet 10-Gigabits	10GBASE-SR	802.3ae	MMF	10 Gb/s	33 m (OM1) a 550 m (OM4)
	10GBASE-LR		SMF		10 km
	10GBASE-LX4		MMF		300 m
	10GBASE-ER		SMF		40 km
	10GBASE-LRM	802.3aq	MMF		220 m (OM1/OM2) a 300 m (OM3)
Ethernet 25-Gigabits	25GBASE-SR	P802.3by	MMF	25 Gb/s	70 m (OM3) 100 m (OM4)
Ethernet 40-Gigabits	40GBASE-SR4	802.3bm	MMF	40 Gb/s	100 m (OM3) 150 m (OM4)
	40GBASE-LR4		SMF		10 km
	40GBASE-FR		SMF		2 km
	40GBASE-ER4		SMF		40 km
Ethernet 100-Gigabits	100GBASE-SR10	802.3bm	MMF	100 Gb/s	100 m (OM3) 150 m (OM4)
	100GBASE-LR4		SMF		10 km
	100GBASE-SR4		SMF		70 m (OM3) 100 m (OM4)
	100GBASE-ER4		SMF		40 km
Ethernet 400-Gigabits	400GBASE-SR16	P802.3bs	MMF	400 Gb/s	70 m (OM3) 100 m (OM4)
	400GBASE-DR4		SMF		500 m
	400GBASE-FR8		SMF		2 km
	400GBASE-LR8		SMF		10 km

Figura 11. Padrões para aplicações de fibra para Ethernet (os padrões em progresso são mostrados de vermelho)

Os padrões da indústria oferecem as regras de design que podemos usar para determinar se a topologia de nosso data center cumprirá de maneira confiável com os requisitos de design das aplicações. Se observamos de forma retrospectiva os requisitos de design dos data centers, podemos avaliar cada topologia de links para determinar os

comprimentos máximos de links e as perdas de sinal máximas permitidas. Qual é a perda total a partir de toda a conectividade no link? Como se comparam a combinação entre perda e comprimento com os limites estabelecidos pelo padrão da aplicação? A comparação de cada caso com os padrões gerará uma decisão favorável ou desfavorável para nosso design.

Para determinar as perdas dos links do sistema é necessário compreender os componentes implementados. Estas características podem variar de um provedor a outro, e inclusive dentro de um determinado lote de produção. Claro que estamos interessados nos piores valores quanto à perda de inserção, para assegurar que não se superem os níveis de tolerância permitidos pelos equipamentos de networking. As fibras de alta largura de banda podem suportar links muito mais longos, enquanto as fibras de qualidade mais baixa requererão comprimentos mais curtos para funcionar de maneira confiável. Basear seu design nos padrões e nos dados sobre a performance dos componentes que o provedor oferece, deixa para você -o designer do sistema de cabeamento- a responsabilidade de realizar todos os cálculos por link.

Os requisitos de design subsequentes exigirão, pelo menos, que possa se suportar a próxima velocidade de rede mais alta com base na topologia do design inicial. Existem certas combinações a considerar.

Buscamos a perda máxima (não a média, nem a típica) que qualquer componente de cabeamento gerará para o link que estamos projetando. Também deve se considerar a largura de banda das fibras — o OM3 oferece menos largura de banda que o OM4, por exemplo. Podemos considerar a possibilidade de links multifibra paralelos no futuro. Finalmente, podemos considerar o impacto do tamanho e do crescimento do data center: como os requisitos quanto ao comprimento dos links limitam as opções que temos quanto a velocidades de redes de próxima geração?

Para realizar a análise dos dois cenários antes descritos com os componentes padrão se requerem os mais baixos valores quanto à perda de inserção para todos os componentes no canal. Neste exemplo, os módulos LC/MPO têm uma perda de inserção de 0,50 dB e os cabos dos troncais de fibra se classificam com uma perda de inserção de 3,5 dB/km. A suposição é de que os cabos de interconexão (patch cords) de fibras dúplex têm poucos metros de comprimento; portanto, não contribuirão de forma substancial para a perda de inserção geral.

