
更高速的布线

超越 10 GB 以太网的数据中心基础设施



在 2006 年 7 月，YouTube 报道每天有 1 亿次视频浏览。在 2010 年 5 月，每天的视频浏览次数攀升到 20 亿。在 2010 年 2 月，Twitter 报道每天有 5 千万 Tweets。同样在 2008 年 6 月，iTunes 有 50 亿的歌曲下载，而在 2010 年 2 月，下载量攀升到 100 亿。Facebook 有超过 5 亿的活跃用户，人们每个月在社交网站上花费 7 千亿分钟。有超过 1 百万个网站已经与 Facebook 平台结合在一起。这些应用以及其他一些类似的应用正在创建新的互联网和新的商务习惯。这些应用的核心就是数据中心，伴随着日益增长的高速带宽和存储的需求。

同样地，增长带宽的需求也成了企业数据中心的核​​心。短短五年前，没有人能够预测碳排放税，增加的标准和报告的需求，或者通过企业网络存储和移动的巨量信息产生的影响。通过多个服务器和存储设备共享一个或者两个网络连接的虚拟化是高速网络的驱动力。企业的运营需要多种广泛的软硬件平台和越来越多的合并协作、视频和其他先进的应用。IT 不再是一个不可避免的灾难，相反地是一个竞争优势。

不管是为了传送接下来的巨量应用，还是管理不断增长的即时数据的需求，数据中心正在提高速率，重新评估现有的应用，并通过合并和虚拟化的方式寻求更加绿色的处理。为了获得所需的更高的数据传输速率，新的标准和传输媒介已经应运而生。每一个应基于建筑式样、设计和技术优势、端到端的成本和性能考虑来评估。

IEEE® 40/100GbE 标准的正式颁布增加了高速传输的选项。同样地，由短距离双芯同轴和光纤组件支持的柜顶 (ToR) 交换机也提供了一个相对新的选项。然而，这些系统的设计与基于工业标准的结构化布线（使用两芯光纤的 10GBASE-SR/SX 和/或者使用四对双绞线铜缆的 10GBASE-T 系统）有着显著的区别。针对如今不断涌现的数据中心的高速传输，这本白皮书将提供一个概述，并且比较各种不同的选项。

从 10GbE 到 40/100GbE

从 40/100GbE 开始，IEEE 802.3ba 标准在 2010 年 7 月 17 日获得了高速任务小组的批准。最初的项目授权要求 (PAR) 是基于如下参数得到批准的，并在斜体字部分作了显著的修改：

- 在底版上传 1 米 (40GBASE-KR4)

-
- 在铜缆上传送 (40GBASE-CR4/40GBAE-CR10) 从 10 米减至 7 米*
 - 在 OM3 上传 100 米 (40GBASE-SR4/100GBASE-SR10)
 - 在 OM4 上传 150 米 (40GBASE-SR4/100GBASE-SR10)**
 - 在单模光纤上传 10 公里 40GBASE-LR4/100GBASE-LR10
 - 在单模光纤上传 40 公里 100GBASE-ER4

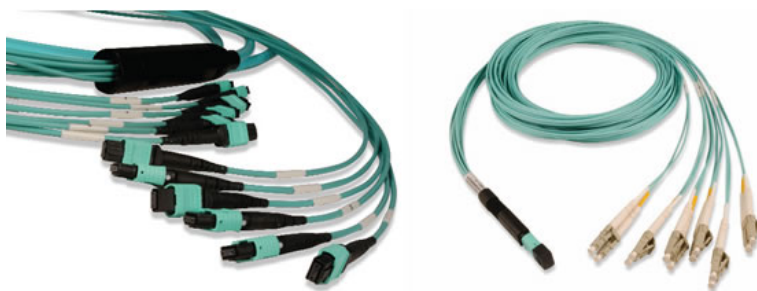
*在任务小组的标准开发过程中，对最初的 PAR 作了一些修改。双芯同轴的 CR4/CR10 距离被减至 7m.

