

市场看法

光电一体化封装（CPO）技术引领数据中心网络向全光演进

Kai Cui

执行概况

图 1

光电一体化封装（CPO）技术引领数据中心网络向全光演进

生成式AI的快速普及正在推动数据中心网络需求的指数级增长。光电一体化封装（CPO）技术以其高带宽密度、低功耗和可靠性优势，成为满足AI时代网络性能需求的关键方案。CPO通过光电融合显著提升网络带宽和能效，同时降低运营成本和通信时延。随着技术成熟和产业生态完善，CPO将在未来几年逐步从试点部署走向规模化应用，支撑AI工作负载的扩展和创新。

Key Takeaways

- AI驱动网络需求激增：生成式AI应用使数据中心网络面临带宽、功耗和可靠性瓶颈。
- 传统网络架构失效：现有100G/200G网络无法满足AI集群东西向流量的爆发式增长。
- CPO技术优势显著：CPO通过光电融合提升带宽密度、优化能效，并降低时延与故障率。
- 市场增长潜力巨大：IDC预测中国生成式AI相关网络硬件支出2023—2028年复合增长率达38.5%。
- 部署窗口临近：2025—2026年是CPO试点部署的关键时期，超大规模数据中心将率先验证其价值。

Recommended Actions

- 提前规划技术路线：在2025—2026年试点部署CPO技术，抢占AI时代的战略先机。
- 聚焦高价值场景：优先在AI训练集群和超大规模云数据中心中应用CPO技术。
- 推动产业协同：与供应链合作加速CPO技术标准化和生态体系建设。
- 全面评估投资回报：结合长期TCO优化，制定全生命周期的成本效益分析。
- 加强技术验证：通过试点部署验证CPO性能和可靠性，为未来规模化应用奠定基础。

来源：IDC, 2025

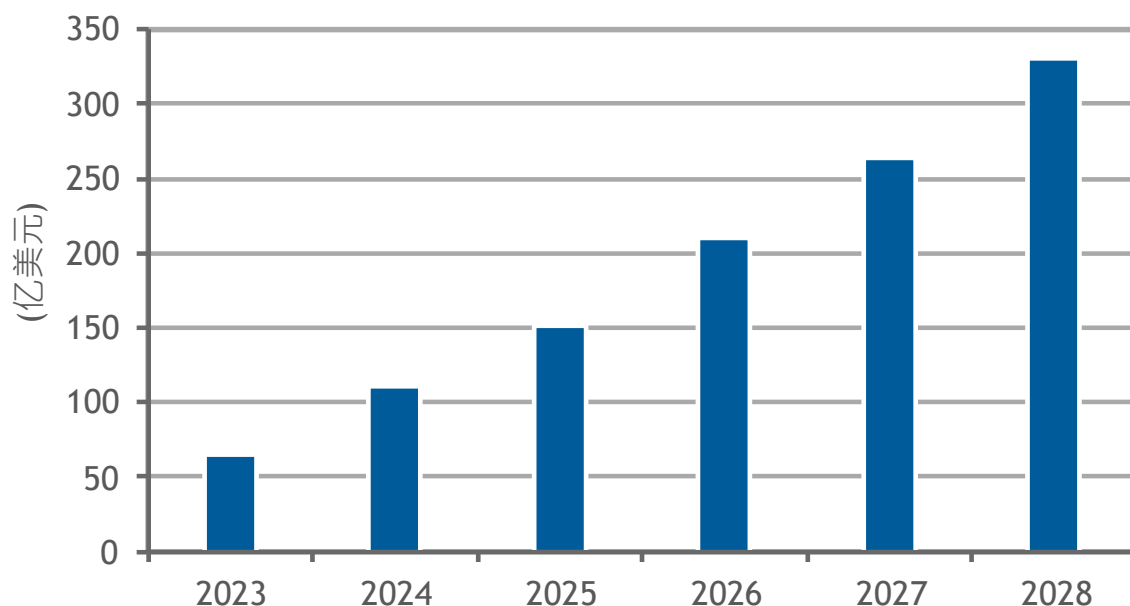
AI 应用普及加速驱动高性能网络需求

人工智能应用在工作和生活场景中迅速普及，从企业业务决策到个人数字助理，各行业纷纷引入 AI 以提升效率和创新。如：媒体公司采用生成式 AI 技术制作视频；教育公司也基于生成式 AI 构建解题思维链，深化个性化自主学习；汽车行业不断优化自动驾驶系统；金融机构利用生成式 AI，提供更加精准的个性化金融服务，提升销售等；IDC 统计，2024 年下半年中国大模型商用落地日均 Tokens 消耗量增长近 10 倍，2024 年中国公有云大模型调用量累计 114.2 万亿 Tokens（不包含使用海外 MaaS 平台的调用量）。

