

400G 全场景光模块白皮书

400G 时代的到来

5G 时代见证了新服务的出现，4K VR、物联网（IoT）和云计算对网络的带宽、并发率和实时性提出了更高的要求。实际上，网络流量的年均增长率仍保持在 26% 的高水平。



图 1-1 网络流量快速增长

100GE 技术在 2013 年出现，标志着该行业的技术转折。因为 100GE 端口无法满足流量爆发所带来的不断增长的传输需求。如果将多个 100G 链路捆绑在一起，光纤资源将滞后于带宽资源的增长，并且可能会出现负载不平衡的情况。因此，需要具有更高带宽能力的高性能端口。于是随着越来越多的网络中使用具有更高速率的相干传输技术，用于远程传输的 100G 端口需要升级。

根据 LightCounting 在 2019 年 10 月的市场预测报告，以太网和 WDM/OTN 侧的 400G 光互连端口的增长率远高于其他端口。

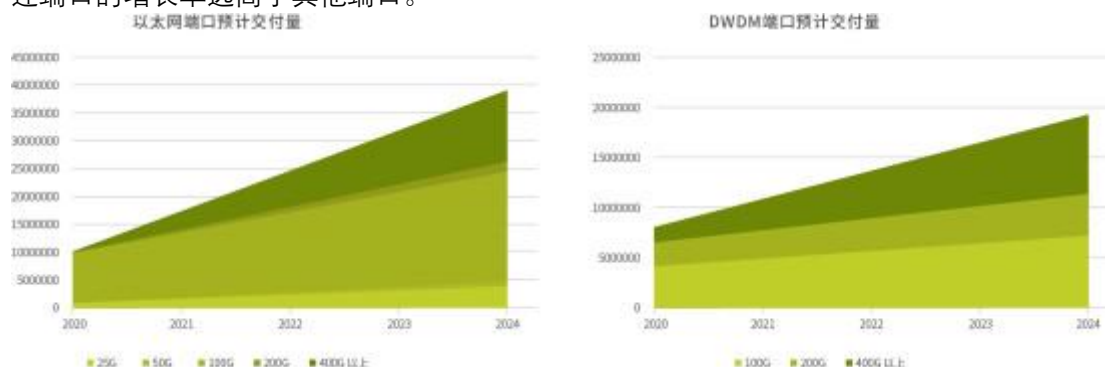


图 1-2 不同速率端口的发货量预测

400G 是下一代主流端口技术，具有显著提高网络带宽和连接利用率的能力，可帮助运营商和 OTT 客户有效应对数据流量的爆炸性增长。

400G 光互连应用场景

通信网络标准正在不断发展以支持 400G，三个国际标准组织（IEEE，ITU 和 OIF）已经发布了 400G 标准。IEEE802.3bs（400GE 标准基准）已于 2015 年初完成，最终版本于 2017 年 12 月完成。与此同时，ITU-T SG15 着重研究了与光传输网络（OTN）相关的标准，ITU-T 已完成了超越 100G OTN 标准化的定义。

自 2013 年以来，光纤互联网络论坛（OIF）一直在从客户端到线路端发布有关 400G 互连的白皮书。

随着行业标准和网络要求的日趋成熟，400GE 引领了 ICT 行业的发展，尤其在以下三种情况下获得了越来越多的关注和应用前景：数据中心、城域承载网、长距离大容量传输网络。

400G 数据中心网络光模块解决方案

背景：东西向流量的快速增长

到 2021 年，约 70% 的数据中心东西向流量将留在数据中心内，预计增长率将大大高于南北流量和数据中心之间的流量。

云计算的渗透导致传统的数据中心被云数据中心取代，大大提高了对高速光模块的需求。

- 传统数据中心使用 10G 低速光模块，而云数据中心主要使用 100G 高速光模块。

- 传统数据中心主要传输南北向流量，而云数据中心主要传输东西向流量，从而驱动了大量的东西向连接，并显著增加了单个服务器上光模块的使用率。因此，主要依赖东西向流量的服务（例如机器学习和人工智能）需要更多的计算和存储功能。

当前，云计算数据中心已完全采用 25G/100G 光互连解决方案。然而，在 2018 年下半年，北美的大型 OTT 客户，导致了向 50G/400G 光互连解决方案的转换。同时，已有公司计划在 2020 年开始部署 400G 解决方案。从 2019 年上半年开始，中国 OTT 客户也开始了 400G 研究，在中国发起了涉及多个 400G 光模块供应商的首次 400G 测试。