Baseado nestes valores, a perda de inserção total é de 3,34 dB, o que supera a perda de inserção máxima de 2,19 dB para Fibre Channel 8G. Assim projetado, é provável que o link falhe ou experimente excessivos erros de bit.

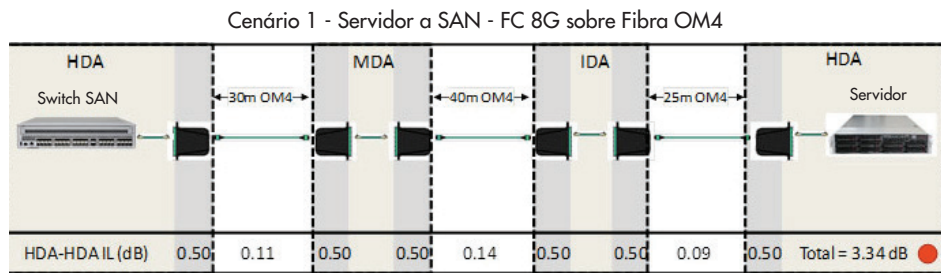


Figura 12. Cálculo estimado de perda para o link de servidor a SAN sobre fibra OM4

Uma análise similar é realizada para o segundo cenário, que inclui um link de servidor à rede Ethernet 10G que opera sobre uma fibra OM4 de 130 metros de comprimento. Neste cenário, a perda total estimada é de 2,39 dB, que está abaixo do limite de perda para esta aplicação sobre fibra OM4. Baseado nesta análise, o link deveria operar de forma apropriada.

Algo que normalmente acontece nas operações dos data centers é a inserção de conexões adicionais à medida que a rede cresce e os novos corredores de dados começam a funcionar. Neste exemplo, acrescentou-se uma conexão adicional à área IDA, o que aumenta o comprimento total do canal a 150 metros e adiciona mais dois módulos LC/MPO. Como se mostra abaixo, a nova perda total de inserção é agora de 3,53 dB, o que supera o valor máximo permitido. Este link, assim projetado, poderia falhar ou experimentar excessivos erros de bit.

Cenário 2 - Servidor a Rede: 10 GbE sobre fibras OM4

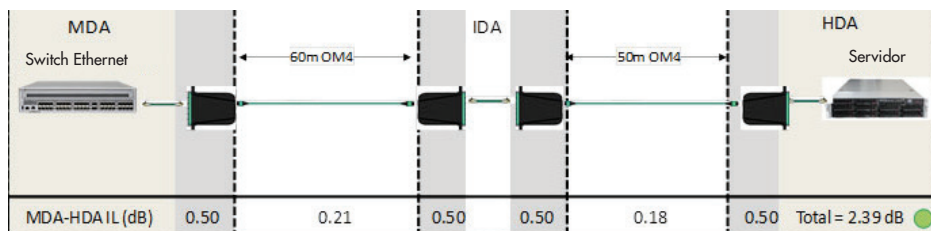


Figura 13. Cálculo estimado de perda para o link de servidor à rede sobre fibra OM4

Cenário 2a - Servidor a Rede - Acréscimo de trecho adicional

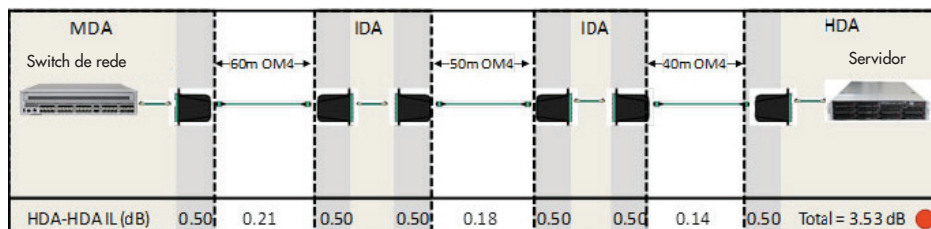


Figura 14. Cálculo estimado de perda para o link de servidor à rede sobre fibra OM4 - acrescentado de trecho adicional

