

云智算光互连发展报告



牵头编写单位

中国移动云能力中心、中国科学院计算所、中国移动研究院

联合编写单位

锐捷网络、华为、中兴通讯、盛科通信、新华三、
曦智科技、奇点光子、图灵量子

前言

本发展报告面向未来智算中心超大规模扩展、AI 大模型极致性能与高效部署的核心需求，联合产业合作伙伴共同提出先进光互连技术架构与演进路径，旨在突破传统电互连在带宽、距离与能效方面的根本性瓶颈，构建高带宽、超低时延、低功耗及高可靠性的新一代智算中心互连底座，为人工智能、高性能计算及云服务等相关业务的持续跃升提供坚实支撑。

本发展报告的版权归中国移动云能力中心所有，并受法律保护。转载、摘编或利用其它方式使用本发展报告文字或者观点的，应注明来源。

目录

前言	1
目录	2
1. 背景与需求	4
2. 智算中心光互连技术概述	5
2.1 新型可插拔模块	5
2.1.1 线性可插拔光学	5
2.1.2 线性接收光学	6
2.2 光电共封技术	6
2.2.1 板上光学	6
2.2.2 近封装光学	7
2.2.3 共封装光学	8
2.2.4 光输入/输出	9
2.3 光交换	9
2.3.1 光线路交换	9
2.3.2 光分组交换	11
2.3.3 光突发交换	12
3. 智算场景下光互连技术的应用研究	13
3.1 LPO 在 AIGC 算力网络中的应用	13
3.2 CPO 交换机在智算场景下的应用	14
3.3 OCS 在 AI 集群参数面的应用	15
3.4 光互连技术在 GPU 超节点的应用	16

4. 移动云在智算场景下的光互连应用展望	18
5. 产业生态与标准化	21
5.1 光电领域互连标准	21
5.1.1 CPO 领域标准	21
5.1.2 LPO 领域标准	22
5.1.3 Chiplet 领域标准	22
5.2 光电领域交换标准与产业生态	24
5.2.1 光交换标准发展现状	24
5.2.2 光交换产业生态进展	27
6. 发展趋势与发展建议	28
6.1 发展趋势	28
6.2 产业发展建议	28
附录:	30
常见缩略语	30

1. 背景与需求

在 AI 大模型、云计算及智能应用普及的推动下，全球算力需求正经历前所未有的爆发式增长。基于铜缆的互连技术在带宽密度、传输距离与能耗效率上的瓶颈日益凸显，光子作为光互连技术的信息载体和物理基石，具有极低传输损耗、超高频率、抗干扰等物理特性，使得光互连技术在带宽、距离、抗扰、功耗、密度等方面具有压倒性优势，拥有巨大潜力。

光互连技术的应用范围正从传统的电信骨干网和城域网，快速向数据中心内部、高性能计算集群等更广泛的领域渗透。特别是在数据中心内部，随着服务器端口速率向 400G、800G 乃至 1.6T 演进，光互连技术方案正迅速取代铜缆，成为数据中心以及超节点场景下的优选方案。随着 LPO、CPO 等技术引入数据中心架构，光电协同设计已成为芯片集成的核心技术需求，芯片-封装-系统级的多维协同优化成为新的挑战。与此同时，随着全光交换技术的逐步小规模应用，为光互连技术的演进方向提供了新的思路。

本发展报告聚焦光互连技术在智算中心和数据中心等典型应用场景下的技术演进，为行业提供兼具前沿性与实践性的技术参考。

2. 智算中心光互连技术概述

随着智算中心的飞速发展，数据吞吐量激增，对底层硬件互连提出了前所未有的挑战。在此背景下，光互连技术以高带宽、低时延、低功耗等方面的优势，有望成为未来算力时代不可或缺的基础设施。智算中心场景下的光互连技术具体包括新型可插拔模块、光电共封以及光交换三个核心技术方向。

2.1 新型可插拔模块

2.1.1 线性可插拔光学

随着数据中心传输速率的不断攀升，传统光模块的功耗和成本急剧上升，已成为制约数据中心扩展的瓶颈。

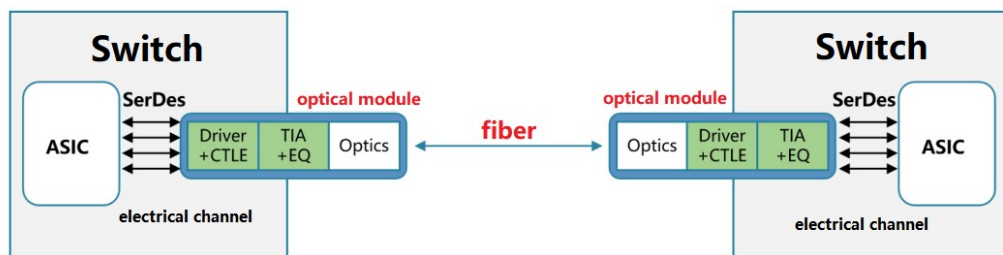


图 1 线性可插拔光学结构

在传统光模块的功耗中，DSP 模块占了很大的比例，因此在 LPO 技术中，直接去除了传统光模块中的 DSP，在发射端使用具有高线性度的 Driver，在接收端使用高线性度的 TIA，从而构建一个纯模拟的、“线性直驱”的光信号处理通道，如图 1 所示。虽然去除了传统光模块中的 DSP，但是 DSP 功能并未消失，而是转移到了交换机 ASIC 中，这意味着 ASIC 的 SerDes 模块必须具备更强的线性驱动能力和信号处理能力。