客户类型项目	传统企业数据中心	中国的云数据中心	北美的云数据中心
规模	小	中/大	大
网络架构	传统的三层网络架构	脊/叶网络架构	脊/叶网络架构
流量与方向	小 主要是南北向流量	中/大 主要是东西向流量	大 主要是东西向流量
每台服务器的光模块数量	少	中/多	多
光模块速率	1G/10G	25G/100G	25G/100G→50G/400G

表 3-1 不同类型数据中心的比较

趋势：单位成本和功耗更低

通常，用户体验会随应用程序场景而变化。在长距离 WDM 场景中，用户对性能更敏感，期望更长的传输距离和更高的频谱效率。相反，考虑到距离，体积和功耗，短距离数据中心方案中的用户对成本更加敏感。

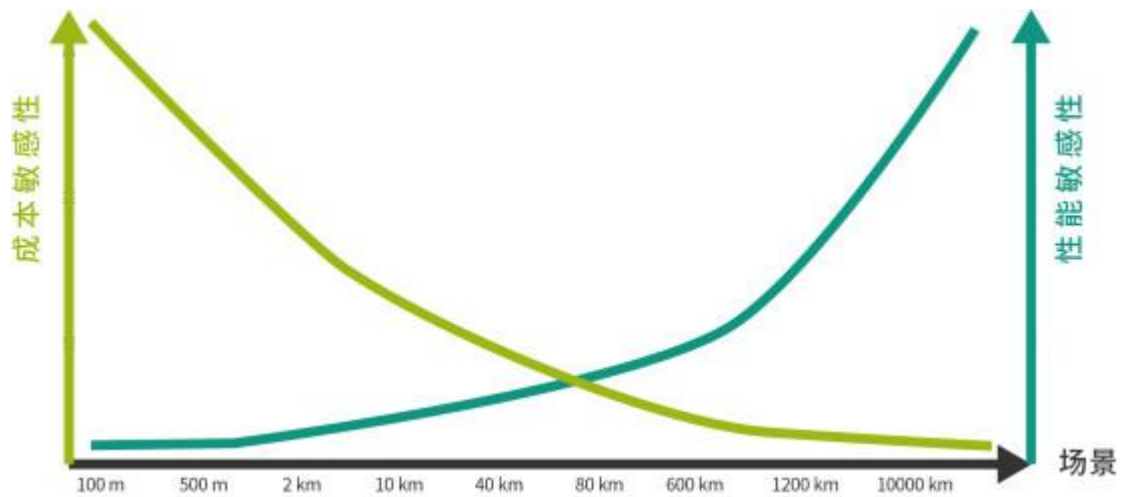


图 3-1 城域综合承载网架构

英特尔创始人戈登·摩尔（Gordon Moore）提出了摩尔定律，他预测微芯片上的晶体管数量每 24 个月翻一番，这在过去半个世纪中使半导体行业大为受益。还有一个光学摩尔定律适用于光电领域，该定律指出，短距离光学模块每四年发展到下一代，从而使单位比特成本和功耗减少一半。