生成式 AI 的快速发展不仅显著提升了企业应用的智能化体验与整体运营效率，同时也对底层数字基础设施提出了全新且更高的要求。随着数据中心内 Token 用量持续攀升激增，东西向流量大幅增长、通信能耗不断加大，网络正成为为 AI 基础设施的决定性瓶颈。400G-800G-1.6T 的网络演进节奏显著加速，传统“先算力、后网络”的升级逻辑已经失效。AI 服务提供商唯有把网络规划前置，通过高带宽、高可靠、低能耗的全栈创新，才能真正释放 GPU 算力红利，并在 AI 平台竞争中占据领先地位。2024 年，中国高端以太网（≥200G）端口出货量突破 600 万，未来将保持 45.6% 的复合增长率，在 2029 年中国高端以太网端口出货量将超过 4300 万个。IDC 预测，中国生成式 AI 相关网络硬件支出将持续加速，从 2023 年的 65 亿美元，增加到 2028 年的 330 亿美元，复合增长率达到 38.5%。

图 2

中国生成式 AI 相关网络硬件支出预测，2023—2028



来源: IDC, 2025

当前数据中心网络正在成为 AI 效能提升的瓶颈

虽然 AI 算力飞速提升，但现有数据中心网络在支持 AI 应用时暴露出多方面瓶颈，主要痛点包括：

网络吞吐能力不足：大规模 AI 集群往往由数千上万颗 GPU/AI 加速卡组成，节点间需要交换庞大的训练数据和模型参数。如果网络带宽不足，GPU 将因等待数据而空转，整体算力难以有效发挥。然而许多现有智算中心仍以 100G/200G 连接为主，无法充分满足 AI 集群东西向流量爆发式增长的需求，网络已成为制约 AI 训练和推理规模扩展的瓶颈。

功耗成本居高不下：高速网络带来的功耗激增正成为数据中心运营的沉重负担。电力是数据中心最大的经常性开支，占企业数据中心支出的 46.3%。IDC 预测，AI 数据中心的能耗预计将从 2022 年的 23.0TWh 增长到 2027 年的 146.2TWh，年均增长率为 44.7%。当数据传输速率攀升至 800G 和 1.6T 时代，传统主动式电接口和可插拔光模块的功耗变得更大，这进一步推高了功耗和散热负荷。高功耗不仅意味着更高的电力和冷却成本，还会受到数据中心供电与散热能力的限制，成为扩展 AI 基础设施的一大障碍。

