

云数据中心互联中的 创新光纤和光缆解决方案

■ 纽约州康宁光纤通讯部市场技术开发经理 董浩

摘要：近年来，人们对于更高带宽、更快服务和更广泛接入的网络需求，不断推动着数据中心相关基础设施的显著进步；超大规模数据中心所采用的新型网络拓扑脊叶架构（Leaf-and-Spine）正在快速取代传统的三层网络模型。正是由于这些趋势，推动大芯数光缆迅速成为超大规模云数据中心园区 10km 范围内网络互联的必需品。另外，城域数据中心互联（Metro DCI）近年来也引起行业内广泛关注和讨论，在该应用领域，互联传输距离通常被限制在 80-120 公里以内，密集波分复用（DWDM）技术通常被优先采用以获取更高的频谱效率。在 Metro DCI 应用中，超低损耗光纤与基于 DWDM 的光传输模块配合使用，可大幅扩展网络覆盖能力，从而帮助用户将数据中心选址在土地价格低廉或可再生能源丰富的地区。本文所述研究重点说明，创新的光纤和光缆解决方案可以更好地为云数据中心互联提供服务，大幅降低联接成本，提高互联效率。

关键词：云数据中心互联；光模块；光纤；光缆；解决方案

一、云数据中心互联

目前，互联网已渗入到我们生活的方方面面，如电子商务、社交网络、在线游戏和视频流等，正深刻改变着人类社会面貌，并成为全球经济发展的基础。企业和政府机构也不断将关键的工作负载放入云端，以确保随时随地访问相关内容和服务。据思科全球云计算指数报告预计，随着虚拟现实和物联网等新兴应用的出现，全球数据中心流量还将在未来五年内增长近三倍。

面对数据中心通讯业务的爆炸式增长，分布式云数据中心成为主流云服务提供商的共同选择。在较早期的数据中心中，数据中心运营者往往通过添加服务器来实现数据中心的容量扩展，但是随着数据中心规模的迅速增加，这种线性的扩容方式不断受到来自供电、土地供应、容灾等方面的挑战。因此，数据中心运营商通常更倾向于在分布式的数据中心园区建造单独的数据中心单元，然后通过数据中心互联（DCI）将物理的分散数据中心单元融为一体（图 1）。这种分布式数据中心互联方式首先缓解了土地供应的限制，尤其是在洛杉矶或伦敦等人口稠密的大都市地区，该优势体现得尤为突出。更为重要的是，这种分

布式云数据中心互联方案要求多个变电站向不同的数据中心单元供电，从而提供了足够的供电冗余，极大地降低了电力故障对云数据中心可能带来的灾难性后果。在许多实际应用中，分布式云数据中心架构被云服务提供商广泛采用，用来应对越来越多的存储和互联业务所带来的巨大挑战。例如，在 2014 年度 AWS Re:invent 大会上，亚马逊透露，AWS（亚马逊云服务）单个数据中心单元的最佳规模为 25-32 兆瓦，服务器数量被限定在 5 万到 8 万台之间，当业务量增长计算能力出现瓶颈时，AWS 会单独追加建设一个新的数据中心单元。

因此，针对分布式云数据中心的结构特征，云数据中心互联需要提供低成本、高带宽和低延迟的 DCI 解决方案，从而融合物理的分散数据中心单元，有效构建虚拟化云数据中心平台。顾名思义，DCI 就是连接云数据中心园区内部、园区之间，甚至远距离内的多个数据中心单元，连接长度可为数百米至数百公里不等。ACG Research（2015 年下半年全球光学数据中心互联预测，2016 年 4 月）研究表明，全球高达 70% 的 DCI 连接距离为 80 公里或以下，因此，超长距离的数据中心互联不在本文的讨论范围之内。