由于移除了传统光模块中的 DSP 模块，LPO 技术能够将功耗降

低 30%~50%，并能够降低延迟。于此同时，由于去除了 DSP 模块，能够在一定程度上节省成本，并且 LPO 技术保留了可插拔的产品形态，有比较好的可维护性。

2.1.2 线性接收光学

LRO 在接收端移除了 DSP，发送端保留了重定时器，是相较于 LPO 的折中方案，如图 2 所示。重定时器可以对信号进行整形、重新计时和放大，发送端保留重定时器能够补偿了信道损耗，确保了发送信号的质量，并提供了更好的互操作性和链路诊断能力。而接收端采用线性模拟组件直接接收来自主机 ASIC 的信号，这在一定程度上减少了的光模块的功耗，以实现节能降本的功效。

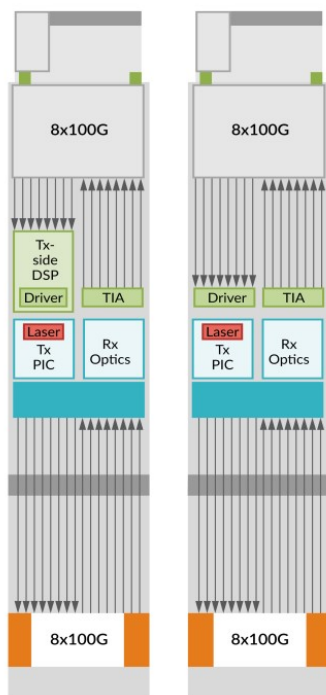


图 2 线性接收光学结构示意图，左侧为 LR0，右侧为 LPO

2.2 光电共封技术

2.2.1 板上光学

OB0 的核心思想是：将光引擎从传统的可插拔模块中解放出来，

直接安装到系统主板上，但交换或计算的电芯片仍然保持独立的封装，如图 3 所示。光引擎与电芯片通过主板上的精密走线进行互连。

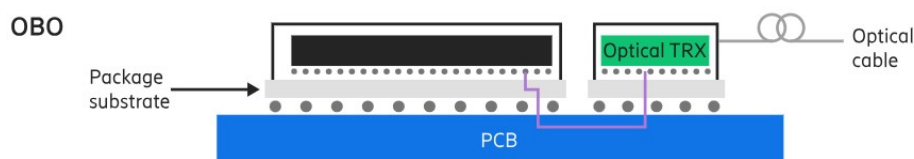


图 3 板上光学结构

由于移除了可插拔模块的“金手指”接口、外壳以及部分重复的电路，缩短了电芯片与光引擎的电气路径，从而降低了信号驱动的功耗。同时，如果光引擎损坏，可以单独进行更换，而无需更换昂贵的电芯片或整个主板。

在性能方面，OBO 虽然优于可插拔模块，但由于主板上的互连距离仍然较长，因此在超高速率传输场景下的优势并不明显。

2.2.2 近封装光学

NPO 的核心思想是：将光引擎非常靠近电芯片放置，但并不像 CP0 那样与电芯片共封装在同一基板或中介层上。它通常将光引擎安装在同一基板上，通过极短的高性能电气链路与电芯片相连，形成一个高度集成的系统，如图 4 所示。

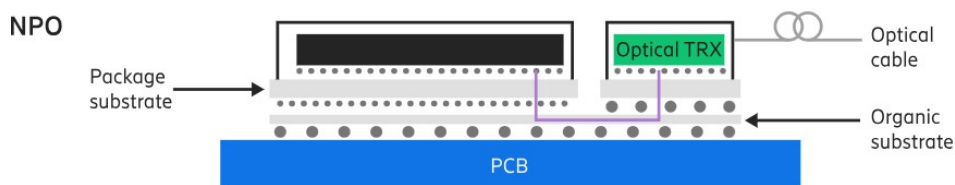


图 4 近封装光学结构

NPO 将光引擎与电芯片物理分离，避免了电芯片的高温热量直接冲击光器件，散热设计更简单、高效。由于电芯片本身是巨大的热源，工作时温度很高，而激光器等光器件对温度极其敏感，所以，

光引擎与电芯片共封装会导致波长漂移和性能下降。同时，由于光引擎未与电芯片共封装，NPO 在可维护性层面具有优势，如果光引擎失效，只需更换光引擎子模块即可，避免了大量的维护成本。

相比激进的 CPO 技术，NPO 技术是一种更务实、风险更低的路径。并且，NPO 与传统光模块相比，其性能远超传统光模块，其主要优势包括以下几点：

- NPO 的光引擎拥有更大的可布置面积和更灵活的走线方案，可以方便地使用 LGA 封装，且有利于光引擎散热；
- NPO 不影响电芯片原有设计，只对 PCB 或基板做差异设计即可满足不同需求；
- NPO 与电芯片解耦，能够避免形成电芯片垄断问题；
- NPO 可单独测试 TP1 的电信号质量，可归一化设备的驱动与固件，可测试性更好。

2.2.3 共封装光学

CPO 相对于 NPO，光引擎与电芯片共封装在同一个插槽或基板上，集成度更高，电互连距离更短，如图 5 所示。同时，相对于传统光模块，CPO 能够显著降低功耗，并通过节省设备面板空间，可克服面板 I/O 密度限制。