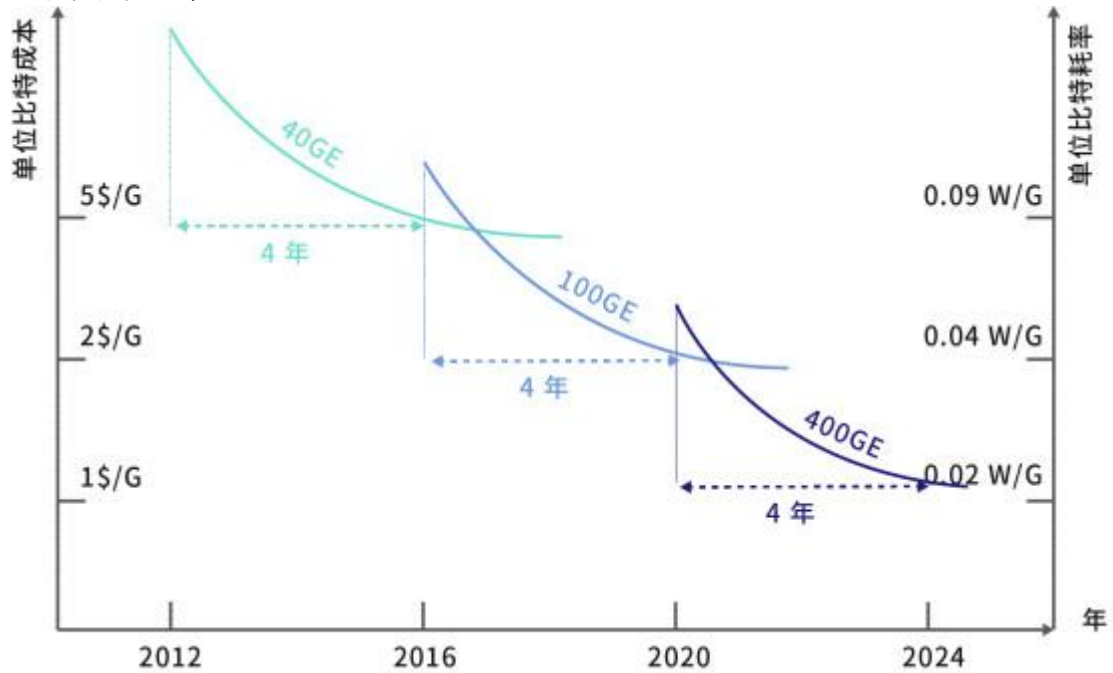


图 3-2 数据中心网络中的光学摩尔定律

通过三种方法，光模块可以达到更高的速率以满足光学摩尔定律所描述的要求：提高光学组件的速率（更高的波特率），使用更多的通道或通过高阶调制降低单位比特的传输成本。

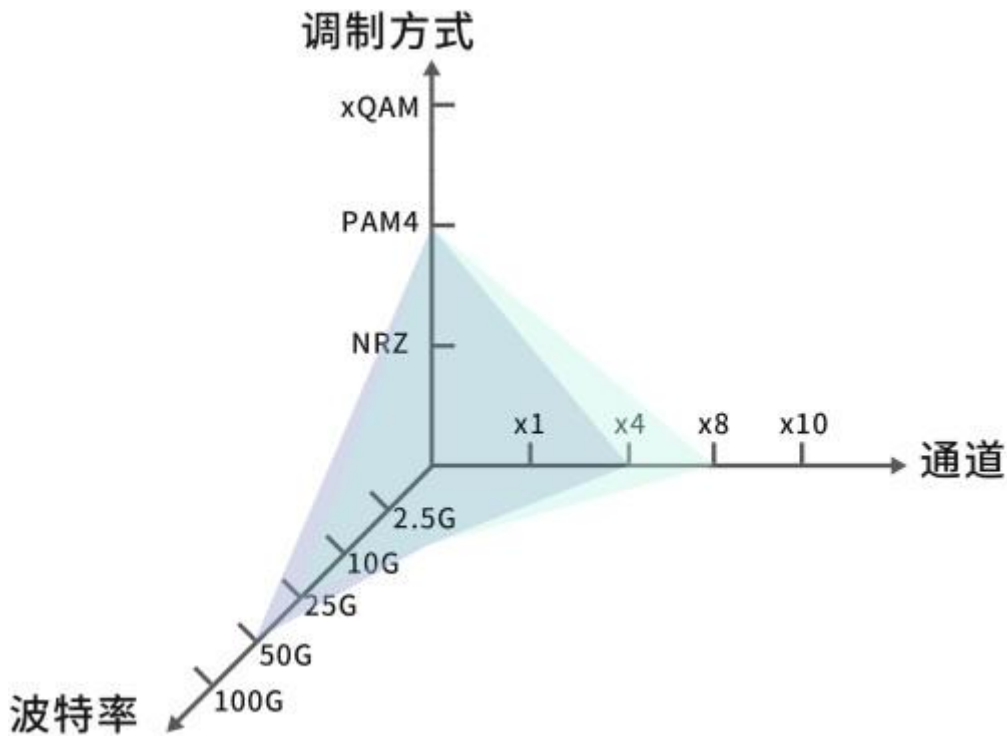


图 3-3 提高数据中心网络光模块速率的技术方向

●使用 PAM4 高阶调制

4 级脉冲幅度调制 (PAM4) 技术效率更高, 可以有效提高带宽利用率。PAM 是不归零 (NRZ) 编码之后最流行的信号传输技术。

NRZ 信号使用高电平和低电平来表示数字逻辑信号中的 1 和 0, 并且可以在每个时钟周期中传输一位逻辑信息。相反, PAM4 信号使用四个信号电平进行传输, 并且可以在每个时钟周期中传输两位逻辑信息 (00、01、10 和 11)。因此, 在相同的波特率下, PAM4 信号的比特率是 NRZ 信号比特率的两倍, 使其成为 400G 模块的最佳选择, 以提高传输效率并降低传输成本。