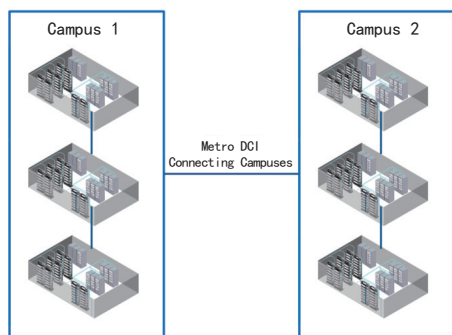


图 1 分布式数据中心架构

二、DCI 光收发模块技术

DCI 链路可以获得的最大传输容量取决于可用的光收发模块技术，在过去几年中，光收发模块迅速向大容量、小型化、低功耗、低成本的方向发展。尽管目前许多数据中心所采用的光收发模块传输速率仍为 40Gb/s 或以下，但 100G 技术正迅速成为 DCI 各个应用层面上的主力技术。同时，多个国际标准组织也已经开始酝酿 400G 传输解决方案，例如 IEEE 协会 P802.3bs 以太网工作组已经公布了 400G 模块的标准，该标准针对连接距离 10km 以内的 400G 以太网光收发模块，详细描述模块物理层面光传输规范。针对不同的应用，高速 DCI 光收发模块的价格相差甚大。总体而言，数据速率更高、传输距离更长的收发模块价格通常更高。图 2 列出了多种现有和新兴 100G、400G DCI 收发器及其市场估价。估价基于公众论坛中的数据或专业咨询公司过去 18 个月公布的数据。另外，我们还对光学部件（如激光器和接收器）的数量、带宽、调制方式和波长网格（如 CWDM 与 DWDM）进行了评估，将这一数据与这些光收发器技术实施中的差异关联在一起。值得注意的是，基于 DWDM 技术的 100G PAM4 收发模块有可能后来居上，取代一部分现有 Metro DCI 光收发模块市场（相干或者灰光模块）。尤其是针对某些数据中心需要租赁光纤资源的情况，100G PAM4 可以减少物理上相互分离的园区之间互联的运营支出（OPEX）。100G PAM4 光收发模块之所以具有较高成本优势，是因为其采用了直接检测技术，具有较高的频谱调制效率（每符号 2 比特数），同时对电芯片的带宽要求相对较低。100G PAM4 光收发技术能否在 Metro DCI 领域取得进一步成功，未来几年将非常关键，其模块成本走势以及市场接受程度值得我们持续关注。

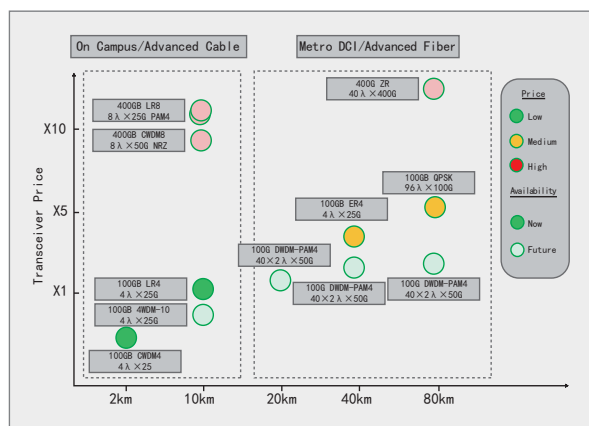


图 2 100G 和 400G DCI 收发器估计相对定价

图 2 所示的光收发模块价格 - 传输距离分析表明，在不久的将来，数据中心运营商仍会继续选择更为经济实惠的 1310nm 工作波长的 100G 光收发模块，支持距离在 10 公里以内的传输。这些收发器需要一对专用光纤才能实现 100 Gb/s 的双向容量，而进一步增加容量需要相应增加光纤，因此云数据中心园区内互联所需光纤量会随着通讯业务量的增加而持续增长。

对于 400G 以太网光收发模块，由于传输速率的提高，特定传输容量所需光纤数量会降低——这是因为短距 400G 以太网光收发模块普遍采用 PAM4 调制方式，具有更高的光纤利用效率。但是，400G 光收发模块小型化及封装标准并未完成，大规模商用尚需时日，未来 400G 光收发模块技术的成功将取决于供应商资源多样化程度以及模块性价比走势。同时模块技术的进步与通讯业务量增长之间存在时间差，鉴于数据中心通讯业务量的爆炸式增长，我们预计在未来 400G 技术得到普及的同时，光纤的需求量仍将持续增加，以适应 DCI 总体传输容量的更高速增长。

根据以上光收发模块综合分析，我们得出如下结论：

- 在 20-80 公里范围内，基于 DWDM 技术的 100G PAM4 有可能取代现有的 100G ER4 或 100G 相干 QPSK 解决方案。
- 10 公里的 DCI 应用领域将出现多种技术标准以及多源协议（MSA）之间的相互竞争：基于 IEEE 标准的 100G-LR4 将继续大量存在，但会有更多数据中心用户考虑采用基于 CWDM 的光收发模块产品。
- 在可预见的将来，低成本 100G CWDM4 收发器将继续主导 2km 应用市场。

·同时，设备供应商以及模块厂商正在积极研发更高速率的解决方案，以满足未来云数据中心互联的需求，例如针对 Metro DCI 的应用，许多模块厂家已经展开单载波 400G 相干解决方案的研发。尽管目前该类光收发模块成本仍然居高不下，但随着技术的成熟，未来价格一定会回归到市场可以接受的范围。