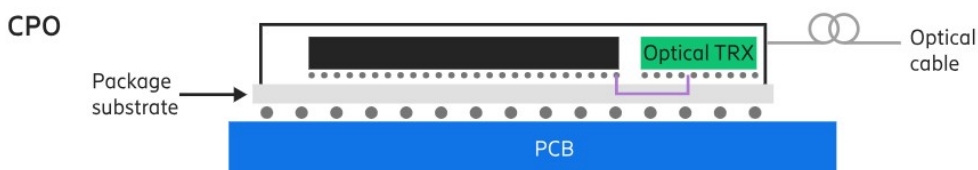


图 5 共封装光学结构

由于光引擎和电芯片紧密共封装，任何一部分的故障都可能导致整个封装体的更换，因此对光引擎的良率、可靠性以及可维护性方面提出了极高要求。

CPO 目前仍处于发展初期，但其在超高带宽、低功耗、高密度互连方面的巨大潜力使其成为未来光通信，特别是 AI 算力集群和超大规模数据中心不可或缺的技术方向之一。

2.2.4 光输入/输出

OIO 的核心理念非常具有颠覆性，它彻底摒弃传统的铜线电气 I/O，将光互连直接集成到计算芯片的封装内部或紧邻位置，使芯片能够直接通过光信号进行数据处理，如图 6 所示。

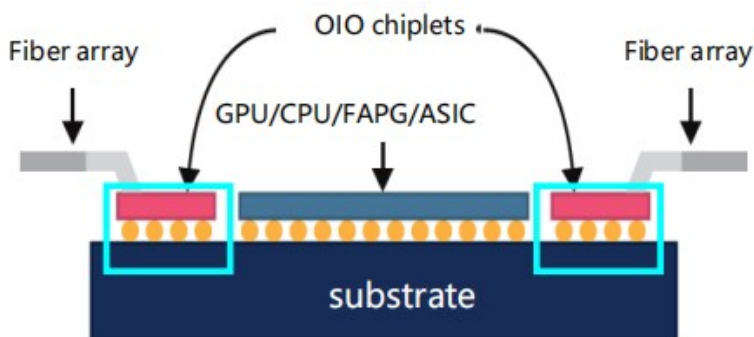


图 6 光输入/输出结构

OIO 的优势主要在于消除了板级电气走线瓶颈，能够大幅提升传输带宽，并将延迟降低至纳秒级，能够更好的契合 AI 模型训练的需求。另外，由于消除了电气走线带来的巨大能量损耗，OIO 将带来颠覆性的能效提升。

2.3 光交换

2.3.1 光线路交换

OCS 的本质是通过光学器件直接操控光信号的传输路径，实现输入端口到输出端口的连接。与需要光电转换的传统电交换不同，OCS 全程在光域操作，因此具有协议透明性、超低延迟、高能效等优势。但 OCS 也存在光学器件累计损耗、稳定与可靠性方面的劣势，

另外，OCS 是一种基于端口的光交换技术，导致其扩展能力受限。

目前 OCS 领域的主流技术包括：直接光束偏转（DLBS）、数字液晶（DLC）和微机电系统（MEMS）。其中，基于 MEMS 的光交换技术具有大端口、低插损、快速切换和低成本的优势，是目前各厂商选择的主流技术路线。

MEMS 光交换是一种基于微机电系统的光交叉连接技术。它通过利用微纳技术制造微小尺寸的光学组件和机械结构，实现光信号的灵活路由和交叉连接，如图 7 所示。MEMS 光交换是微镜反射型，方便集成和控制，易于组成光交换阵列，是 MEMS 光交换研究的重点。微镜阵列芯片是 MEMS 光交换系统中的核心组件，负责在光纤间建立和断开连接。

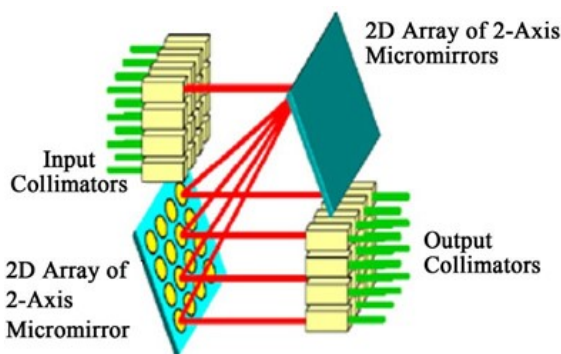


图 7 MEMS 微镜原理

OCS 与流量调度系统结合，能够显著提升资源利用效率。目前国外主要由谷歌主导，谷歌在其自研 TPU 集群中已批量应用 OCS，自 TPUv4 沿用到今年发布的 TPUv7。每 64 个 TPU 通过电互连组成一个 cube，cube 间使用 OCS 进行互连，如图 8 所示。在该组网下，通过 OCS 灵活调整拓扑的能力，隔离故障节点提升可用性，集群可减少 50 倍停机时间，并通过按需定制拓扑，提升 30% 的吞吐量。除此之外，OCS 在网络成本和网络功耗都有明显收益。