●使用更多通道

根据历史数据, 具有八个以上通道 (例如 x10 和 x16) 的解决方案导致较低良率和较低可靠性, 而多通道 (x4 或 x8) 架构则更具成本效益, 并降低了功耗。

100G (4x25G) CWDM4 和 100G SR4 模块已成为上一代数据中心光互连的主流解决方案。

●使用波特率更高的光电芯片

基于包括直接调制激光器 (DML) 和垂直腔面发射激光器 (VCSEL) 在内的 25Gbaud/s 光学芯片产业链, 数据中心 100G 光学模块采用 NRZ 信号和 x4 架构来实现。包括 DML, 集成了电吸收调制激光器 (EML) 和 VCSEL 在内的各种 25Gbaud/s 光学芯片正在不断发展, 以支持更高的 56Gbaud/s 波特率。当前, 有 56Gbaud/s EML 产业链可用, 而 56Gbaud/s DML 和 VCSEL 芯片仍在研究中。

场景	100G解决方案	400G解决方案
TOR到叶(100 m)	100G SR4 4路:25Gbaud VCSEL+NRZ	400G SR8 8路:25Gbaud VCSEL+PAM4
叶到脊(500 m)	100G PSM4/CWDM4 4路:25Gbaud SiP/DML+NRZ	400G DR4 4路:56Gbaud EML/SiP+PAM4
叶到脊(2 km)	100G CWDM4 4路:25Gbaud DML+NRZ	400G FR4/DR4+ 4路:56Gbaud EML/SiP+ PAM4

表 3-2 数据中心对光电芯片要求

用于城域网的 400G 光模块解决方案

背景：新服务推动 5G 时代的流量增长

随着 5G 的快速发展，包括 4K VR、物联网、超高清视频、VR、增强现实（AR）和车联网（IoV）在内的新服务对承载网络提出了以下新要求：超高带宽，多连接通道，超低延迟和高可靠性。由于城域网的 100GE 端口无法支持需要超宽带的汇聚层和核心层，因此基于 400GE 的光互连解决方案是必要的。

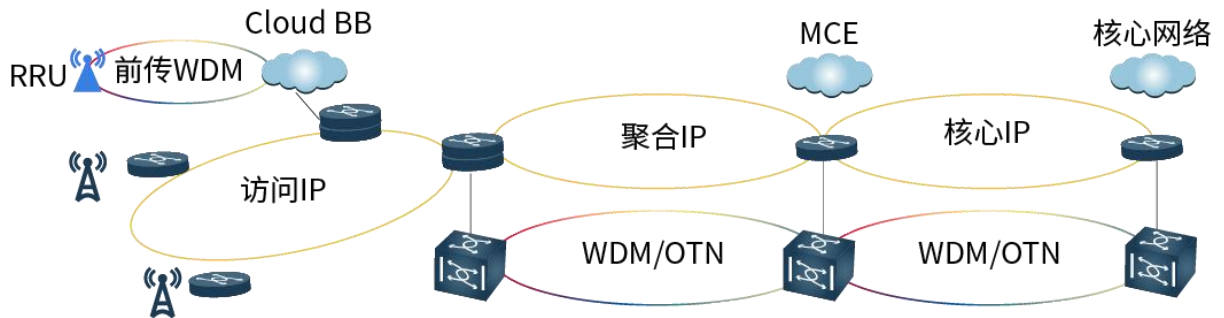


图 4-1 5G 承载网络架构

以 5G 承载网为例。根据下一代移动网络（NGMN）联盟的带宽评估，在城域网大规模 5G 商业化期间，接入层带宽将发展到 50GE，而核心层带宽将发展到 200GE/400GE。