通过对光收发模块的分析不难看出，针对不同的应用场景，数据中心运营商会灵活选择不同的光收发技术，因此对光纤光缆也会相应提出不同的要求。创新的光纤光缆产品可以为云数据中心园区内（2-10 公里）和 Metro DCI（20-80 公里）互联提供高性价比的解决方案。

三、数据中心园区内 DCI 应用的创新型光缆

目前，许多云数据中心园区互联使用 864 光纤数（864FC）或 1728FC 光缆，而更高芯数的光缆例如 3456FC 光缆也正在越来越多地被云服务商会所采用。这一现象主要由光网络层面的两大趋势所驱动，一方面，机器对机器通信所带动的“东西向”网络通讯业务成指数倍增长；另一方面，网络架构由于采用了脊叶架构（Leaf-and-Spine）而变得扁平，并易于扩展。

传统上，数据中心采用三层拓扑架构，包括核心路由器，汇聚路由器和接入交换机。传统的三层拓扑架构虽然已经成熟并且被广泛部署，但已逐渐不能满足现代化数据中心的通讯需求——尤其是云数据中心日益增长的工作负载和低时延需求。作为回应，今天的超大规模数据中心网络拓扑结构正在向 Leaf-and-Spine 转变。在 Leaf-and-Spine 中，网络可被划分为两个层面：“Spine”层面用于汇聚和发送数据包到最终目的地，“leaf”层面用于连接终端服务器和负载均衡。在理想的情况下，每个“leaf”交换机扇出到每个“spine”交换机，从而实现最大限度地将服务器互联为一个有机的整体，因此，脊叶架构的数据中心网络中需要大量高基（high-Radix）的光交换机作为硬件支撑。

由于网络架构趋于扁平，并且采用了高基交换机，数据中心网络正在变得更大，更加模块化，同时更易于扩展。例如，对于一个采用了架顶式（TOR）加 Leaf-and-Spine 架构的数据中心网络，64 基（64-Radix）光交换机理论上可支持 65,000 台服务器，而 64 基光交换机基

至可支持多达 50 万台服务器，以及 260 万组光纤互连，不难看出，这正是驱动云数据中心园区内互联需要大量光纤资源的深层原因。

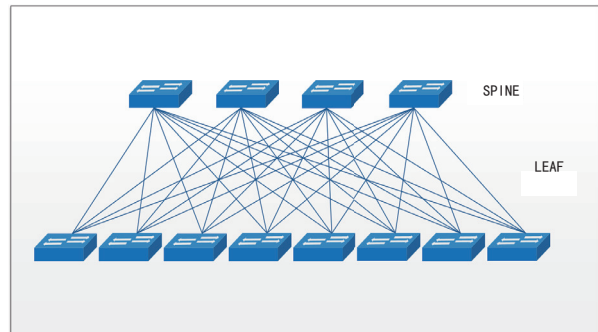


图 3 Leaf-and-Spine 网络架构以及高基光交换机导致网络结构中需要大量的光纤互连

而铺设于云数据中心园区内互联的光缆所面临的另一个限制条件便是其能否匹配现有管道。在北美地区，这些管道的内径通常为 2 英寸（约 50mm）。无论采用什么收发技术，均需满足这一条件。为了平衡云数据中心园区中大量的光纤互连需求和管道内径限制之间的矛盾，业界不得不考虑增加 DCI 应用光缆的光纤密度，光缆制造商正生产更高光纤数和更小直径的高密度光缆，以应对这些挑战。

近年来，光缆行业提出了 1728FC 和 3456FC 等创新的高芯数光缆设计来满足云服务商会的需求，例如，包括康宁在内的少数光缆企业已经开始提供 3456 芯的高密度光缆产品，这些光缆外径通常不超过 34 毫米，兼具高密度、便于安装和操作性好等优点，适用于云数据中心园区内 DCI 应用。

四、Metro DCI 应用的创新光纤

针对分布式的云数据中心结构，云服务商会往往需要在城域范围内进行数据中心园区之间点到点的互联，即 Metro DCI。在 Metro DCI 的应用领域，1310nm 的低成本光收发模块通常不足以支持所需的传输距离。因此，业界通常选择使用基于 DWDM 的相干通讯解决方案，其中，相干 400G 目前成为行业广泛关注的一个热点。