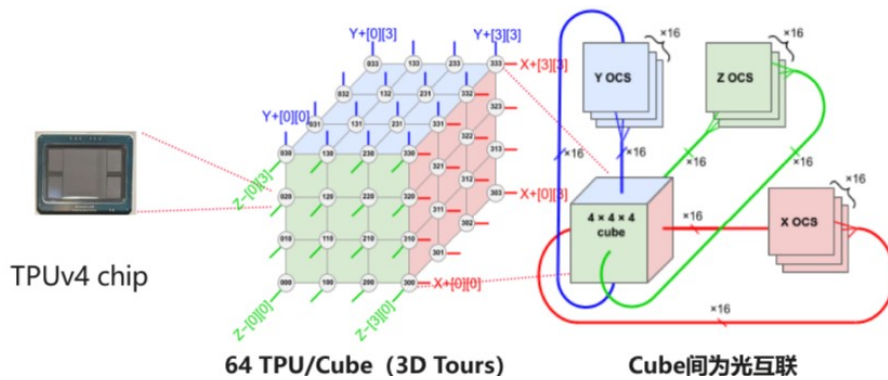


图 8 谷歌 TPU 的超节点架构

国内主要由华为主导，目前推出了 DC-0XC 解决方案，其通过上层的算网协同模块，对底层链路进行流量调度。通过在智算集群等场景实际测试，在降低延迟、降低功耗以及增加可靠性等方面，取得了较好的效果。

2.3.2 光分组交换

OPS 是光网络领域的远期技术路线，它描绘了一个带宽效率最高、灵活性最强的全光网络终极蓝图，其宏伟目标是在光域内，以单个分组数据作为为最小交换单元，实现全程无须光电转换的存储、路由和转发，如图 9 所示。但受制于光缓存/同步难题，目前暂未实现工程化。

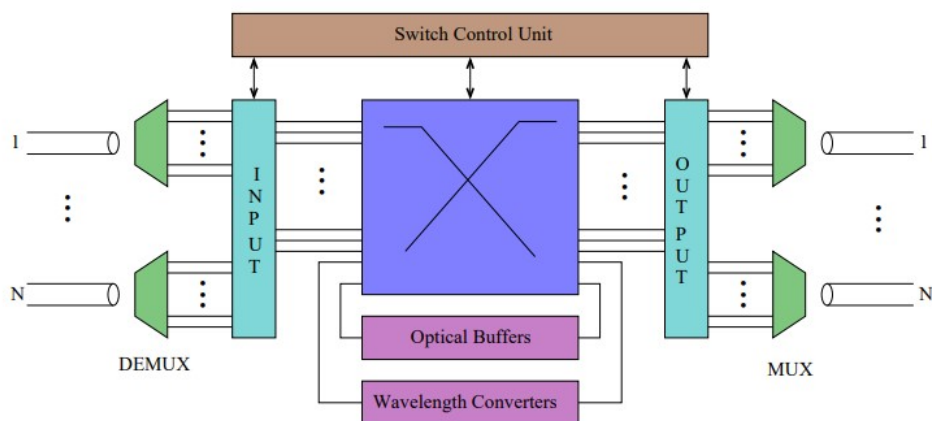


图 9 OPS 组成结构

2.3.3 光突发交换

OBS 的核心思想是一种折中与优化：它将数据流分割成较大的突发数据包（Burst），并在传输前先发送一个控制分组来为这个突发数据包预留网络资源。OBS 最关键的特点是数据在光域传输，而控制信令在电域处理。但由于缺乏光随机存储器，以及控制平面的复杂性等原因，该项技术尚未推出商业化产品。

3. 智算场景下光互连技术的应用研究

3.1 LPO 在 AIGC 算力网络中的应用

LPO 虽然采用 TIA 和 Driver 芯片替换 DSP，但同时引发了误码率提高的问题。行业普遍认为，LPO 只适用于特定的短距离应用场景。例如，数据中心机柜内服务器到交换机的连接，以及数据中心机柜间的连接等。目前，锐捷网络公司聚焦 AIGC 算力网络场景规划设计了三款自研 LPO 光模块，满足以下三种网络架构的互连需求，如图 10 所示。



图 10 锐捷 LPO 模块在三种网络架构下的互连应用

千卡 GPU 集群为 AI 训练和高性能计算提供算力，其内部数据交互需要高带宽、低延迟的网络互连。LPO 技术在此类场景中的应用优势主要体现在：能耗降低、成本减少以及良好的可维护性等方面。以下是千卡 GPU 集群场景下 LPO 模块的典型应用，如图 11 所示。

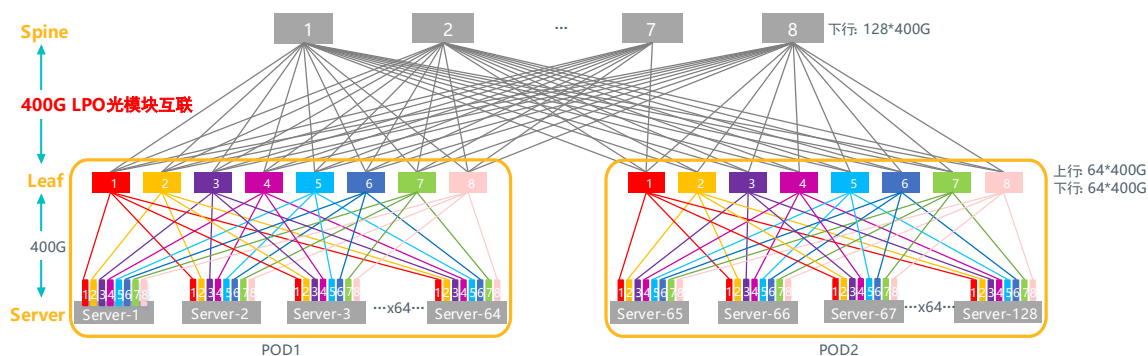


图 11 锐捷 LPO 模块在千卡 GPU 场景下的应用

场景规模为 128 台 AI 服务器，且每台搭载 8 卡 GPU 和 8 个单口 400G 网卡。采用二级组网，Spine 和 Leaf 互连两端可以使用 400G LPO 光模块。每 64 台 GPU 服务器为一个 POD，千卡集群共需 2 个 POD，16 台 Leaf 和 8 台 Spine 互连共需 2048 个 LPO 光模块。按功耗减低 50%（3.5W）估算，光模块功耗总共下降约 7kW。