Se considerarmos a possibilidade de melhorar este link, levando-o de 10 GbE a 40 GbE usando as óticas 40GBASE-SR4, o cálculo da perda de inserção é o que continua abaixo, conforme detalhado. Nota: a perda de inserção geral se reduziu ao mudar as óticas seriais de 10 GbE por óticas paralelas de 40 GbE, baseado na substituição dos módulos LC/MPO com painéis adaptadores MPO simples. No entanto, apesar da perda de inserção mais baixa, o link superou a perda geral estimada para 40GBASE-SR4 de 1,5 dB para aplicações sobre fibra OM4. Como tal, é provável que o link falhe ou experimente erros quando for atualizado a 40 GbE.

Cenário 2b - Servidor a Rede - Atualização a 40 GbE sobre Fibra OM4

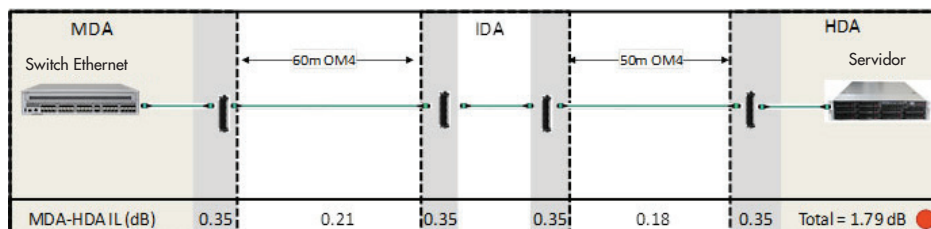
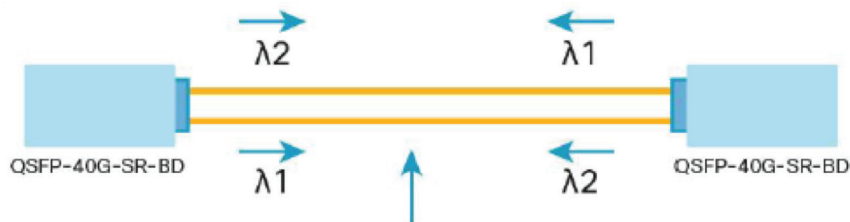


Figura 15. Cálculo estimado de perda para o link de servidor à rede sobre fibra OM4 - atualização a 40G

As ilustrações anteriores mostram os cálculos básicos considerando a perda a partir dos conectores, do cabeamento e dos limites de especificação. Nosso atual projeto de data center inclui uma ampla variedade de requerimentos de patchamento, de tal forma que devem ser consideradas todas as conexões e a soma dos comprimentos de cabos para as várias combinações de links que precisamos suportar.

Para responder estas perguntas, podemos assumir um enfoque simplista. Ao acrescentar as perdas totais e comparar os resultados com os requisitos padrão poderá surgir uma decisão favorável ou desfavorável. A repetição deste processo para cada topologia de links e tipo de aplicação nos oferece uma compreensão integral do data center em sua totalidade. Este processo leva tempo. Mude de provedor, por exemplo, e a perda máxima para cada componente poderia variar. Mesmo assim, outros provedores não fazem a cotação dos valores máximos — em troca, estimam valores de perda típicos, que não são úteis para este exercício.



Fibra multimodo dúplex com conectores LC em ambas as pontas

Figura 16. Transmissão bidirecional

Existem outras tecnologias patenteadas específicas conforme o provedor, que apresentam limites específicos para os links, mas que não são definidas pelos padrões da indústria. O esquema Cisco BiDi - QSFP-40G-SR-BD (que é ilustrado abaixo) é um exemplo de um novo design de links que utiliza dois comprimentos de onda a cada duas fibras, para uma capacidade total de 40 Gb/s. Neste caso, não existem limites padronizados com os quais comparar os designs dos links. O design dos links depende do que os provedores estipularem e está sujeito à informação de design que eles fornecerem para as várias topologias de cabeamento.