针对这一趋势，OIF 已经启动了适用于 Metro DCI 的 400G-ZR 项目。在众多技术解决方案中，基于 PM-64QAM 和 PM-16QAM 的单载波 400G 相干方案是最具代表性的两种，图 4 比较了 PM-64QAM 和 PM-16QAM

两种方案的主要特性。考虑到符号速率、带宽要求、频谱效率和最大传输距离等因素之间的平衡，在不久前的 OIF 会议上，OIF 宣布支持 PM-16QAM 成为 400G-ZR 项目所采用的标准调制格式。



Modulation	PM-64QAM	PM-16QAM
Constellation		
Spectral Efficiency (bits/Hz)	8	5.3
Channel Spacing (GHz)	50	75
Single-λ line rate (bits)	400G	400G
Baud Rate	42GBaud	63GBaud

图 4 两种单载波 400G 解决方案的比较
PM-64QAM 与 PM-16QAM

在 Metro DCI 应用中，针对目前所广泛采用的 100G 相干传输，往往只需铺设常规的单模光纤（0.20 dB / km 左右的光纤衰减）即可满足大部分数据中心园区间互联的需要，这是因为 100G 相干（QPSK）光传输系统具有足够色散容限和偏振模容限，因此无需考虑线路传输色度色散和偏振模色散影响，从而给网络建设和运维带来更多便利。所以即使在日渐老化的光纤设备上也能基本保证足够的传输性能。然而，随着 400G 时代即将到来，情况会发生很大改变——由于调制密度的增加导致光信噪比（OSNR）必须更高，以致于传输能力会大幅度下降。例如 16QAM 的 400G-ZR 相干传输系统中，传输距离会比 100G DP-QPSK 的情况骤减 10 倍。因此，在 400G 单载波的 Metro DCI 应用场景中，需要考虑通过采用新型光纤来解决传输距离受限的问题。

我们采用高斯噪声模型进行了计算，结果表明，若采用衰减为 0.20dB/km 的常规 G.652 单模光纤，400G-ZR 单载波传输的单跨段(无中继)传输距离仅为 70km。然而，若选用衰减为 0.17dB/km 的超低损耗光纤，其传输距离可延长至 90km 左右。

在 Metro DCI 应用中，单跨段传输距离的提升有利

于帮助用户将数据中心选址在土地价格低廉，或可再生能源丰富的地区。新数据中心或次级数据中心选址通常需要考虑两个主要因素，即用地成本和能源情况。需说明的是，目前新建云级数据中心可能需要数公顷用地，约相当于几个足球场的面积。数据中心选址用地成本每低 100 美元 / 平方米，则每公顷可节约成本 100 万美元。

另一方面，可再生能源发电成本较低，且在许多地区可享受税收优惠，因而可再生能源越来越被数据中心运营商所青睐。事实上，低成本的可再生能源具有高度可用性，例如越来越多的企业选择在美国太平洋西北地区（水力发电）、冰岛（地热发电）和瑞典（水力和风力发电）建立数据中心。

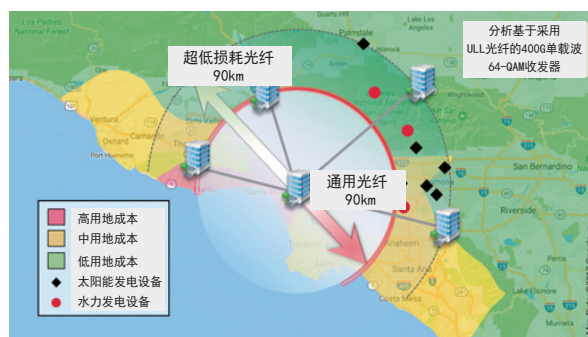


图 5 超低损耗光纤扩大数据中心选址范围

例如，一个位于美国西海岸洛杉矶地区的分布式云数据中心，其主数据中心位于洛杉矶市区，采用超低损耗光纤可使次级数据中心的选址半径扩大 20km，从而可设于土地成本更低的乡郊地区（见图 5）。另外，用超低损耗光纤还可扩大覆盖范围 50%，使新数据中心具有更大的选址范围。以洛杉矶地区为例，可再生能源供电站基本上均位于距市中心 70km 以外的区域。因此，若采用超低损耗光纤，运营商便可靠近可再生能源建立数据中心，并实现清洁能源的既定目标。

综上所述，随着全球云数据中心通讯业务的继续增长，各主要经济体必将建立更多的云数据中心，且这些数据中心的互联将扩大对大容量链路的需求。光纤和光缆的创新必将推进云服务发展及其应用，而 DCI 和光收发技术的发展也必将为全球化时代更深层次的互联互通做出贡献。