3.2 CP0 交换机在智算场景下的应用

CP0 交换机具备高吞吐量、低时延、低能耗等多项技术优势。如图 12 所示，新华三 CP0 交换机是基于 51.2T 高性能芯片设计的 800G CP0 交换机，最大支持 64 端口 800G 端口，可扩展为 128 * 400G/200G/100G 端口，单 POD 可支持超过 8000 张 400G 网卡，最大可支持 3.2 万张 100G 网卡的计算集群。在 AIGC 智算场景下，可与业界最新的 800G GPU AI 芯片对接，实现最高规格的 AIGC 组网方案。该 CP0 交换机采用外置光源集成方案，即通过外部独立光源模块提供纯净光信号，将光源耦合到光通路后再耦合至调制器。在外部光源的数量设计上，该设备也有充分的技术考量，除了总体需要满足 512 通道 FR4 标准输出光功率外，还需要综合考虑单个外置光源模块的光功率规格问题，因此采用了 16 个 20dBm RLM 外置光源接口方案。

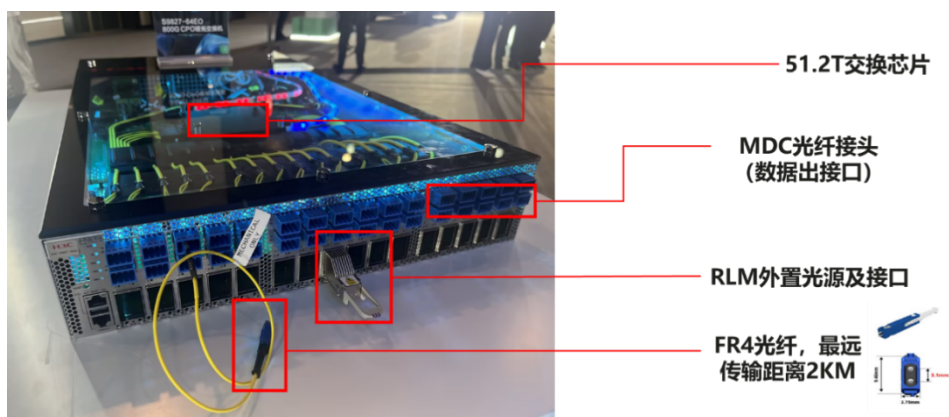


图 12 新华三 CP0 交换机前面板实物图

3.3 OCS 在 AI 集群参数面的应用

华为在 2024 全联接大会上发布数据中心全光交换机，旨在打造面向 AI 的新一代光电融合智算数据中心架构，将全光交叉技术引入数据中心内部。如图 13 所示，光交换机替代顶层电交换机，是智算场景下的典型组网。

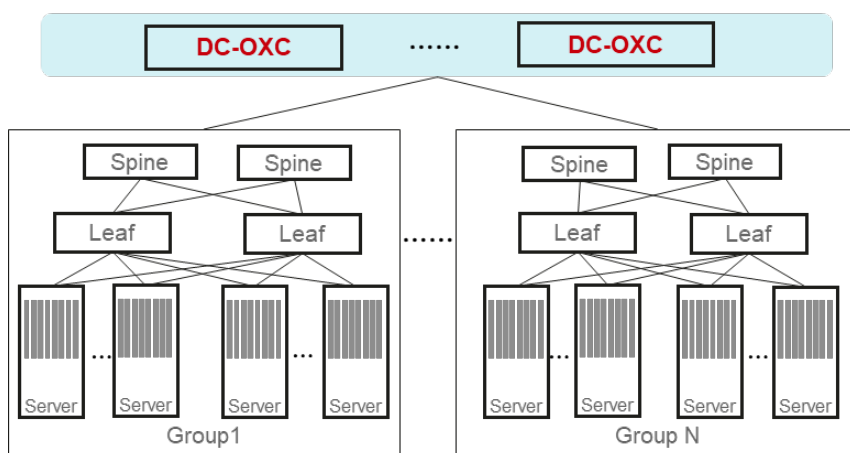


图 13 OCS 在智算场景下的应用

在该应用场景下，组网的收益点如下：

- 1) 大规模弹性组网：全光交换机端口密度高，支持按 POD 粒度分期建设，算力资源支持按需灵活组网，弹性可扩展。
- 2) 平滑演进：全光交换机本身协议无感知，支持向 800G、1.6T 甚至更高速率演进，可跨代际长期复用。
- 3) 绿色节能：全光交换机无光电转换，设备功耗百瓦级，网络功耗相比三层胖树降低 20%⁺。

- 4) 高可靠：全光交换机无需光模块，可以省去一层光模块，集群故障率可降低 15%⁺。

3.4 光互连技术在 GPU 超节点的应用

光跃 LightSphere X 是以分布式光交换（d0CS）芯片为核心的国内首个光互连 GPU 超节点解决方案。它的原理是取代传统机柜当中的交换机层，给每一个 GPU 模组配备一个光交换模组，使得单个 GPU 具备交换功能。光交换模组上同时封装了光引擎，可以和外部的互连光纤相连接，因此使得单个 GPU 具备了和任意 GPU 高速通信互连的功能，从而使得超节点的规模不再受到服务器或机柜的物理结构限制，在实际部署上更加灵活，可以根据模型算力需求灵活调整超节点的大小。另一方面，由于其带来的互连灵活性，使得在遇到单个 GPU 或服务器故障时，系统可通过软件的协同，迅速与闲置的 GPU 或服务器进行拓扑重组，大大提升了实际应用中的系统性能及 GPU 的利用率。