**在标准开发过程中，一种新的光纤规格，即 OM4 被定义了。根据 TIA-492AAAD 和 ISO 11801:2002 Amendment 2, OM4 增加的带宽允许把支持距离扩展到 150m。由于 40/100GbE 标准是新的，OM4 已被包含在内。但是对于 10GBASE-SR/SX 应用，基于 IEEE 传输距离的扩展需要对 802.3an 标准作一个修订版。

更高速的光纤传输

当升迁至 40/100GbE，在主干和水平多模应用中最重要区别是光纤芯数的不同。40GBASE-SR4 使用 4 芯来传送和接受，共需 8 芯。100GBASE-SR10 使用 10 芯来传送和接受，共需 20 芯。SMF 还是保持 2 芯应用，虽然光纤相对便宜，但是 SMF 的光电设备却可能贵上 10 倍。在数据中心和主干，有可能会需要 8 芯或 20 芯光纤。然而，如果这些光纤芯数采用了不同的路径从一端到另一端，则可能会引起延迟偏移并导致比特误差。因为这个原因，40/100GbE 标准要求采用预端接光纤，并采用 MPO 或者 MTP® 多芯光纤接头。在这种情况下，所有的芯数会是相同的长度。参考"parallel optics," 这种结构将使延迟偏差减至最小，允许接受端的每芯光纤都可以同时接受到信息。

MPO (Multi-fiber push-on) 和 MTP (Mechanical Transfer Push-on) 是以 12 芯或者 24 芯的结构用在预端接光缆的两端。MTP 设计是性能更好的 MPO。有专利的 MTP 接头更加坚固，椭圆形，不锈钢排成一列的针的末端改善插入指导，并减少引导孔的磨损。MTP 接头还提供一个浮动的金属箍来保持外加负载时的物理接触，改善了机械性能。MPO/MTP 预端接主干缆同样也支持 10GBASE-SR/SX 的 2 芯应用。在这种情况下，预端接主干缆连至模块和/或者扇出跳线，这里把多芯光纤分成 2 芯的接头（通常是 LC 或者 SC）。



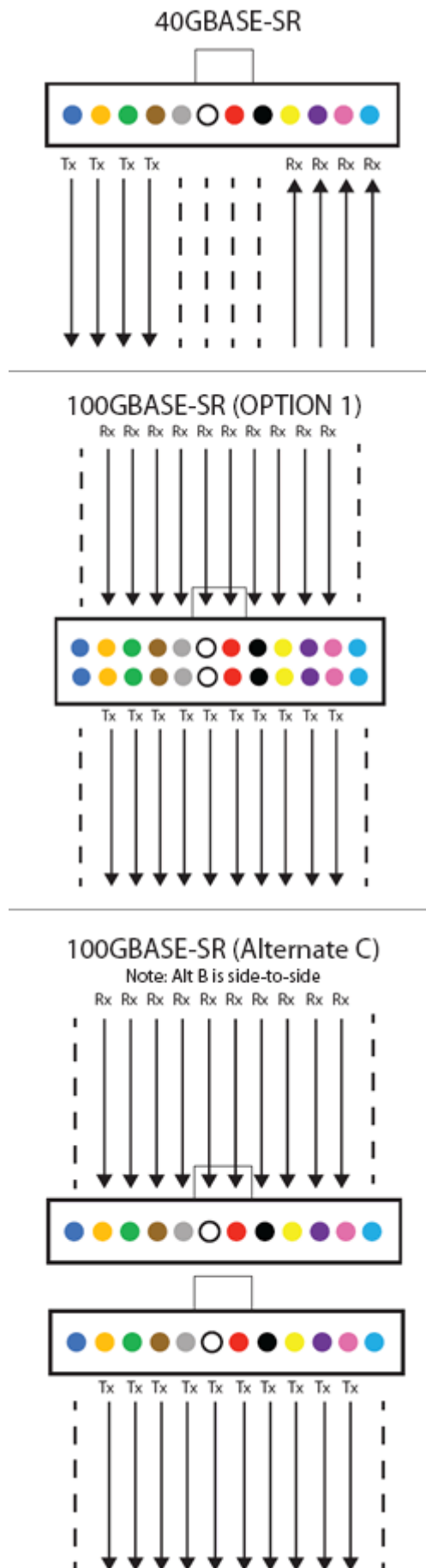
高速光纤结构的第二个不同是极性。对于像 10GbE 这样的 2 芯应用，管理极性是非常简单的，只需在信道的某个地方做一个反转（即从发送到接受）。因此