Existem várias opções, dada a volatilidade no design das redes e a grande quantidade de tipos de ativo a partir dos quais se pode optar. Projetar redes fabric baseando-se nos critérios aplicáveis aos links não é uma tarefa simples. Apesar de as redes fabric incluírem tolerância às falhas, os links físicos não deveriam permitir a introdução de nenhum sinal de risco. Optar por comprar componentes baseados nos padrões requer que o usuário final avalie os designs de links em geral e que depois determine se são adequados para tal fim. Não existem garantias do provedor de que o design do usuário final funcionará como se requer. Estes provedores apenas certificam a performance do componente — não a função geral do link.

Design da topologia fabric —suporte de aplicações

Os elementos prévios em matéria de capacidade, topologia de cabeamento, densidade, alcance e requerimentos de hardware de rede contribuem em seu conjunto para permitir um design de link ou aplicações de rede particulares. Manter as opções abertas significa considerar as permutas e combinações que fizerem sentido dentro de seu data center. A escolha de uma solução patenteada colocará limites a suas opções no futuro?

A CommScope desenvolveu o InstaPATCH 360, uma solução modular pronta para usar que permite todas as combinações de tipos de fibras, quantidade de canais e estratégias quanto à topologia para o funcionamento das redes fabric. Para apoiar ainda mais esta solução, a CommScope oferece suporte para o design das aplicações. Você pode determinar como oferecer suporte a qualquer aplicação de rede —baseadas nos padrões ou não— e adequá-la à topologia modular que seu data Center requer. Misture, combine e compare a performance dos custos do hardware de rede e o desempenho das soluções de forma rápida e fácil. Reduza os erros de design e se antecipe aos planos de capacidades futuras baseado no Application Assurance Guide (Guia de Garantia de Aplicações) da CommScope. A CommScope oferece suporte para trajetos distantes de links baseado no tipo de fibra, na quantidade de conectores e nas aplicações.

As seguintes tabelas ilustram o suporte garantido sobre links projetados para umas poucas opções de links fabric muito comuns, incluindo as aplicações descritas nos cenários 1, 2 e 2a antes citados. Baseado no seguinte quadro, o cenário 1 (Fibre Channel 8G sobre 95 metros de fibra LazrSPEED® 550 [OM4], com seis conectores MPO e seis conectores LC) seria plenamente compatível. Como se indica, esta topologia poderia operar sem inconvenientes em uma distância de até 150 metros.

Fibre Channel 8 Gigabits, Serial 850 nm “com receptor restritivo” (FC-PI-4 800-MX-SN)

Distância suportável em pés (m)

LazrSPEED 550 com Conexões LC

# Conexões LC com:	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
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)

Figura 17. Performance de Fibre Channel 8G sobre LazrSPEED 550

O 40GBASE-SR4 utiliza conectores MPO. Os transceivers baseados nos padrões e o correspondente alcance máximo com uma determinada topologia de cabeamento pode ser lido diretamente da tabela. Um link com seis conexões MPO pode ser configurado com no máximo 140 metros de cabos troncais OM4 LazrSPEED 550. Se isto for comparado com a tabela de 100GBASE-SR4, surge que o comprimento máximo do link é de 115 metros. Projetar o alcance desde o primeiro dia para no máximo 115 metros ofereceria uma rota de melhorias possível rumo aos 100G utilizando a mesma infraestrutura de cabeamento.

Ethernet 40 Gigabits, Paralelo 850 nm (40GBASE-SR4)

Distância suportável em pés (m)

LazrSPEED 550

# Conexões MPO	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
Distância em pés (m)	570 (175)	560 (170)	540 (165)	510 (155)	490 (150)	460 (140)

LazrSPEED 300

# Conexões MPO	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
Distância em pés (m)	460 (140)	440 (135)	430 (130)	410 (125)	390 (120)	380 (115)

Figura 18. Performance das aplicações —40GBASE-SR4 sobre fibras LazrSPEED

Ethernet 100 Gigabits, 4-pistas Paralelo 850 nm (100GBASE-SR4)

Distância suportável em pés (m)