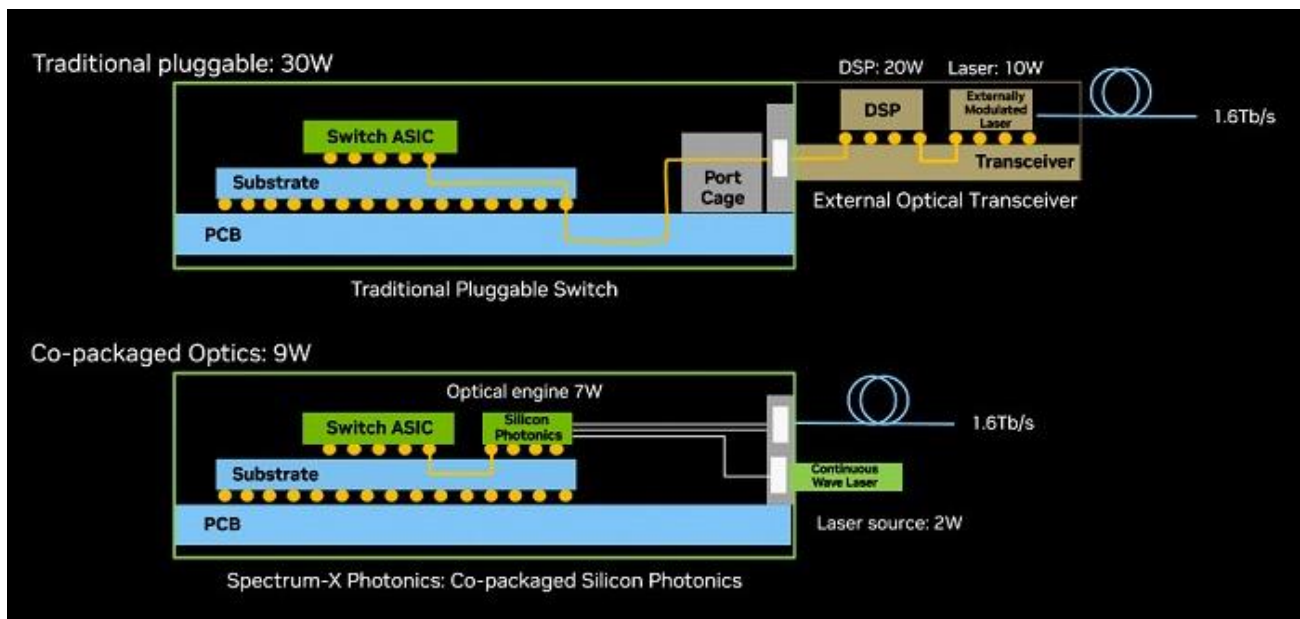
网络可靠性与扩展性不足：在传统架构下，大型 AI 集群往往需要成千上万个光模块和连接器，这些组件增加了故障点并降低了网络可靠性。光模块和连接器是网络中常见的故障源，即使高质量器件硬故障率仍在每十亿小时百次量级，灰尘等因素导致的软性故障更为频繁。当连接数量以 AI 规模成倍增长时，任何单点失效都可能导致训练任务中断或性能下降。此外，传统多级交换架构下数据需要经过多跳转发，路径的复杂性也增加了通信延迟，难以提供 AI 工作负载所需的稳定低时延网络。当前种种可靠性挑战表明，现有网络架构在面向 AI 时已接近极限，业界亟需新的方案来简化互连结构、提升稳定性。

CPO 技术加速商用，并引领数据中心网络向全光演进

面对上述痛点，光电一体化封装（Co-Packaged Optics, CPO）技术和产业正日趋成熟，成为未来 5 年数据中心大规模横向扩展（scale-out）网络架构的关键技术。CPO 通过将光引擎与交换 ASIC 芯片集成封装在同一基板上，实现“光电融合”，从物理层根本上缓解了传统架构的瓶颈。这一创新设计带来了多方面显著优势：

图 3

CPO 连接与传统光模块连接硬件技术方案区别



来源: 网络公开资料, 2025

更高带宽密度：由于省去了前面板可插拔接口，CPO 大幅缩短了交换芯片到光引擎的连接距离，使得在有限封装空间内可以支持极高的 I/O 带宽密度，在紧凑封装中提供远超以往的吞吐能力。这意味着单台交换机可承载更多光通道，支持更大规模集群的互连需求。

功耗和能效优化：CPO 通过将光学收发器移至 ASIC 侧，省去了长距离高速电连接和中间器件，极大提高了能效。Cisco 的统计显示，与传统 DSP 驱动的可插拔光模块方案相比，在 51.2T 交换机中引入 CPO 技术，可使整机功耗降低 25%~30%。有效解决高速、高密度互连的能耗难题。功耗的大幅下降不仅缓解了数据中心的电力和冷却压力，也意味着单位带宽的运营成本显著降低。

可靠性与低时延：由于取消了可插拔光模块和从交换芯片到光引擎之间的大量高速电连接，CPO 从源头上减少了插拔连接器和焊接点等故障点，潜在提升系统可靠性。光引擎贴近芯片还缩短了信号传输路径，减少了信号衰减和转接环节，从而降低了传输时延并提高信号完整性。虽然 CPO 整体可靠性仍需经过大规模部署验证，但没有松动接插件的固定式设计天生避免了灰尘污染等问题，有望提供更稳定的链路通信。根据英伟达 (NVIDIA) 在 GTC 大会上发布其 CPO 交换机时的官方数据，相比目前的可插拔交换机，NVIDIA 的 CPO 交换机可以在可靠性上提升 10 倍和时延上降低 3 倍。