时信道是由单芯组成的。在 12 芯一组的预端接光缆，标准中有三种建议的极性（如下表所示）。

方式	主干缆	模块	跳线
A	两端一样	两端一样	某一端的极性反转
B	两端一样	某一端的极性反转	两端一样
C	某一端的极性反转	两端一样	两端一样

如上所述，当 2 芯应用迁移到多芯平行应用时，非常重要的一点是必须保证选用正确的极性以支持更高的应用。所有的极性方式都应该可以从 2 芯应用转换到 12 芯应用。

还有一点值得注意的是这些极性方式是标准建议的，而不是强制的。但是有一条强制规定：一个极性方式的建立和维护必须贯穿所有的光纤信道，使得每一芯的传送信号都能到达另一端的接受信号。这个规则适用于所有情况，包括多芯传送模式。为了使多芯应用的传送模式更加形象化，可参考如下图表：



40GBASE-SR 使用 12-芯 MPO/MTP 主干缆中的 8 芯(4 芯传送, 4 芯接受), 剩下 MPO/MTP 接头中间的 4 芯不用。设备上的接口将接受 MPO/MTP 接头, 而不是传统的 LC 接头。

100GbE 有 3 种经核准的传输模式包括一个 24 芯 (如左所示) 或者 2 个 12 芯主干, 使用"上和下"(如右图所示)或者 "肩并肩"。利用 10 芯传输, 10 芯接受, 剩下的外侧的 2 芯不用。还有一种可能是采用"Y"型集成, 即把 2 根 12 芯的主干转变到一个 24 芯的接头。不管采用何种方式和支持何种电子设备, 必须考虑正确的极性。

高速 InfiniBand (IB)

以太网也不是唯一使用平行传输的应用。InfiniBand^T 是一个工业标准的协议, 用于连接高性能计算集群到服务器, 交换机, 存储和嵌入式系统。根据 2010 年 11 月 6 号由 Infiniband 同业协会发布的公告, “InfiniBand 代表了前 500 个超级计算机中超过 43% 的系统。InfiniBand 连接在前 100 中占据 61%, 在前 200 中占据了 58%, 在前 300 中占据了 51%。” InfiniBand 的低延迟, 高品质服务, fabric 架构网络使得交换机, 点对点信道的数据传输能通过铜缆和光纤连接达到 120 Gb/s。InfiniBand 同业协会做了互操作性测试, 并在综合列表中提供了数百个产品的清单。

在直连结构中允许主机总线适配器(HBA)与目标通道适配器(TCA)直接通讯, 实质上是扩展了服务器或者存储设备的总线并创建了低延迟的吞吐量。虽然 InfiniBand 以 SDR, DDR 和 QDR 的数据传输速率运行在传统的 CX4 或 8470 铜缆和光纤组件上, 它采用了类似多通道通讯的 QSFP 连接器的 40/100GbE 连接机制。InfiniBand 当前支持多种接口, 在通道上操作, 每个通道等于两股导线 (一个传输, 一个接收)。

