

400G 的兴起将颠覆数据中心格局

作者：Tony Campbell, Molex



数据中心运营商及数据中心供应链将在迈向 400G 的道路上坚持到底。

更高的以太网速度、云计算、物联网以及虚拟数据中心都向数据中心的运营商提出了更多的要求。超大规模数据中心的运营商正在推动着更加广泛的采用 100G 的链路和模块技术。与此同时，400G 的形状系数和光学模块已经处于全面启动的关键点上，预计将在 2019 年逐步展开。数据中心行业内部的这一转变将以惊人的速度使久经考验的 QSFP28（四分之一小形状系数可插拔 28G）模块密度增加一倍，最高可使带宽增加三倍，与四个 100G 的模块相比，400G 模块的总功耗甚至更低。

用于网络交换机的 56G PAM-4 ASIC（专用集成电路）芯片在不断发展，功能日益强大，用户包括博通、Innovium、Nephos 和 Barefoot Networks 之类的企业，因此对于下一代光学互连系统和模块的需求在持续保持增长。

这些新型的 ASIC 可以提供 12.8 Tbps 的带宽，从而下使一代的交换机可提供 32 个 400 Gbps 的端口。或者，如果数据中心的架构需要更高的基数，那么这些 ASIC 可以在可逆变速箱模式下运行，从而提供 128 个 100 Gbps 的端口。传统的 OEM，比如说思科和 Arista，以及智邦科技、QCT 和天弘之类的白盒制造商都在争先恐后的生产这些速度更高的交换机，其中很多都已经上市。随着 400G 的交换机已准备就绪，关键的一点在于，光缆和铜缆互连系统也需要符合规格要求。现货供应可为实际的部署提供支持。

400G 的主要推动力

有哪些因素在推动着新的需求？根据 IDC 的统计，数据中心的存储要求每年上升超过 50%，预计数字信息截至 2020 年将增至 40 泽字节，截至 2025 年则将增至 163 泽字节。对于这种增长，存在着几个关键的贡献因素，包括一波向云存储、开放系统、边缘计算、机器学习、深度学习以及人工智能的转型。

虚拟现实才刚刚开始大范围的受欢迎，并且，在可以预见的未来，无人驾驶汽车的现实情况是将成为主流，将会为数据中心基础设施带来指数级的压力。

为了满足带宽需求，一般超大规模数据中心会每两年对整体网络构架进行一次升级，因此有计划的报废是无法避免的情况。

数据中心的供应链已经加快了发展速度，研发前所未有的、功能更加强大、能效更高的可扩展解决方案迫在眉睫。目前，100G 的技术可以为以太网链路提供速度最快的连接。100G 和 400G 以太网技术的同步实施，在今后的数年内将会继续增长，而后者最终会占据优势地位，成为交换芯片和网络平台上普遍采用的速度。

发展路径

当我们展望未来时，可以看到怎样一种情况？一系列的数据中心基础设施解决方案，给人留下如此深刻的印象，其设计可满足扩张中的超大规模数据中心对于提高带宽与功率的要求。下一代的解决方案充分利用了铜缆和光缆，具有极高的信号完整性并且降低了延迟和插入损耗，从而达到最高的效率、速度与密度。

现有的铜缆（DAC）已经能够达到 400G 的速度，而能够实现 400G 交换连接的光端机则正变得日趋完善，准备用于全面发布。独立的 100G Lambda 光端机和 400G 光端机当前正处于贝塔取样阶段，很快即将投入市场。由于早期的使用者将需要更高的带宽来部署这些产品，400G 将从 2019 年的年中到年底间开始升温，即使供应链成本尚未降低下来、价格侵蚀并未开始，情况也将如此。

很多的数据中心将会继续在更长的链路传输距离上部署 100G CWDM4 光端机，而对于 100G PSM4 的需求正在快速消失，这类供应商也在退出市场。随着独立的 100G Lambda（100G-DR 或 100G-FR）光端机产品已经在 2019 年初开始提供，由于其预期价格较低，因此预计这些产品将抢占 100G CWDM4 的市场，此外，独立的 Lambda 产品还能够在分支拓扑结构中实现与 400G 光端机的直接互操作。

随着带宽迁移到更高的水平，行业将继续见证 10G 和 40G 技术的逐步淘汰，这些技术已经被光端机、直连线缆（DAC）和有源光缆（AOC）所取代，后者在一

系列数据中心之间以及数据中心内部的通信中支持 100G、200G、400G 及更高的速度。QSFP-DD（四分之一小形状系数可插拔双密度）光端机在这一螺旋式上升的过程中将坚持扮演重要的角色。