成本与集成度优势：技术不断成熟，因为高度集成减少了器件数量、简化了系统设计，CPO 在规模效应下将带来整体成本的下降。同时，共封装方案体积更小，有助于缩减设备占用空间并简化布线，这在空间宝贵的高密度数据中心内尤为重要。尽管当前 CPO 硬件初始成本可能仍高于成熟的可插拔器件，但长期来看，通过功耗节省和结构优化，其总拥有成本 (TCO) 具有竞争力。

CPO 技术的快速发展得到了产业界的印证。领先厂商正引领 CPO 产业走向实用化：例如：英伟达（NVIDIA）在 2025 年 GTC 大会上发布了 NVIDIA Quantum-X 和 NVIDIA Spectrum-X 硅光网络交换机（涵盖 InfiniBand 和以太网协议），其中将在 2025 年下半年量产的 NVIDIA Quantum-X 硅光网络交换机可提供 115.2Tb/s 吞吐量和 144 个 800Gb/s 无阻塞多芯连接接口（MPO: Multi-fiber Push On），之后推出的 100T 和 400T 的 Spectrum 以太网 CPO 交换机，在大幅提升 AI 网络的互连带宽和密度同时，为构建百万级 GPU 的高性能 AI 工厂提供了理想选择，英伟达有望成为全球首家在数据中心大规模部署 CPO 网络的厂商；台积电方面推出了"COUPE"先进封装平台，实现硅光器件与半导体工艺的一体集成，为 CPO 大规模量产奠定工艺基础；其他国外厂商以及国内厂商也在积极布局 CPO 领域。产业的百花齐放使 CPO 生态体系日趋丰富完善。

CPO 技术引领数据中心网络向全光演进。可以预见，随着 400G/800G 乃至更高速率以太网时代的到来，CPO 将逐渐从试验阶段走向数据中心网络的核心，引领数据中心网络向全光演进。尤其在大规模横向扩展的 AI 数据中心的 CPO 有望成为主流架构，支撑起指数级增长的带宽需求和严格的能效目标。同时，CPO 的实践还为更远期的"光学 I/O"奠定基础，即未来在算力芯片之间直接采用光互连，将计算与通信深度融合。CPO 技术以其独特优势正在崛起，为化解 AI 时代的网络性能瓶颈提供了一条切实可行的路径。

CPO 技术的挑战与演进路径与部署建议

尽管 CPO 的发展前景被广泛看好，但在大规模商用之前，仍需逐步应对封装工艺、电-热管理、标准互通等多项挑战，并稳步推进技术成熟与生态建设。

封装工艺复杂度：将高速光器件与 ASIC 芯片共封装对制造工艺提出了极高要求。光学元件与电芯片在同一封装中对准连接需要精密的 3D 封装和耦合技术，工艺窗口小、良率提升困难。初期产能和良率的限制可能推高成本、影响供应。

热设计与散热：将光模块移入交换芯片封装后，热功耗在局部高度集中，对散热提出严峻挑战。激光器、调制器等光器件对温度敏感，而高速芯片本身功耗密度就很高，英伟达（NVIDIA）在 CPO 数据中心交换机上率先采用了业界领先的微环调制器（MRM: Micro-Ring Modulator），极大地降低了光路上的功耗，有助于构建高密度端口的交换机。传统风冷方式可能无法满足，需要引入冷板、水冷等更高效的散热技术。

维护和可靠性：CPO 取消了可插拔模块，在维护上可以更加简洁，但是也带来了新挑战。以往光模块故障时可热插拔替换，而在 CPO 方案中，光引擎与交换芯片绑定，单个光通道故障可能需要更换整个交换模块，无法简单更换模块恢复。NVIDIA 探索了可插拔光子分离模块等部分可维护设计，但整体来看 CPO 设备部署初期的现场维护难度较高，需要通过提升整体 CPO 交换模块的可靠性来解决这个问题。CPO 的优势也在于由于光电一体化封装会在无尘的环境中进行，这会极大提高整个模块的可靠性，避免人工插拔光模块带来的故障，同时也可加快 AI 工厂的部署速度。