MPO/MTP 主干缆可以被连至当地的 InfiniBand 接口或者混合集成的缆线上。

目前的 InfiniBand 规范定义了三种数据传输速率。2.5Gbit/通道的单数据速率 (SDR)，5Gbit/通道的双倍数据速率 (DDR)，以及 10Gbit/通道的四倍数据率 (QDR)。由于编码，实际数据传输速率的吞吐量小于设定值。例如，接口采用 8/10 位编码，每 10 位能传输 8 位数据，剩下的 2 位是用来编码的，这比 Ethernet overhead 所要求的位数已经显著变少了。预计将在 2011 年发布两个新的 InfiniBand 规格，即十四倍数据传输速率 (FDR) 和增强数据速率 (EDR)。

更高速的 Fibre Channel (FC)

Fibre Channel 是高速传输的另一个协议，并已变成企业数据中心的存储区域网络(SAN) 的标准。Fibre Channel 同样也有铜的接口。Fibre Channel 和较新的允许使用以太网来传输的 Fibre Channel over Ethernet (FCoE)，已在存储中占据了主导地位。由国际信息技术标准委员会(INCITS)开发的标准，包括了存储、处理、转换、显示、管理组织和信息检索。自 1970 年以来，在 INCITS 里的技术委员会 T11 负责 Fibre Channel 的接口。FCoE 是在 2009 年 6 月 3 日正式获得批准的。

传统上，服务器使用不同的适配器来处理存储和网络流量- 主机总线适配器 (HBA) 运送存储流量，网络接口卡 (NICs) 处理局域网流量。对 FCoE 有一种混合的适配器，一个端口担当 Fibre Channel HBA 接受 Fibre Channel 的连接，另一个端口用于以太网。在 FCoE，其以太网接口与传统的铜和光的以太网接口一样。新的汇聚网络适配器 (CNA) 只需要单个的端口就能同时支持存储和网络流量。使用服务器汇聚网络适配器 CNAs 将减少所需的服务器适配器数量，因此就减少了 I/O 线缆的数量，和交换机端口的数量。这个构造减少了硬件资源，简化了服务器 I/O 配置，降低了电力消耗，并降低了总拥有成本。

Fibre Channel 的拓扑结构包括点对点，即用一根线缆组件连接两个设备，并支持以下距离：

不同类别线缆的支持速度和距离		
类型	速度	距离
OM2	1Gb/s	500m/1,640'
OM3	1Gb/s	500m/1,640'
OM2	2Gb/s	300m/900'
OM3	2Gb/s	500m/1,640'
OM2	4Gb/s	150m/492'

OM3	4Gb/s	270m/886'
OM2	8Gb/s	50m/1,64'
OM3	8Gb/s	150m/492'
Twinax copper	8Gb/s	15m max'

第二种 Fibre Channel 拓扑结构,是仲裁环路 (FC-AL), 可支持多达 127 个设备。但是更通用地, 是 Fibre Channel 的第三种拓扑结构 is deployed in a switched fabric。Fibre Channel Switched fabrics (FC-SW) 随着存储虚拟化的增长而变得流行。一个 switched fabric 结构包含一个大型安装中的 ToR 或者 EoR 的光通道交换, 和一个基于机箱中央交换的 SAN Director。

速度增至 8 Gb/s 的 Fibre Channel 正在开发中, 另外更快速的 16Gb/s 技术也有望在 2011 年发布。最近发布的 16GFC (16 Gigabit Fibre Channel) 是向下兼容的, 并在连接传统设备时会自适应向下到 4/8GFC 。

选择铜或光的 Fibre Channel 连接通常依赖于所需的速度、支持的长度和成本。当地的 16GFC 可在 OM3 光纤上运行 100m 以上, OM4 光纤上运行 125m。更长的距离可通过单模光纤来支持。SAN directors 产生了密集环境。为适应这个密度, 可使用如下图所示的扇出跳线 (把一个 MPO/MTP 接头的 12 芯分成 6 个双工 LC 接头)。八口的 MPO/MTP 适配器可以通过 MPO/MTP 主干连至光纤模块中的 48 个双工 LC 端口。

Fibre Channel 同样支持直连的 SFP+ 双芯铜轴组件, 可为短链路应用提供低功耗、有成本效益的选择。这些高速互连线缆将主要用于 ToR 柜顶交换机应用中。



