

## 可插拔光模块：向更低功耗演进

在人工智能（AI）带来的带宽需求激增推动下，光网络正在经历重大变革。事实上，在数据中心内部，预计到 2030 年，AI 以太网网络将需要 335 Eb/s 的带宽，几乎是 2024 年的 60 倍。

AI 业务与其他流量一起，给网络带来了巨大压力 —— 不仅是带宽，还有功耗和空间。受限于供电、散热、机架与机柜空间，这些资源无法随数据需求线性扩展。2024 年，数据中心耗电量约占全球总用电量的 1.5% (415 TWh)，预计到 2030 年将翻倍以上，达到约 945 TWh。

虽然相干可插拔光模块已针对城域、区域和长距传输进行了优化，但数据中心内部连接（通常小于 500 米）正转向高效率可插拔光模块，以满足严格的功耗和散热限制。

### 降低可插拔光模块功耗

高速可插拔光模块中，影响功耗的一个关键因素是 SerDes 损耗。它取决于电信号从主机 ASIC 到光接口的传输距离。SerDes 通道长度与功耗直接相关：链路越长，数字信号处理器（DSP）为保证信号完整性所需的功耗和处理能力就越大。

因此，减少 SerDes 通道数量或长度，就能降低功耗。光传输的最大距离，是能效与用于校正信号畸变的信号处理能力之间权衡的结果。

以下是面向数据中心内部连接的几种架构：

#### 1. 全重定时光模块（FRO）

在 FRO 中，发送（Tx）和接收（Rx）通道均采用 PAM4 DSP 对信号进行重定时与恢复。

主机 ASIC 与光模块之间的电线路通常会带来高达 22 dB 的损耗，需要通过重定时来保证信号完整性和传输距离。这提升了性能，但也增加了功耗和时延。DSP 还会执行前向纠错，以实现更长距离的可靠传输。FRO 模块目前仍被广泛部署，尤其适用于 10 公里以上的链路。但其每比特功耗较高，不太适合数据中心短距链路（如 500 米以内）。