产业生态与标准：目前 CPO 产业链各环节尚不完全成熟，标准化工作也在推进中。共封装涉及光芯片、激光器、封装基板、主设备等多个领域，未来发展亟需产业链各环节协同。OIF 等组织也在讨论 CPO 接口标准，以便不同厂商方案的互通。随着龙头企业验证 CPO 可行，整个产业链（光器件供应商、代工厂、系统厂商等）正投入资源完善配套，标准和规模效应有望在未来几年逐步显现。

给数据中心（云）服务商的建议

部署时机判断：综合当前技术进展和产业节奏，2025—2026 年将是考虑 CPO 试点部署的关键窗口。IDC 认为，大规模 CPO 商用可能在未来 2—3 年后（约 2027—2028 年）出现，但部分头部玩家将于 2026 年开始初步部署 CPO 以验证技术和运营模式。这意味着有前瞻性的超大规模数据中心应当在 2025 年前后着手调研评估，在 2026 年前后进行小规模试点。提前介入不仅可以在技术成熟时顺利扩展，也有助于与厂商合作塑造技术演进方向，获取先发优势。反之，如果等到标准完全成熟再行动，可能错失利用新技术降低成本和提升性能的战略窗口。

适用的网络规模：从 ROI 角度看，CPO 当前最适合规模足够大的网络环境。AI Factory 级别的超大数据中心（如年投入上万台服务器、数十万个光模块）将是 CPO 优先发挥价值的场景。在庞大的网络下，每节省 1W 功耗、每提升 1% 的可靠性都意义重大。对于这类规模的用户，CPO 的能效和密度优势将带来显著的总体价值。在评估网络规模时，可将 GPU 集群规模、东西向流量占比、光模块用量等作为决策参考指标：当这些指标达到临界点时，引入 CPO 的收益将逐步大于成本。

部署场景选择：优先将 CPO 用于 AI 训练集群、超大规模云数据中心等带宽密集型场景。典型场景包括：大规模 AI 模型训练所用的 GPU 集群内部网络（需要极高的东西向吞吐和无阻塞交换）；公有云提供 AI 即服务时底层的算力池互连；以及高性能计算（HPC）中心内部的超低延迟互连网络等。在这些场景中，网络性能直接决定了任务效率和服务质量，CPO 提供的更高带宽和低延迟可以明显提升整体算力利用率和 AI 作业完成速度。此外，在建设“智算中心”的新型数据基础设施项目中，也可将 CPO 列为架构选项之一，以满足日益增长的 AI 算力调度需求。

ROI 和 TCO 考量：引入新技术需仔细评估投资回报。对于 CPO，其价值主要体现在长期 TCO 优化而非短期 Capex 节省。一方面，功耗降低带来的电力和冷却成本节约是最直接的回报。以 51.2T 交换机为例，CPO 方案可节省约 25%—30% 的整机功耗；在一个部署数百台交换机的大型集群中，这意味着节省数十千瓦的用电，按中国当前电价和 PUE 估算，每年可减少相当可观的运营支出。另一方面，带宽密度提升和端口成本下降将减少基础设施扩容所需的设备数量。CPO 可使单位带宽的光学成本降低最多 40%，加之其架构简化了层级，长期可降低每 Gbps 网络容量的总体建设成本。这些效益需要在 3—5 年的使用周期中体现，建议用户在 ROI 分析时将能源成本、空间成本、运维成本纳入全生命周期模型进行衡量。综合来看，对于引领技术潮流的先行者，CPO 有望通过能效提升和性能优势实现正向的 TCO 回报。更重要的是，它可以为数据中心带来不可量化的战略收益——支持更大规模的 AI 工作负载和更快速的模型迭代，从而抢占行业制高点。