图 14 分布式光交换全光互连模组支持不同拓扑间的灵活切换

此外，对于存量数据中心中，单机柜电力无法提升的情况，分布式光交换全光互连方案可以在不改变数据中心供电条件的情况下，部署大规模光互连超节点，起到盘活存量数据中心的作用，如图 14 所示。

同时，光互连对于不同计算硬件的互连协议并不敏感，从而可以适配各类计算硬件，可灵活应用于各类存量或新建数据中心。

4. 移动云在智算场景下的光互连应用展望

随着 AIGC 智能化程度持续提升，大模型参数量不断增长，模型应用进入爆发式增长阶段，算力需求不断攀升，万卡乃至超十万卡规模的 GPU 集群成为训练大模型的标配，同时，在大模型训练时，海量 GPU 节点间需要持续、高速地同步数据和梯度，任何网络延迟或带宽瓶颈都会导致昂贵的 GPU 空闲等待，大幅降低整体计算效率。在今年混合专家模型流行之后，专家并行的加入还进一步提升了推理场景节点之间的通信量。以 DeepSeek 为例，官方给出的多机多卡策略中，模型需要部署在 4 个计算节点完成推理的预填充阶段，部署在 18 个计算节点完成推理的解码阶段。整体上看，大模型在参数量、模型框架上的进展都会推高计算节点之间的通信量，移动云需要在下一步智算中心的网络通信布局上充分考虑由此带来的通信高吞吐、低延迟、全连接需求。

智算中心的交换网络架构在带宽、时延及能效比等方面都遇到了提升瓶颈，这不仅影响着数据的传输，更会影响分布式计算的执行效率和系统可扩展性。智算中心网络的演进正经历着物理层基础技术的革新。传统的电分组交换机的交换容量受限于 SerDes 速率和 Crossbar 芯片规模，交换容量停滞在 51.2T 量级，光交换技术可以通过光域信号处理突破电互连的物理极限，成为未来智算中心网络架构演进的基石。当前的技术路径中，商用高速光模块已经实现 $4 \times 100\text{G}$ （400G）至 $8 \times 200\text{G}$ （1.6T）的传输能力，单通道速率突破 224Gbps。在研技术的单波 400G 光互连，有望提供 3.2T 光模块和 Pb 级别的交换容量。

在 Scale-Up 层面，移动云计划在 1~3 年的短期内，采用铜缆配合 CPO 光纤互连的方案：于近距离场景（ ≤ 7 米），优先采用铜缆互连方案（如，AEC 增强型铜缆），充分发挥低功耗、低成本及高可靠性优势；当传输距离超出铜缆有效传输范围（ > 7 米），优先选择 CPO 光纤互连方案，在满足长距传输的同时，最大程度降低功耗与成本。从 3~5 年长期来看，超节点 Scale-Up 网络互连将向端到端 CPO 互连方式演进。为满足不断增长的 Scale-Up 网络高带宽需求，XPU 普遍具备单位面积高密度端口，且端口速率更高。光引擎与 XPU 合封是极具潜力的技术方向，随着技术突破与成本下降，端到端 CPO 有望成为主流技术选择。在产业生态建设方面，移动云将加大与相关厂商的合作投入，包括封装厂（OSAT）封装能力建设、基板厂商提升大尺寸基板制作工艺及能力、连接器厂商提供通用的可拆卸光纤接口方案的合作等，通过全产业链生态逐步建立构建端到端全光互连芯片生态。

在具体的实施层面，移动云在超节点 Scale-Up 网络互连技术研发和部署方面会分场景、分阶段推进：针对整机柜超节点方案，推广正交矩阵互连方案或 AEC 铜缆互连方案；针对分机柜或级联超节点方案，优先推广一级和二级交换机之间 CPO 互连方案，为将来的端到端 CPO 化升级积累工程经验。随着 XPU 侧 CPO 技术成熟，逐步推进端到端 CPO 互连方案，即在一级交换机到二级交换机采用 CPO 互连的基础之上，进一步将 XPU 与一级交换机之间的互连方案从铜缆互连切换为 CPO 互连。

在 Scale-Out 层面，移动云将在未来考虑在各平面使用 OCS 替代原有的 Super Spine。OCS 替代 Super Spine 并非单一设备升级，而是移动云算网架构从“电为主、光为辅”向“全光原生”的范式

转变，本质是通过算网底层的全光重构，破解大规模 AI 算力集群的带宽瓶颈、延迟损耗与扩展桎梏，为“N+X”智算节点的弹性组网提供核心支撑。这一升级并非简单的硬件替换，而是覆盖数据、控制、管理多平面的系统性算网协同革新。随着移动云呼和浩特、贵阳等超大规模智算中心的落地，单集群 AI 加速卡规模已突破 2 万张，算力达 6.7EFLOPS，传统基于电交换的 Super Spine 在横向扩展中逐渐暴露性能、成本、扩展性的三重矛盾。未来大规模智算集群性能上限的突破将依靠 OCS 为代表的光互连模式，通过全光算网的 Scale-Out 能力，可支撑未来百万卡级智算集群的落地，为通用人工智能的发展提供底层算力底座。