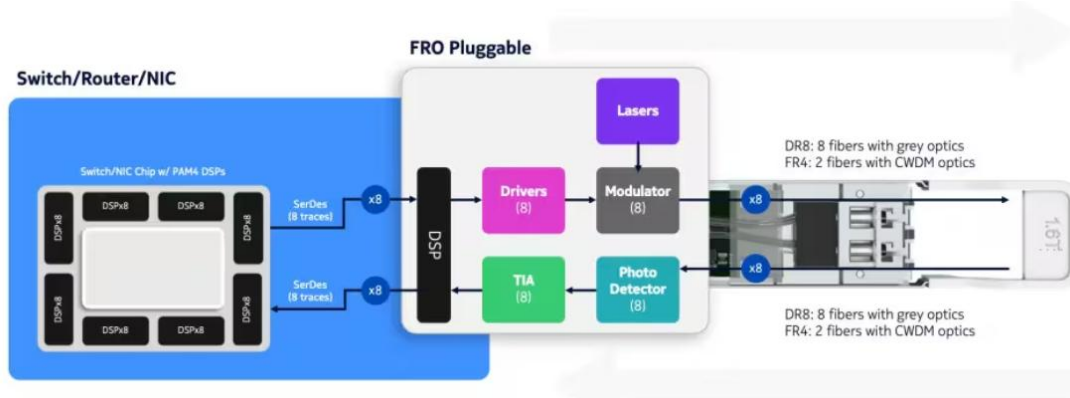


图 1: 全重定时光模块 (FRO) 的组成结构

## 2. 线性接收光模块 (LRO) / 发送重定时光模块 (TRO)

LRO 通过仅在发送端使用 DSP 重定时来降低功耗。由于 DSP 是光模块中功耗最高的部分，去掉接收端重定时可以显著降低模块整体功耗。

在 LRO 中，发送端仍使用 DSP 输出干净信号；接收端则工作在线性模式，将时钟恢复和均衡功能转移到主机 ASIC 处理。

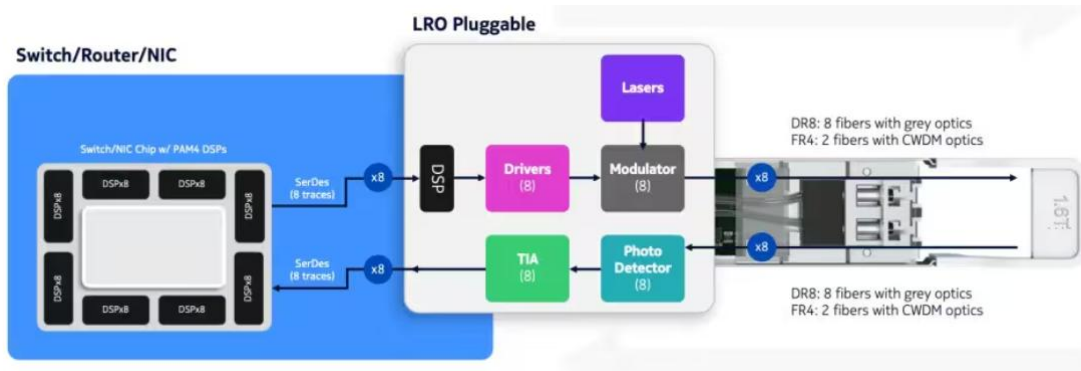


图 2: 线性接收光模块的组成单元

## 3. 线性可插拔光模块 (LPO)

在 LPO 中，发送和接收端均不包含重定时电路。光模块工作在纯线性模式，将未处理的原始信号直接送到 ASIC 的 SerDes 进行均衡和时钟恢复。去掉重定时功能后，模块功耗大幅降低。

由于模块内部没有 CDR/DFE 逻辑，信号校正由主机 ASIC 完成，这使得 LPO 结构精简、功耗高效，非常适合高密度、短距链路（通常多模光纤 100–500 米，单模光纤可达 2 公里）。注意：无 DSP 意味着需要主机 ASIC 具备先进的均衡能力和线性模拟前端。

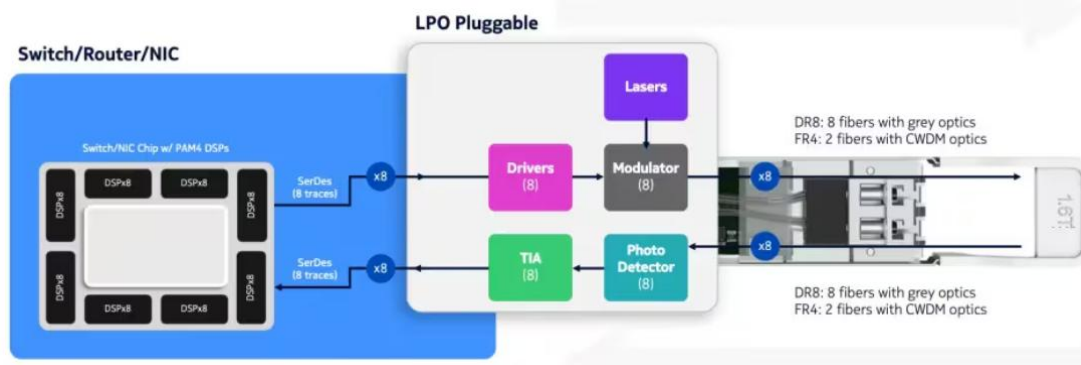


图 3：线性可插拔光模块（LPO）的组成单元

#### 4. 共封装光模块（CPO）

CPO 不再通过主板上的高速 SerDes 线路将电信号从主机 ASIC 传输到光模块，而是将光模块与主机 ASIC 集成在一起，通常在同一基板上。这大幅缩短了电路路径，将 SerDes 损耗降至 1–4 dB。

光信号转换被直接移到交换 / 计算 ASIC 旁边。与传统可插拔方案相比，CPO 显著降低了每比特功耗与时延，专为超短距、高密度链路设计。

虽然理论上 CPO 可达到 500 米，但其最佳应用场景是 scale-up 组网(GPU 直连 GPU)，能将类似铜缆的性能从 2 米扩展到 10–100 米。