支持 10Gb/s, 40Gb/s, 100Gb/s 和 Fibre Channel 应用的高速互联 (HSI) 铜缆跳线



2004 年颁布的第一个支持万兆以太网的铜缆接口是针对 10GBASE - CX4 (IEEE802.3ak) 应用而开发的。该接口符合 X2, Xenpack 和 XPack 多源协议 (MSAs 是相互竞争的制造商之间标准化一种接口形式的多源协议)。SFF- 8470 标准定义了这个屏蔽四通道铜接头的物理要求。作为一个无源组件, SFF-8470/CX4 电缆可达到 15 米。它支持万兆以太网、InfiniBand, FibreChannel 和 FCoE。还支持 I/O 功能的其他串行传输, 如 SATA(串行高级技术附加装置), SAS(串行连接 SCSI) 和 RapidIO。这些跳线使用的 Twinax 双轴电缆, 由屏蔽编织网整体包裹的两根内导体构成。双轴电缆通常用于短距离高速互连, 通过精确的焊接工艺在工厂完成端接, 并提供定制长度。双轴电缆导体的线径通常从支持短距离的 30 AWG(标称外径 6.1 毫米) 到长距离的 24 AWG(标称外径 11.0 毫米)。

由于其低延迟, 这些跳线多用于超级计算集群, 高性能计算和存储。作为 802.3ba 40GbE/100GbE 标准的一部分, 多通道 40GBASE- CR4 和 100GBASE - CR10 也已被定义。按照规定, 使用 4 和 10 通道的双轴电缆跳线, 可以在 7 米的距离上实现 40 和 100GbE 速率。

SFP+ 10Gb/s 跳线



SFP+ 10Gb / s 跳线是市场上的新人, 并同时支持铜缆和光纤格式。SFP +互连跳线支持以太网, InfiniBand 和 FiberChannel 协议。基于铜缆的 Twinax 双轴电缆跳线为 ToR 或短柜中 (MoR) 应用提供了短距离应用支持, 而 SFP+光模块则支持中长距离应用。 SFP 和 SFP+的接口规范由竞争厂家间的多源协议来定义。10Gb / s SFP+光模块、10Gb / s SFP+铜 (铜缆) 接口和插座的电气及机械规格经由 SFF 委员会制定的 SFF -8431 和 SFF-8083 (电气), SFF-8432 (机械) 和 SFF - 8472 (EEPROM) 来定义,。

用于 10Gb / s 应用的短距离直连 10Gb / s SFP+铜缆跳线正在日益流行。和 SFF-8470/CX4 跳线一样，这些产品采用 30AWG 至 24AWG 的同轴电缆，可定制长度。因此在设计上应有相同的走线通道考虑。10Gb / s SFP+无源铜跳线可以在 7 米距离内支持以太网应用，对于其它协议增长更长。实际支持距离根据制造商、导线尺寸和应用而定。

SFP+10Gb / s 的有源铜缆跳线利用芯片来放大信号，可用于提供比直连无源铜跳线更远的应用距离或更细的电缆。SFP+电缆的传输距离有限，适合用于 ToR 或相邻机柜的交换机、高性能计算机、服务器及存储集群等。虽然目前这些交换机端口的功耗低于 10GBASE - T 的铜口，必须注意端口的超额配置，以确保每端口功耗的节省没有被电子产品的成本（更多的交换机和电源）和随后的维护费用（代表日常支出）增加所抵消。应仔细核查的交换机端口，直连跳线和服务器网卡的整体成本。同时，整体功耗，包括高功率服务器网卡，应进行评估。当传统的结构化布线不可用时，这些跳线允许快速互连。随着 SFP+小尺寸连接器的问世，CX4 跳线通常用于铜缆 InfiniBand。在市场等待更省电的 10GBASE- T 交换机期间，SFP+连接器在 ToR 交换机中更受欢迎。