LazrSPEED 550 WideBand e LAZRSPEED 550

# Conexões MPO	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
Distância em pés (m)	390 (120)	390 (120)	370 (114)	370 (114)	350 (108)	350 (108)

LazrSPEED 300

# Conexões MPO	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
Distância em pés (m)	280 (85)	280 (85)	280 (85)	260 (80)	260 (80)	250 (75)

Figura 19. Performance das aplicações -100GBASE-SR4 sobre fibras LazrSPEED

A garantia de aplicações da CommScope também se estende às opções de networking não padrão de provedores específicos. A tabela a seguir mostra a compatibilidade com os links projetados com relação à tecnologia Cisco BiDi 40G. As opções do CSR4 de alcance mais amplo também são mostradas abaixo. Uma comparação das duas opções oferece ao designer o alcance máximo para tais alternativas sobre fibras OM4 LazrSPEED 550 (150 metros para Cisco BiDi vs. 420 metros para 40GBASE-SR4).

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

Distância suportável em pés (m)

LazrSPEED 550 WideBand e LazrSPEED 550 com Conexões LC

# Conexões LC com:	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
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)

Ethernet 40 Gigabits, Alcance Estendido Paralelo 850 nm para Cisco (TRANSCEVEIRS QSFP-40G-CSR4)

Distância suportável em pés (m)

LazrSPEED 550

# Conexões MPO	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
Distância em pés (m)	1380 (420)	1310 (400)	1310 (400)	1310 (400)	1310 (400)	1310 (400)

LazrSPEED 300

# Conexões MPO	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
Distância em pés (m)	980 (300)	980 (300)	980 (300)	980 (300)	980 (300)	980 (300)

LazrSPEED 150

# Conexões MPO	1 MPO	2 MPOs	3 MPOs	4 MPOs	5 MPOs	6 MPOs
Distância em pés (m)	490 (150)	490 (150)	490 (150)	490 (150)	490 (150)	490 (150)

*A quantidade de conexões não inclui a conexão com o dispositivo ativo em cada ponta do canal

Figura 20. Performance das aplicações —aplicações Cisco

Uma revisão dos quatro cenários mostra uma comparação quanto à compatibilidade das aplicações e à garantia de cumprimento com os métodos padrão baseados nos componentes, que usa em análise manual, e com a performance das aplicações garantida pela CommScope. No caso dos designs baseados nos componentes, só um dos quatro cenários poderia cumprir com os requerimentos quanto a design e as estimativas de perdas. Usando o sistema InstaPATCH 360 da CommScope poderia ser cumprir com cada um dos quatro cenários de design.

Cenário	Aplicação	Comprimento Total	LC	MPO	Suportado com componentes padrão	Suporte da CommScope
1	8G Fibre Channel	95m	6	6	Não	Sim
2	10GbE	110m	4	4	Sim	Sim
2a	10GbE	150m	6	6	Não	Sim
2b	40GbE	110m	0	4	Não	Sim

Figura 21. Comparação de cenários

Links fabric —Opções de Ethernet

Opções de cabeamento monomodo, multimodo, paralelo ou dúplex — a opção correta para seu data center dependerá de seu tamanho, do ritmo de crescimento dos serviços que oferece, dos orçamentos para hardware, das tecnologias que os provedores oferecem e outras coisas. O design do cabeamento de rede deve se integrar com a arquitetura da rede, sua topologia e roadmap.

Considere as duas alternativas para links de Ethernet 40G (abaixo):

- 40GBASE-SR4, quatro pistas de 10G sobre oito fibras
- O design patenteado de fibras dúplex BiDi, que multiplexa duas pistas de 20G sobre um único par de fibras

O alcance que estas soluções oferecem é bastante diferente, como pode se observar se forem comparadas as tabelas de soluções segundo as aplicações para cada caso. A solução CSR4 pode chegar até os 420 metros sobre fibras OM4 LazrSPEED 550 —vs. os 150 metros para a solução BiDi. Estas distâncias e topologias que se mostram nas tabelas de aplicações se baseiam no uso da solução de fibras pré-terminadas InstaPATCH 360 da CommScope. Este exemplo permite aos designers comparar e projetar links com dois transceivers não padrão, enquanto também se compara o alcance e as capacidades da topologia com relação aos transceivers baseados nos padrões.