总而言之，AI 时代的数据中心网络正面临前所未有的性能和可靠性挑战，光电一体化封装（CPO）技术以其高带宽密度、低功耗和高度集成的优势，有望成为破解网络瓶颈的关键方案。尽管 CPO 当前尚处于产业化初期，存在工艺、散热、维护等方面的挑战，但这些问题正在随着技术进步和生态协同而逐步解决。我们预计在未来两年内，会有领先的超大规模数据中心率先部署 CPO 进行验证，并在随后的 2025—2026 年形成一批标杆应用案例。作为面向未来的决策者，中国的 CIO、CTO 们应密切关注这一领域的发展动态，结合自身业务需求制定相应的技术路线图。在审慎评估的基础上适时投入试点，不仅能在技术成熟时迅速扩展，享受降本增效之利，更能向业务部门交付支撑 AI 创新的强大网络能力。CPO 将成为支撑 AI 普及和数字化转型的新一代网络基石，引领数据中心进入光电融合的新时代。

相关研究

- *Generative AI's Impact on Data Generation and Storage* (IDC #US53317225, 2025 年 4 月)
- *Worldwide Semiconductor Advanced Packaging Market Forecast and Analysis, 2025–2029* (IDC #US52212025, 2025 年 2 月)
- *Worldwide Generative AI Datacenter Ethernet Switching Forecast, 2024–2028: Asserting Our Bullish View* (IDC #US52399523, 2024 年 7 月)

大纲

本 IDC 研究重点关注人工智能 (AI) 应用的快速普及正在推动高性能网络需求的加速增长。生成式 AI 的广泛应用显著提升了企业智能化体验和运营效率，但也对数据中心网络提出了更高要求。现有网络架构面临吞吐能力不足、功耗成本高昂、可靠性与扩展性不足等瓶颈，制约了 AI 训练和推理的规模扩展。光电一体化封装 (CPO) 技术作为一种创新解决方案，通过将光引擎与交换 ASIC 芯片集成封装，提供更高的带宽密度、优化功耗与能效、提升可靠性与低时延，并降低长期总拥有成本 (TCO)。尽管 CPO 技术仍面临封装工艺复杂、散热挑战、维护难度及产业生态不成熟等问题，但其独特优势使其成为未来数据中心网络的关键技术。IDC 预测，2025—2026 年将是 CPO 试点部署的关键窗口，超大规模数据中心将率先验证其价值。CPO 技术不仅能满足 AI 时代指数级增长的带宽需求，还将引领数据中心向全光网络演进，成为支撑 AI 普及和数字化转型的核心基石。

IDC 中国电信行业与网络市场助理研究总监崔凯表示："AI 的指数级增长正在重塑数据中心网络的需求格局。光电一体化封装 (CPO) 技术以其高带宽密度、低功耗和高度集成的优势，为解决网络性能瓶颈提供了切实可行的路径。尽管面临技术和生态挑战，CPO 的潜力不可忽视。超大规模数据中心应抓住 2025—2026 年的试点窗口，提前布局这一技术，以抢占 AI 时代的战略制高点。CPO 不仅是技术创新的突破，更是推动 AI 普及和数字化转型的关键驱动力。"

关于 IDC

国际数据公司 (IDC) 是在信息技术、电信行业和消费科技领域，全球领先的专业的市场调查、咨询服务及会展活动提供商。IDC 在全球拥有超过 1,100 名分析师，他们针对 110 多个国家/地区的技术、IT 对标研究和采购以及行业发展机遇和趋势，提供全球化、区域性和本地化的专业意见。IDC 的分析和洞察有助于 IT 专业人士、业务主管和投资界做出基于事实的技术决策，实现他们的关键业务目标。IDC 成立于 1964 年，是国际数据集团 (IDG, Inc.) 的全资子公司。

IDC China

IDC 中国 (北京)：中国北京市东城区北三环东路 36 号环球贸易中心 E 座 901 室

邮编：100013

+86.10.5889.1666

Twitter: @IDC

blogs.idc.com

www.idc.com

版权声明

这份 IDC 研究文件作为 IDC 持续情报服务的一部分发布，提供书面研究、分析师互动以及网络会议和会议活动记录。请访问 www.idc.com 了解有关 IDC 订阅和咨询服务的更多信息。要查看 IDC 全球办事处列表，请访问 www.idc.com/about/worldwideoffices。请致电 +1.508.988.7988 或访问 www.idc.com/?modal=contact_repsales 联系 IDC 报告销售人员，了解有关应用本文档价格购买 IDC 服务的信息，或有关其他副本或网络权利的信息。2025 IDC。未经许可，不得复制。保留所有权利。