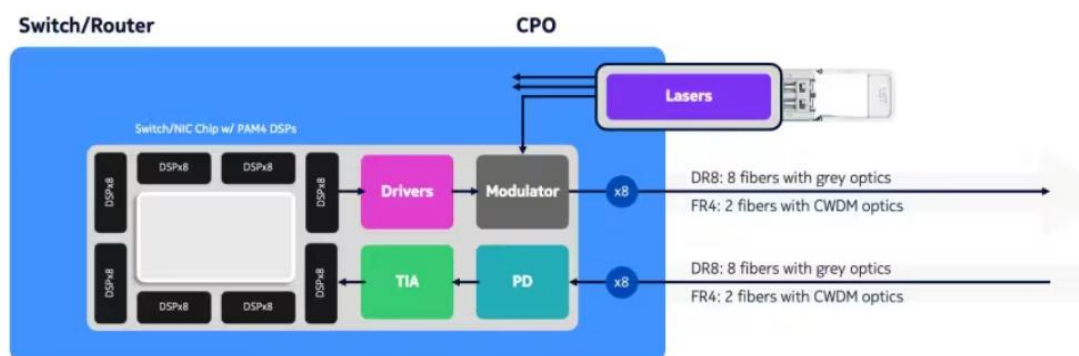


图 4：共封装光器件的组成单元

表 1 总结了这四种设计方案之间的差异。

对比项	FRO (全重定时光模块)	LRO (线性接收光模块)	LPO (线性可插拔光模块)	CPO (共封装光模块)
SerDes 损耗	20–22 dB	16–18 dB	16 dB	1–4 dB
DSP 使用	全量 (发送端 + 接收端)	半量 (仅发送端)	无, 由主机 ASIC 处理	集成至主机 ASIC
功耗	25 W / 模块	16 W/module	10 W/module	5 W*

对比项	FRO (全重定时光模块)	LRO (线性接收光模块)	LPO (线性可插拔光模块)	CPO (共封装光模块)
(约)	(0.016 W/G) (16.0 pJ/bit)	(0.01 W/G) (10.0 pJ/bit)	(0.0063 W/G) (6.3 pJ/bit)	(0.0031 W/G) (3.125 pJ/bit)
传输距离	30 m 至 >80 km	2–10 km	100 m–2 km	2–500 m
优势与挑战	<b>优势:</b> 行业标准;长距离信号完整性高;成熟的多厂商生态 <b>挑战:</b> 功耗与成本高(因使用DSP);时延高于线性方案	<b>优势:</b> 相比 FRO 功耗与成本更低 <b>挑战:</b> 需要主机 ASIC 具备接收端处理能力	<b>优势:</b> 极低功耗;低时延 <b>挑战:</b> 传输距离短;厂商互操作性有限;需要主机 ASIC 具备收发两端处理能力	<b>优势:</b> 超低功耗;低时延;最高密度 <b>挑战:</b> 需要高度复杂的主机 ASIC;仅激光源可现场替换;多厂商互操作性受限

\* 注: CPO 功耗数值通常存在差异,若使用外部激光源(ELS),可能不包含激光功耗。

### 诺基亚的数据中心内部光互联方案

凭借位于美国的光半导体工厂,诺基亚推出了 Infinite Capacity Engine – Data Center (ICE-D) 数据中心光互联技术,将多种光功能集成在单片芯片上,实现业界领先的高密度、低时延和高能效。

该技术支持 3.2Tb/s 以上速率,可将每比特功耗降低最高 75%,并可用于 FRO、LRO 和 LPO 架构,帮助数据中心运营商应对由 AI 应用驱动的内部容量指数级增长。

### 优化每比特功耗:数据中心内部光模块的发展方向

功耗效率是驱动可插拔光模块演进的核心因素。随着速率不断提升,传统架构(如 FRO)难以满足严苛的功耗和散热限制,尤其在高密度数据中心环境中。

LRO 和 LPO 通过将信号处理转移到主机 ASIC,实现了显著的功耗节省。CPO 则通过将光转换集成到主机 ASIC 内部,彻底改变了设计思路,但对主机 ASIC 提出了更复杂的要求。未来,多种创新架构将根据传输距离、成本、功耗需求进行优化,并在相当长时期内共存,以最大化能效并提升网络灵活性。