由于这些跳线通常只被限制用于同一机柜或相邻柜间，SFP+铜跳线可以帮助那些走线通道匮乏的数据中心减少通道挤塞情况。主流电子制造商在交换机上开发了更多功能，同时支持光纤和铜缆 SFP+适配器，使得任一媒质都可连入电子插槽。在大多数数据中心，人们会看到多种技术的混合（ToR 和结构化布线系统），提供基于应用需求的两全其美的技术运用。需要注意的是，虽然直连铜缆和光纤的接口形状相同，不同厂家之间的产品可能因主流电子设备制造商的加密方案而不兼容。有些加密方案需要使用专用的电缆，通常远远超过了通用成品跳线的价格。

QSFP+ (四通道增强型小型可热插拔接口)

更进一步发展的互连电缆是 QSFP+（四通道增强型小型可热插拔）跳线。支持与 SFP+类似的应用，这些四通道高速互连接口为最远 10 米的高密度应用设计，每通道的传输速率为 10Gb / s。一根 QSFP+跳线

可以取代多达 4 根标准的 SFP+连接，提供更大的密度，并降低系统成本。802.3ba 的 40GbE 标准认可 QSFP+连接器(又称 Style 1 ,SFF-8436 Rev 3.4)。对于 Style 1，多通道 40GBASE - CR4 指定的 QSFP+ 4 通道跳线支持距离可达 7 米。40Gb / s 的 QSFP +有源光缆，铜缆跳线和插座的电气及机械规格经由 SFF 委员会制定的 SFF- 8431（电气），SFF -8436（机械）和 SFF - 8472（EEPROM）来定义，SFF 委员会制定的规范。对于 100GbE 连接，通道数量会增加，连接器由 SFF-8642 Rev 2.4 定义。

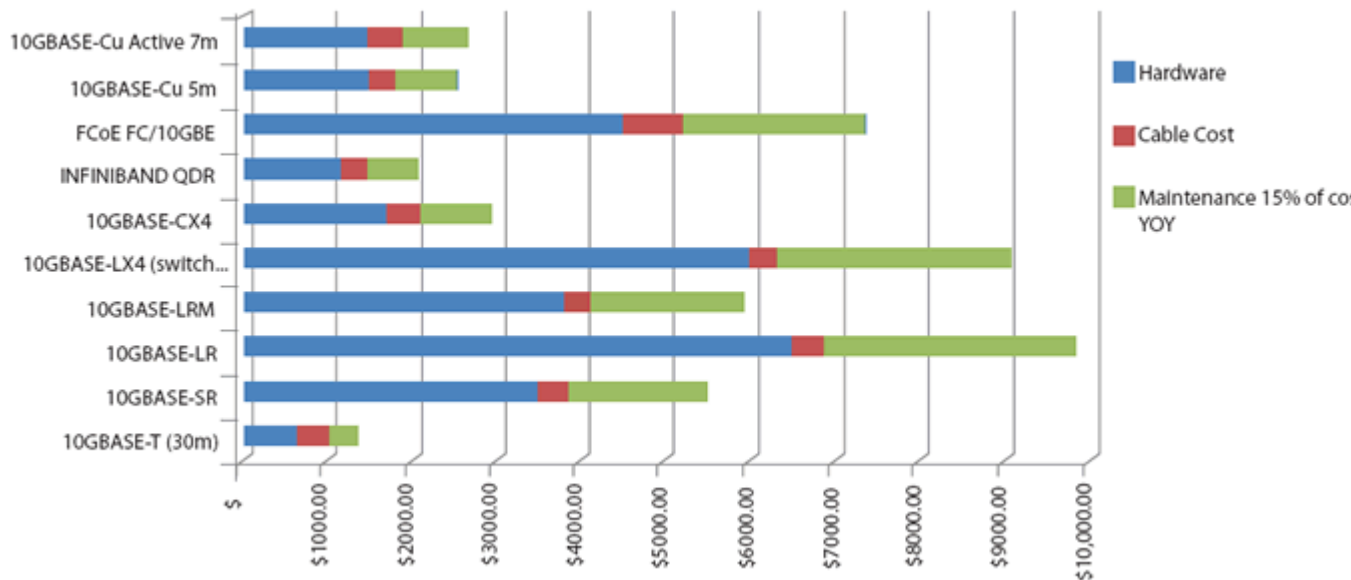
高速互连有源光缆跳线 (AOC)