O sistema InstaPATCH 360 suporta transceivers 40GSR4 baseados em padrões através de seis conexões sobre uma distância de 140 metros de fibras, comparado com os requerimentos baseados nos padrões de 125 metros —muito maior alcance e flexibilidade da topologia para design de redes fabric. Se pensarmos a futuro e considerarmos o 100GSR4, existe compatibilidade para distâncias de 108 metros com seis conexões.

Links fabric —próximos passos

Antes, discutimos sobre as vantagens dos links fabric de alta capacidade (40G e superior) que permitem a mais servidores e dispositivos de armazenamento operar em conjunto sobre uma rede com uma capacidade geral superior. Enquanto hoje os links fabric de 40G são rentáveis e eficientes para muitos designs de redes fabric, provavelmente seja uma questão de tempo antes que as velocidades das portas de acesso para servidores subam a 25G, ou talvez a 50G. Nos próximos anos, os links fabric, obviamente, precisarão aumentar a 100G, ou talvez a 400G.

Se olharmos para frente, existe uma certa quantidade de opções para links de velocidade mais alta. Alguns provedores defendem as soluções ópticas monomodo. Outros defendem as ópticas multimodo. Para cada tipo de ativo que se escolher, existem potenciais opções de canais paralelos e dúplex. O custo relativo destas opções continua evoluindo rapidamente. Continuam se mantendo de pé algumas relações básicas — os sistemas ópticos multimodo continuam sendo os de custos de capital mais baixos e talvez mais fáceis de manter e operar, em comparação com os sistemas ópticos monomodo.

Links fabric — novas opções de ativo

As fibras multimodo OM3 e OM4 suportam links de 40G com um alcance e uma flexibilidade de topologia que satisfazem as necessidades de praticamente todos os data centers, menos aqueles mais amplos, do “tamanho de um armazém”. De cara ao futuro, se considerarmos os 100G e mais, um dos métodos mais prometedores para aumentar as capacidades da rede inclui acrescentar mais canais de comunicação a cada par dúplex de fibras multimodo. Enquanto a tecnologia de multiplexação por divisão de comprimento de onda (WDM) esteve disponível para ópticas monomodo por um alto custo, novos transceivers WDM de onda curta (SWDM) combinarão quatro canais sobre um par de fibras, ganhando assim quatro vezes mais capacidade sobre fibras multimodo. Esta alternativa de custo mais baixo combina a facilidade de instalação e operação das fibras multimodo com uma largura de banda que pode aumentar até suportar o crescimento que as redes fabric requererão no futuro.



Figura 22. Multiplexação por divisão de onda curta sobre MMF de banda larga

Em respaldo da tecnologia SWDM, a CommScope — junto com outros membros da Aliança SWDM— desenvolveram uma nova ativo de fibras multimodo “de banda larga”, ou WBMMF. Esta fibra foi projetada para expandir a capacidade disponível das ativos multimodo, permitindo mais canais de comunicação por fibra com um maior alcance. A WBMMF será usada para oferecer taxas de transmissão de dados mais altas, atingindo 100 GB/s e 400 Gb/s, enquanto se reduz a quantidade de fibras requeridas para suportar as futuras redes fabric de alta capacidade.