5. 产业生态与标准化

5.1 光电领域互连标准

在光电领域互连标准中，目前主要囊括了：CP0、LP0、Chiplet 几个领域。

5.1.1 CP0 领域标准

在国际上 CP0 领域相关标准有：

OIF CP0 工作组发布了三份执行协议：2022 年 2 月发布的《Co-Packaging Framework Document》对光电合封系统框架进行了说明和定义。2023 年 3 月发布的《Implementation Agreement for a 3.2Tb/s Co-Packaged (CP0) Module》定义了用于以太网交换机的 3.2T CP0 模块，光口 FR4 和 DR4、电接口 32xCEI-112G-XSR、光机械模块规格、电气规格以及通过增强现有 OIF CMIS 规范来实现的控制和管理接口等。2025 年 1 月发布的《External Laser Small Form Factor Pluggable (ELSFP) Implementation Agreement》定义了前面板可插拔外部激光光源规格，以及对机械、热、电气和光学参数的互通性，标准的功率范围和光纤结构等进行了定义。

由 IPEC Form Factor 封装协议工作组起草的《OIO Pluggable External Laser Source (PELS) Implementation Agreement V1.0》，于 2024 年 9 月正式批准发布，该标准规范了用于网络设备应用的 OIO 可插拔外置激光光源的外形、光学接口、电气接口、热管理与通用管理接口等，适用范围包括 CP0、NP0 及 OIO 等场景。

在国内，中国科学院计算所联合电子标准院牵头制定了《半导体集成电路 光互连接口技术要求》(T/CESA 1266-2023)标准草案，

并于 2023 年 7 月正式发布。这是我国自主制定的首个面向 CPO 场景的标准，描述了用于微电子芯片光互连接口的共封装模块的设计技术要求，包括概念说明、电学特性、光学特性、数字管理接口、机械结构设计要求等。同月，中国科学院计算所还定义了 CPO 外置光源池 ELS 标准，明确了 ELS 的应用场景、标准规范、关键技术及典型产品形态。该标准通过统一提供高功率、稳定的外部激光光源，为 CPO 等新型互连架构在交换机和数据中心中的应用提供支撑，内容涵盖 ELS-DR 与 ELS-FR 的封装、波长、光功率、功耗等参数，并提出了高效率耐高温激光芯片、双透镜高效耦合、电流源反馈控制等关键技术，同时给出了 CWDM、PSM、DWDM 等典型产品规格。

5.1.2 LP0 领域标准

在 OFC 2025 大会期间，LP0 MSA 正式发布了其《100G-DR-LP0 单模线性可插拔光学数据传输规范》。该规范定义了构建 LP0 兼容的交换机、网卡（NIC）和光模块产品生态所需的光学和电气要求，覆盖了 100Gbps、200Gbps、400Gbps 和 800Gbps 的以太网并行单模链路，旨在解决行业普遍面临的降低功耗、成本和时延的挑战，同时确保用于 AI/ML 应用的高速光互连的可靠性。2025 年 2 月举行的互操作性测试活动中，来自系统、模块和芯片制造商的参与者共同展示了高于规范要求的链路性能裕量，并成功实现了与多样化链路伙伴的互连互通。此外，LP0 MSA 已经启动了 200Gbps/通道线性可插拔方案的规划工作，并计划与标准组织如 IEEE、OIF 等合作，在更高速率下继续推动降低功耗、降低成本和减少延迟的技术实现。

5.1.3 Chiplet 领域标准

近年来，国内外在 Chiplet 标准化方面同步推进。国内发布的

《小芯片接口总线技术要求》奠定了芯粒互连规范基础，兼顾现有协议与本土工艺；国际上，UCIe 自 2022 年 1.0 版本起持续迭代至 2025 年 3.0 版本，逐步强化可靠性、可管理性、3D 封装及高速互连能力。两者共同推动 Chiplet 生态走向互操作、低成本与高性能。

2022 年 12 月，中国首个原生 Chiplet 技术标准发布《小芯片接口总线技术要求》（T/CESA 1248-2023），这项标准描述了 CPU、GPU、人工智能芯片、网络处理器和网络交换芯片等应用场景的小芯片接口总线（Chiplet）技术要求，包括总体概述、接口要求、链路层、适配层、物理层和封装要求等，以灵活应对不同的应用场景、适配不同能力的技术供应商，通过对链路层、适配层、物理层的详细定义，实现在小芯片之间的互连互通，并兼顾了 PCIe 等现有协议的支持，列出了对封装方式的要求。小芯片设计不但可以使用国际先进封装方式，比如 CoWoS，也可以充分利用国内封装技术积累，实现一种或者几种成本低廉、重点针对 Chiplet 芯片架构、可以覆盖 80% 以上应用场景的先进封装手段。

在国际方面，UCIe 1.0 规范，于 2022 年 3 月正式发布，定义了完整的标准化 Die-to-Die（芯粒间）互连，涵盖物理层、协议栈、软件模型和合规性测试。该规范使终端用户能够在多供应商生态系统中灵活组合 Chiplet，用于构建标准或定制化的片上系统（SoC）。

UCIe 1.1 规范，于 2023 年 8 月正式发布，在 Chiplet 生态中进行了功能增强，扩展了可靠性机制，支持更多协议与更广泛的应用模型。规范增加了汽车电子场景相关的改进（如预测性故障分析和运行健康监测），并引入了更低成本的封装实现方式。同时，它补充了架构规范属性，用于定义系统配置和测试寄存器，以支持合规性测试和互操作性验证。该规范与 UCIe 1.0 完全向后兼容。