网络层项目	城域接入层		城域汇聚层	城域核心层
	5G前传	5G中传	5G回传+DCI	5G回传+DCI
距离	<10/20km	<40km	40-80km	40-80km
光模块	10G/25G灰光 Nx25G WDM彩光	25G/50G灰光 Nx25G/50Gb/s WDM彩光	100G/200G灰光 Nx100G WDM彩光	200G/400G灰光 Nx100G/200G/400G WDM彩光

表 4-1 5G 承载网光模块带宽要求

同时，数据中心正在从骨干层迁移到城域网的边缘。因此，数据中心流量的快速增长提出了更高的带宽要求，从而加速了数据中心互连接口的带宽从 100G 升级到 400G。

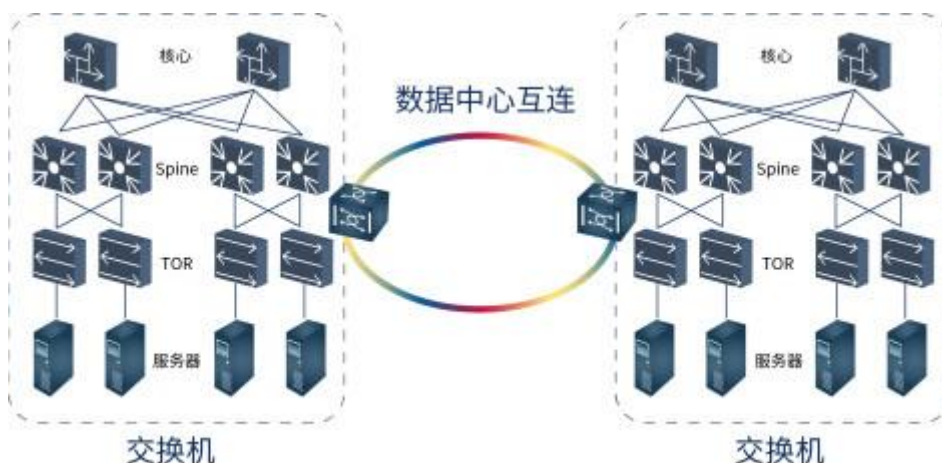


图 4-2 数据中心互连组网图

趋势：单位比特成本更低，可靠性更高，距离更长

在 5G 时代，预计 5G 站点将比 4G 站点多 1.5 到 2 倍，而城域网中使用的光模块数量有望达到数千万。光模块成本与设备成本之比也在不断增加，导致运营商在网络建设投资中优先考虑较低的光模块成本。运营商也迫切需要 400GE 光模块仅需要 100GE 光模块的功耗和单位比特成本的一半。另外，载波承载网络比数据中心具有更高的可靠性和性能要求。

顾客 关键指标	操作员	数据中心
生命周期	> 10年	3-5年

表 4-2 不同场景下光模块的生命周期要求

与传输距离短的数据中心网络相反，传输距离长的城域集成承载网对光模块的传输性能有更高的要求。

场景	100G解决方案
DCN	100m-2km
城域承载网	10-80km
DCI	80-120km

表 4-3 不同场景下的传输距离

城域集成承载网采用数据中心组网场景中的技术，以实现 400G 的更高传输速率，降低单位比特的传输成本。

- 提高光学元件的波特率

可以通过增加光学组件的比率来提高线路容量，但是由于 III-V 半导体激光器的性能瓶颈，该方法受到限制。根据城域网的要求，已批量应用了 25Gbps 的光学组件（DML 和 EML）。但是，在 56Gbps 级别，仅 EML 可用。

●更多通道

鉴于光学组件的带宽和性能改进受到限制，可以并行使用多个通道来实现 100G 传输，如 4x25G NRZ 技术所示。25G 组件的成熟产业链可以重复使用，以提前提供大容量解决方案，从而满足实际应用需求。在 400G 时代，多通道解决方案将是必不可少的，考虑到成本和功耗，x4 或 x8 多通道架构是最合适的。

●使用高阶调制

当使用高阶调制格式时，可以在不增加信号波特率的情况下提高频谱效率，从而提高总承载率。当前，PAM4 具有 25G 的发射器和接收器，是业界主流的调制方式，因为它有助于实现单通道 50G 传输速率，从而降低了单位比特功耗和成本。

鉴于城域集成承载网络对光模块的可靠性和性能要求很高，因此一直受业界高度重视。

●使用更可靠的组件

在城域集成承载网中，光模块主要应用于运营商级场景。它们要求具有 10 年的生命周期，在 0°C 至 70°C 的温度范围内能充分工作，并采用密封包装以确保其可靠性。降低成本对于数据中心模块变得越来越重要，从而推动了非密封封装技术的发展。但是，此类技术在电信级场景中仍面临以下挑战：