发展情况

QSFP-DD 光端机配备了一个八通道的电气接口，每个通道的数据传输速率能够达到 50G。QSFP-DD 模块可实现高达 20 瓦的功率（按照 QSFP-DD MSA 第 5.0 版），在创新性的散热器功能的帮助下，可以在一系列的传输距离内达到 400G 的性能。这一点非常重要，由于先进的 ASIC 消耗的功率更多、发散的热量也更多，而 QSFP-DD 的形状系数则可以通过行之有效的热管理策略，高效的耗散掉这些热量。

OSFP（八进制小形状系数可插拔）的形状系数更宽、更深，也为 400G 提供支持。QSFP-DD 光端机与 OSFP 相比，一个主要的优势就是完全向下兼容现有的 QSFP+ 和 QSFP28 光端机。56G PAM-4 技术被广泛认为是实现 QSFP-DD 和 OSFP 形状系数的光端机的关键所在。现在正在引入整合了 QSFP-DD 和 OSFP 光学模块形状系数的平台，在云应用中为 400G 以太网提供支持。这些新的平台提供与 100G 端口的向下的兼容性，在数据中心或企业中可以实现交错的实施方式。

在不考虑形状系数的情况下，400G 的光端机要求使用 DSP 的“变速箱”，从八条 50G 通道中创建四条 100G 的光信道。这在供应链中将成为一个关键的组成部分，在光端机供应商的能力方面发挥重要的作用，使其提供相应的产品与产量来应对数据中心消费者提出的巨大需求。随着光端机供应商对产品差异化的追求，对于能量需求较低的 7 纳米 DSP 来说，其在 2019 年的供货能力将进一步的对这一供应链造成干扰。

Molex 已经演示过了符合 100G Lambda MSA 要求的 100G FR QSFP28 和 400G DR4 QSFP-DD 产品。下一代网络设备的技术生态系统将推广 112G PAM-4，将其作为一个基础来为大容量的数据中心支持 400G 的解决方案。MSA 的规范强调了在采用每波长 100G 的 PAM-4 技术时，为实现光学接口而在技术设计上遇到的挑战，以及不同供应商的互操作性。PAM-4 技术可实现 100G 的光学通道，传输距离为 2 至 10 公里，而对于 400G 的通道来说，在双工单模光纤上则可达到 2 公里的传输距离。PAM-4 平台可以有效的奠定初步的基础，以高成本效益的方式完成向 400G 的迁移。该技术平台聚合起了四条每波长 100G 的通道，可以为分支应用支持 400G 的版本，例如 400G DR4、400G FR4 和 4X100G 等。

400G 的演进

现代的通信网络需要更高的带宽来满足全球范围内数据爆炸提出的要求。如此一来，数据中心交换机和光端机市场正在飞速的增长和演化。对于构建 400G 的网络设备来说，高速光端机、高度灵活而又可以扩展的光传输产品、紧凑型连接器以及光纤管理都属于至关重要的能力，从而服务于大量的电信提供商、企业以及超大规模数据中心。

对 400G 及更高速度下的数据中心光纤管理进行评估，这非常重要，而 Molex 光纤聚合设备之类的产品可以提供高效的解决方案，适用于高光纤密度的系统，实现有组织的光纤管理。这类产品可以减少或者消除不通的信道，提供一个无源的交换位置，极其紧凑而又无需电源或冷却。它们还可以在当前的 LC 双工插拔和下一代的 MPO 高密度连接器解决方案之间架设起桥梁，弥补连接上的差距。这方面的一个例子就是一处原先配有 LC 双工光纤设备的数据中心，在 100G 下采用了 CWDM4 光端机，但是该中心现在正在迁移到 400G 下的 DR4，需要并行的光纤基础设施。

超越 400G

数据中心交换机的 ASIC 供应商已经宣布，56G PAM-4 12.8Tbps ASIC 现在可以普遍供货，并且当前正在研发 112G 的 PAM-4 25.6Tbps ASIC，能够推动 32 端口交换机的发展，每个端口可支持 800 Gbps 的速度。然后，ASIC 的这一功能将产生一系列的挑战，与信号完整性、热、功率和损耗等等主题有关，此处略举数例，同时还会提出这样一个问题，那就是互连系统是否可以或者是否应当继续做到模块化，并且，如果是这样的话，那么哪种形状系数实际上可以为此提供支持。

随着数据中心的运营商正在制定计划从而以高成本效益的方式快速的实现扩展，通过与那些具备当今的数据中心需要的能力、专家经验及可扩展性的供应商开展紧密的合作，100G 和 400G 基础设施的设计可以得到优化。实施过程需要良好的安排，从而协调数百或者数千个组件之间的数据传输，实现最优的数据中心结构，在未来减轻整体风险并满足各种动态的需求。