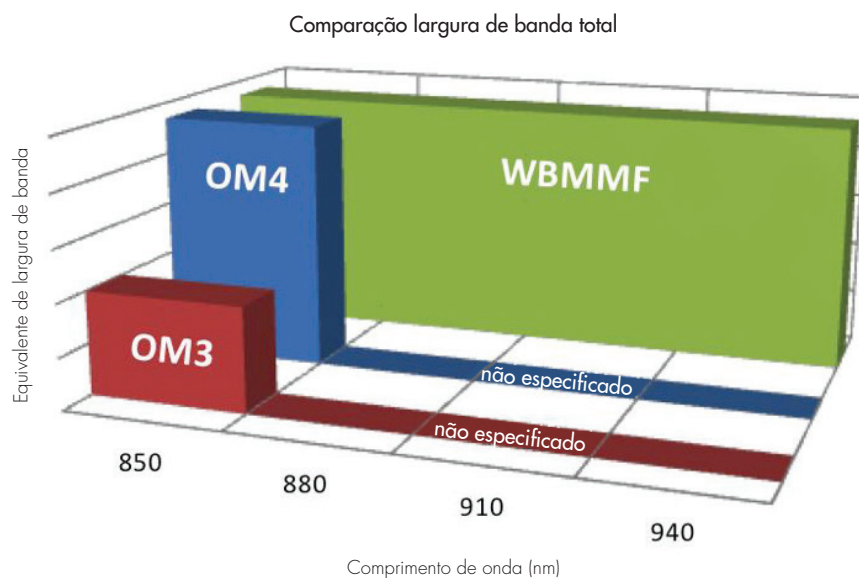


Figura 23. Uma comparação da largura de banda total entre OM3, OM4 e WBMMF

A WBMMF apresenta retrocompatibilidade com as fibras OM3 e OM4. A padronização da WBMMF progrediu rapidamente dentro do comitê TR-42 da Associação da Indústria das Telecomunicações (TIA), com a publicação da norma antecipada em 2016. A CommScope lançou a solução de cabeamento LazrSPEED 550 Wideband ao longo de todos os componentes da plataforma InstaPATCH — a via de próxima geração para redes de alta capacidade e baixo custo está disponível hoje.

Conclusões

Em resposta à demanda de maiores capacidades e custos mais baixos, os data centers estão se adaptando aos novos sistemas baseados nas redes fabric que permitam o funcionamento de sistemas informáticos e de armazenamento hospedados no cloud. As topologias de cabeamento dos data centers estão aumentando sua densidade para suportar links de comunicações multidirecionais de baixa latência normalmente requeridos pelas aplicações de cloud distribuídas.

O design de links de alta capacidade é mais complexo, já que o número de links de rede deve crescer e a velocidade da rede está em aumento. Prover mais capacidade ao data center significa pressionar até o extremo os ativos existentes e as tecnologias dos canais de comunicação. Os designs dos encaixes de fibras e a tecnologia WBMMF também estão evoluindo para prover capacidades de próxima geração e uma densidade física que se amolde perfeitamente às arquiteturas das redes fabric. As fibras monomodo são compatíveis com maiores comprimentos de canais.

O design das aplicações e as soluções de links fabricadas pela CommScope asseguram redes confiáveis de alta velocidade, que foram criadas para satisfazer as mais estritas demandas quanto a requisitos de capacidade da rede presente e futura. Os sistemas InstaPATCH 360 fornecem maior alcance para links de alta capacidade, liberdade de topologia de design para crescer de acordo com a magnitude dos ambientes complexos e performance garantida das aplicações, tanto para sistemas patenteados emergentes quanto para aqueles baseados nos padrões.

As soluções desenvolvidas fazem com que os complexos designs de redes fabric sejam simples de criar, implementar e manejar. Os sistemas pré-terminados de alta performance suportam ativos de redes de próxima geração e aplicações modulares dúplex e de múltiplas fibras, enquanto reduzem o tempo e os custos da gestão de sua implementação.



www.commscope.com

Visite nosso site ou fale com seu representante local da CommScope para obter mais informação.

© 2016 CommScope, Inc. Todos os direitos reservados.

Todas as marcas comerciais identificadas com ® ou ™ são marcas comerciais registradas ou marcas registradas, respectivamente, da CommScope, Inc. O presente documento será utilizado apenas para efeitos de planejamento e não é seu propósito modificar nem complementar nenhuma especificação ou garantia relativas aos produtos ou serviços da CommScope. A CommScope conta com as seguintes certificações: ISO 9001, TL 9000 e ISO 14001.

TP-110117.1-EN (04/16)