激光二极管（LD）的可靠性风险：非密封封装对 LD 端面的涂层要求更高。

热电冷却器（TEC）的冷凝和化学腐蚀风险：LAN 波分复用（LWDM）组件使用 TEC 进行温度控制。结果，LWDM 组件的某些部分（TEC 所在的位置）的温度低于环境温度，在非密封环境中存在凝露的风险，长期暴露于水蒸气中可能会导致电化学腐蚀。

光路污染风险：在非密封封装和 LD 芯片与光口之间的空间中，非密封包装无法避免以下问题：由于凝结和污染，光表面功率可能会发生变化；并且组件内部的胶水可能会吸水并导致光路的相对位置发生变化，这也会改变光功率输出。

●使用高性能 LWDM 发射器

粗波分复用（CWDM）波长间隔为 20nm，并且不需要 TEC 进行冷却，从而大大降低了成本。CWDM 发射器支持 2km 的应用，并且是数据中心方案的主流解决方案。与 CWDM 相比，LWDM 具有 5nm 的波长间隔，并具有低色散损失和更高的传输性能。因此，它是城域网中运营商级方案的首选。以 100G/通道解决方案为例。传输 10km 之后，CWDM 的色散窗口和色散损失（2.5dB）远大于 LWDM（1dB）。在 400G 时代，CWDM 无法支持 10km 以上的远程传输，因此 LWDM 发射器将成为城域集成承载网络的主流解决方案。

●将高性能接收器与雪崩光电二极管（APD）结合使用

通常，25G 发射器的输出功率为 0-3dBm，25G PIN 接收器的灵敏度约为 -7dBm，无法满足电信级光模块的长距离（40km）传输要求。因此，需要高性能的 APD 来提高接收器的灵敏度。

	灵敏度	接收器	成本
400G LR8	-6.6dBm	PIN	低
400G ER8	-16.1dBm	APD	高

表 4-4 光模块的光接口规格

●引入相干技术

受组件性能的限制，400G PAM4 解决方案不支持 80km 以上的远程传输。为了克服这一挑战，应该引入已经成功部署在长距离传输网络上的成熟的相干技术。此外，SiP 和 InP 集成

技术以及互补金属氧化物半导体（CMOS）技术不断发展，以支持具有更低功耗的更小的相干光模块。借助紧凑的 ICT/ICR 和低功耗 7nm oDSP，可以实现 400GE 紧凑的 QSFP-DD 相干光模块。2017 年，OIF 400ZR 项目启动，该项目定义了串联前向纠错（CFEC）模式和低复杂度 400G-16QAM 调制模式。性能和功耗都被认为可以支持 80-120km 的城域网和数据中心应用。

400G ZR 具有低功耗和紧凑封装，可用于城域边缘访问场景。但是，原始 400G ZR 定义的性能无法满足城域边缘访问场景的要求。为了解决这个问题，一些光模块供应商正在采用基于 400G ZR 算法和 FEC 算法的高性能算法，以满足 200-300km 的传输要求。但是，更高的性能会导致更高的功耗，进而对系统的散热造成压力。作为响应，一些供应商在 ZR+ 方案中采用 CFP2 封装尺寸。

光模块	调制	光学元件波特率	副载波数	传输距离	模块包装	符合标准
400GE-LR8	PAM4	25Gbaud	8	10km	QSFP-DD	802.3bs
400GE-ER8	PAM4	25Gbaud	8	40km	QSFP-DD	802.3cn
400GE-LR4-6	PAM4	56Gbaud	4	6km	QSFP-DD	802.3cu
400G ZR	PDM-16QAM	61Gbaud	1	30km (无EDFA)	QSFP-DD	OIF 400ZR
				120km (EDFA)	CFP2	/
				120km (EDFA)		