由于点对点直连结构中的距离限制，在 ToR 交换机中通常会出现端口超额认购，在一个机柜中能支持的设备数量会受到电力和制冷的条件限制。为了在一列或者数据中心的增加交换机的端口使用率，有源光缆跳线(AOC)得到了开发。美国西蒙的 MorayT 线缆组件采用标准的小型 QSFP+ 接头，每个接头上均包含了四芯单模光纤的光电转换收发器，每一芯能支持 10 Gb/s 速率，每根跳线能支持 40Gb/s。Moray AOC 能支持到 4,000 米 (12,960 英尺)，其标准长度可达 300 米，用户定制长度可达 4,000 米。该跳线能支持 InfiniBand，以太网，Fibre Channel 和其他应用。

设计考虑

对每一种技术的评估应建立在互操作性、应用效益、未来可扩展性、维护成本和每端口总传输成本的基础上。这种分析应包括跳线、交换机端口和服务器或存储设备的网络接口卡的成本。虽然 TOR 解决方案具有低延迟，可降低水平铜缆成本的特点，但由于其电子设备的昂贵成本以及因此带来的高维护费用，大多数情况下，他们仅被运用在数据中心的某些特殊应用领域。考虑到特制跳线和有源设备的高成本，大部分数据中心仍考虑安装支持 10GBASE - T 的布线系统。英特尔®在 2011 年初供货的集成 10GBASE - T 芯片的主板，大幅度降低了电能消耗。节能型以太网 (IEEE 802.3az) 将在不久的将来进一步降低 10GBASE - T 端口的功耗，其原理是将闲置的端口设为“睡眠”模式（低功耗模式），从而降低每端口的净功耗。下表比较了各种互连和布线的万兆技术：



（上图是基于思科® MSRP，英特尔® MSRP 和布线系统的安装成本。价格以 11/30/10 为准，可能会随时间而改变。仅计算收发器模块成本，不包含交换机机箱）

在数据中心应用中，传统的铜缆信道被用于监控、集中式 KVM 和设备管理。而新的国际标准 ISO 24764 和即将颁布的 TIA 942-A 则建议在数据中心中安装 6A 类/EA 级铜缆布线来支持 10GBASE-T，在主干中应该使用预端接光纤 MPO/MTP 连接来支持未来的 40/100GbE 应用升级。直连跳线将根据不同应用的距离限制被用于单个机柜内或机柜列内。通过采用支持 10GBASE-T 的“集中式”布线设计，可以避免端口超配及昂贵的移动，增加和变化。同样，服务器若使用集成了 10GBASE-T 端口的主板就可以节省额外的网卡成本。另外非常重要的一点是，对于 SFP, SFP+ 和 CX4 模块，由于它们本身属于收发器模块，其质保期一般只有 90 天，与此相对应的交换机的端口有一年的质保期，而结构化布线系统一般至少承诺 20 年质保期。

通过利用不同的铜缆和光缆选项，可以设计一个强壮的、可升级的、灵活高效的数据中心来满足现在和未来的需求。美国西蒙提供最广泛的铜缆和光缆高速互连跳线、各种级别的线缆和线缆管理解决方案。想要了解更多的有关西蒙互连解决方案 (SIS)，可访问 www.siemon.com/sis。想要了解更多的有关西蒙铜缆、光缆以及管理解决方案，包括 Zero-U 配线的 VersaPOD 数据中心机柜解决方案，可以访问 www.siemon.com/us/versapod/。