表 4-5 不同传输距离的光模块规格

DWDM 网络的 400G 光模块解决方案

背景：流量增长为长距离传输增加带宽压力

网络流量的增长导致传输网络上端口带宽的增加。对于长距离和高带宽传输，基于波分复用器（WDM）的相干传输技术提供了最佳解决方案。

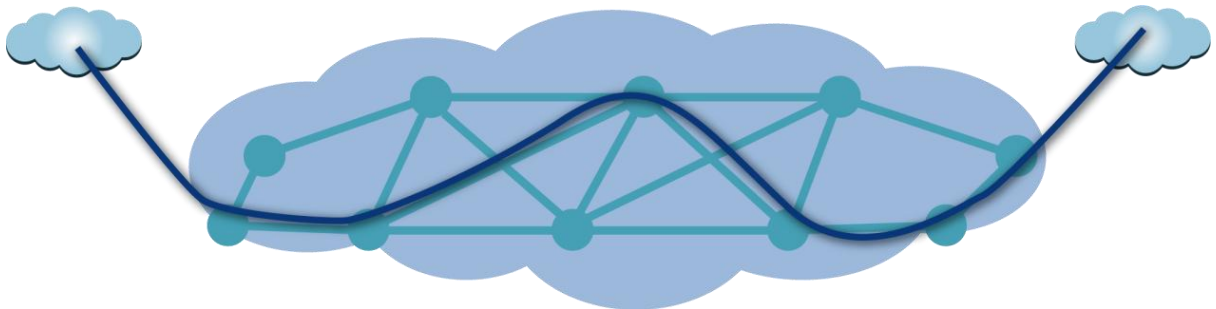


图 5-1 远程传输图

随着 400G 一致性解决方案的成熟，到 2020 年以后，对 400G 一致性端口的需求将迅速增长。推动 400G 一致性端口增长的动力有两个：网络带宽的增长和客户端 400GE 端口数量的增加。事实证明，使用一个 400G 波长来承载 400GE 服务是最经济高效的方法。根据 LightCounting 的预测报告，将在越来越多的网络中使用 400G 相干端口，并将在未来 5 年内看到最快的增长。

随着网络流量，波长总数以及一个网络中波长的持续增加，网络运营商还将增加网络管理和调度的灵活性要求，从而促进可重配置光分插的大规模部署。多路复用器（ROADM）和光交叉连接（OXC）。借助波长选择交换（WSS）技术，运营商可以根据需要动态配置波长路径，从而通过光路实现点对点连接，并减少等待时间和功耗。由于这些好处，越来越多的运营商开始采用这种解决方案。例如，2017年，中国的运营商在长江中下游建造了ROADM网络，该网络具有多达364个波长。

灵活的速率调制和灵活的网格技术使DWDM网络更灵活，更具弹性，而传统的DWDM系统使用固定的50/100GHz栅格，中心频率和通道宽度。如果可以使用弹性调制和栅格技术，则可以根据容量和传输距离自定义每个端口的调制格式和通道宽度，从而提高频谱效率和传输容量。

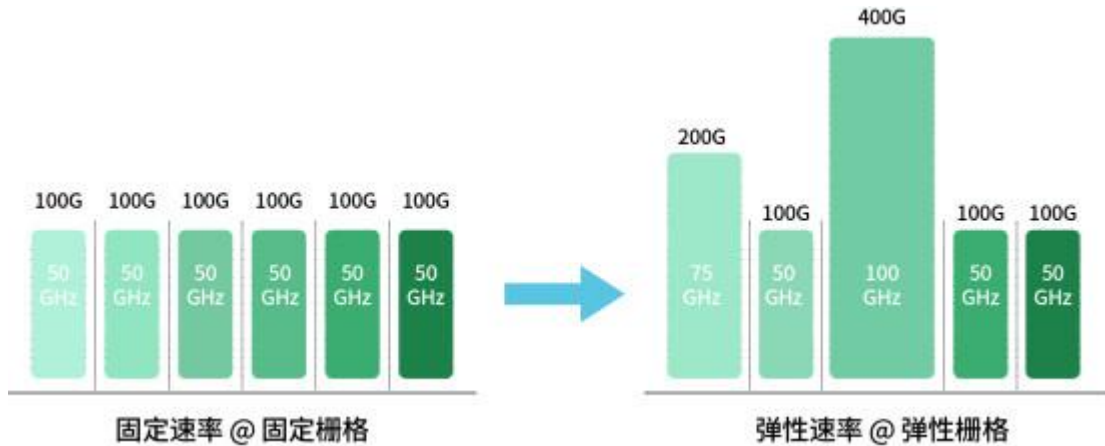


图 5-2 弹性速率和栅格，用于灵活的网络配置

网络架构的变化需要更灵活的线端光模块，以支持弹性速率和弹性栅格。

趋势：更高的光谱效率，接近香农极限

相干光模块在三个方向发展：

- 频谱效率：根据 oDSP 算法的进步，提高频谱效率和单纤容量。
- 波特率：提高单波长波特率以获得更高的单端口带宽，从而降低单位比特成本和功耗。
- 更小的尺寸和更低的功耗：使用集成的光电组件，先进的制造工艺和专用的 oDSP 算法。

由于香农极限，64Gbaud 400G 波长无法实现长距离光传输所需的性能。有必要使用更高的波特率和更复杂，更强大的 oDSP 算法，以满足城际（区域）和长距离骨干网的需求。例如，对于长距离链路（>1000km），400G 波长的波特率应高于 90Gbaud，并且 oDSP 中的 ADC 和 DAC 速率需要同时提高。

但是，随着波特率的增加，光纤的传输代价更高，甚至更难以补偿。因此，需要更强大的补偿算法来补偿物理通道损伤。考虑到 ROADM 已被广泛使用，端到端波长链路需要穿过数个甚至数十个 ROADM，其中包含波长选择开关（WSS）。WSS 滤波叠加效应使链路的有效带宽变窄，从而导致对 oDSP 中的补偿算法有更高要求。

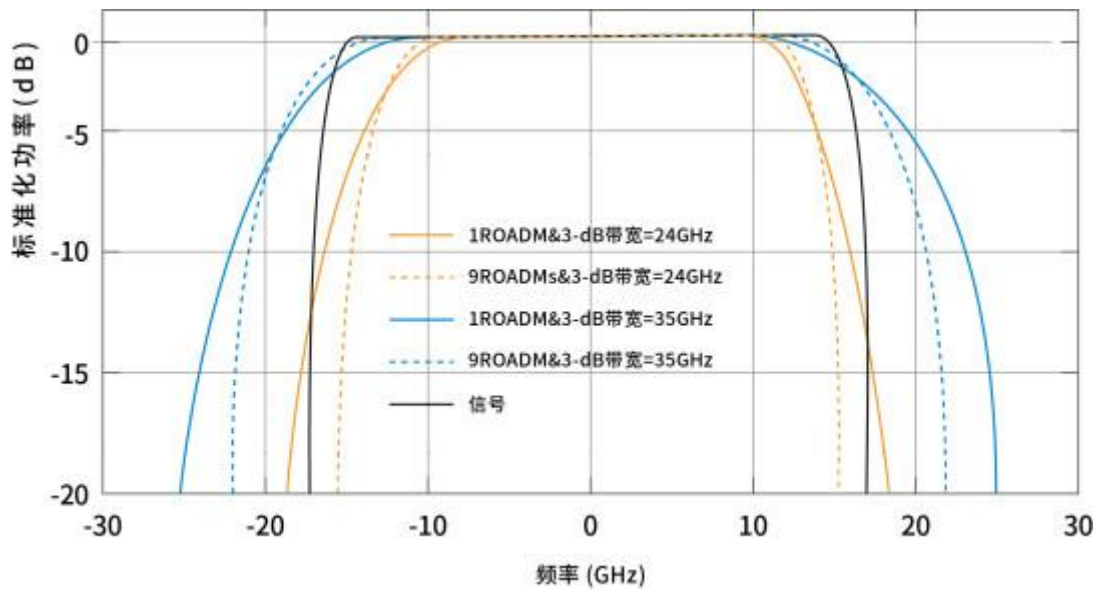


图 5-3 多层 ROADM 对光通道带宽的影响

此外，许多运营商希望基于端口速率和传输距离灵活配置调制格式和波特率。例如，为 400G 长距离传输部署 400G 16QAM，为数十 km 的城域数据中心互连部署 800G 64QAM，以提高频谱效率并降低单位比特成本。借助这种灵活的调制技术和光学层的弹性栅格，可以最大程度地提高光纤容量，从而节省了光纤投资。

结论

对更高容量，更低单位比特成本和更低功耗的需求正在推动光模块的传输速率越来越高。作为上一代的主流技术，100G 已进入其生命周期的成熟和稳定阶段，单位比特成本的降低受到极大限制。

当前，主流 400G 光模块已经用于各种网络场景中，例如数据中心网络，城域集成承载网络以及大容量和长距离传输网络。