UCIe 2.0 规范，于 2024 年 8 月正式发布，引入了标准化的系统架构以增强可管理性，全面解决了多芯粒系统封装（SIP）生命周期（从芯片测试到现场运维）中的可测试性、可管理性和可调试性（DFx）挑战。规范新增了可选的可管理性功能，并提出 UCIe Dfx 架构（UDA）：在每个 Chiplet 中嵌入管理结构，实现测试、遥测和调试功能，从而通过统一的 SIP 管理和 Dfx 方法，实现供应商无关的互操作性。此外，UCIe 2.0 增加了 3D 封装支持，相比 2D/2.5D 架构，显著提升带宽密度与能效。UCIe-3D 针对混合键合（hybrid bonding）优化，凸点间距可从 10 - 25 μm 缩小至 1 μm 甚至更小，具备更高灵活性与可扩展性。该规范与 UCIe 1.0、1.1 完全兼容。

UCIe 3.0 规范，于 2025 年 8 月正式发布，标志着开放式 Chiplet 标准演进的新阶段。该规范显著提升性能，新增 48GT/s 和 64GT/s 数据速率，并引入架构更新，以满足业界对高速、可互操作 Chiplet 互连的需求。改进包括更高的带宽密度、能效和系统级可管理性，推动可扩展的多芯片 SIP 设计，加速模块化半导体创新。

5.2 光电领域交换标准与产业生态

5.2.1 光交换标准发展现状

在光交换领域，国际上，OCS 器件相关标准主要由国际电信联盟电信标准化部门（ITU-T）承载。2023 年 2 月，在 ITU-TSG15 Q6 中间会上，联通文稿《Discussion of AWGR used for hybrid optical and electronic switch》，率先介绍了 $N \times N$ AWGR 结合可调谐激光器在数据中心实现光电混合交换的应用场景和架构。在 2023 年 4 月的 ITU-T SG15 Q6 全会上，联通进一步探讨了 AWGR 与可调谐激光器的结合在光交换中展示了极大的潜力，能够提供灵活、

高效的交换能力，并增强网络的适应性和性能，并建议在 G.671 中立项修订 $N \times N$ AWGR 和可调谐激光器标准。在 2023 年 7 月，ITU-T 中间会议上，会议通过了 $N \times N$ AWGR 的在 G.671 的标准立项修订，周期一年。目前该标准增补工作已在 2025 年 3 月的 ITU-T SG15 全会上完成发布。

OCS 配套光模块主要由 IEEE 802.3、OIF 等标准组织共同制定。2024 年 11 月，OIF Q4 会议上，由 Meta，Google 牵头立项 1.6T coherent light (CL) 项目，涵盖 OCS 场景，最大传输距离为 2km，最大链路损耗定义为 8dB，对模块功率的要求与 1.6T-FR4 直检光模块相当，时延小于 300ns，需前向兼容未来 3.2T CL 光模块。在 2025 年 5 月的 OIF Q2 会议上，由 OCS 器件厂商 Lumentum 牵头，明确将 O 波段 1.6T LR/OCS 应用场景写入 1.6T CL 项目 Editor' Worklist。截至目前，该项目仍在讨论中，预计 2025 年 Q4 完成初版 baseline 定义。

国内的相关标准主要有以下的进展：2021 年 12 月，CCSA TC6 WG1 通过了《面向下一代数据中心内部网络光交换技术研究》立项申请，该研究报告主要开展下一代数据中心内部网络的需求及背景，系统架构和关键技术的要求，快速切换相关技术和研究，光交换机相关技术和研究，相关控制策略和协议等研究。在 2024 年 Q4 CCSA TC6 WG1 结题印发了该研究报告，引起了业界对 OCS 在算力中心网络应用的热点讨论。

2024 年 12 月，CCSA TC6 WG1 又通过了《面向智算的光交换应用场景和组网技术研究》研究报告立项，意在对面向智算中心的光交换应用场景研究和面向智算中心的光交换组网技术，包括设备的功能性能需求、组网要求、管理控制要求等方面做进一步研究，该

研究报告预计 2027 年 Q1 结题。同年 12 月，CCSA TC6 WG4，也立项了《面向智算中心的光交叉模块研究》，该研究报告面向智算中心，开展光交叉模块技术研究，包括：研究智算中心场景对光交叉功能和性能需求，研究端口级和波长级光交叉技术方案，并给出发展建议，该研究报告预计也会在 2027 年 Q1 结题。

在行标方面，2023 年 12 月，CCSA TC6 WG4 通过了《平面光波导集成光路器件 第 4 部分：阵列波导光栅路由器 (AWGR)》和《 $N \times N$ 阵列光交换矩阵开关》两个项目的立项申请，同意制定基于 AWGR 的光电融合交换方案中 AWGR 器件和基于 $N \times N$ 全连接交叉矩阵光器件标准。《平面光波导集成光路器件 第 4 部分：阵列波导光栅路由器 (AWGR)》定义了 AWGR 相关内容，像通道波长的选择、差损的限制和通道串扰等关键指标。该标准是国内首个关于数据中心光交换器件的标准；与此同时，《 $N \times N$ 阵列光交换矩阵开关》将规范大端口阵列光交换矩阵开关的技术要求，主要包括光学特性、外形结构、测试方法、可靠性试验、检验规则及标志、包装、运输、贮存等要求，主要面对针对数据中心应用场景。这两个行标，预计 2025 年 Q4 完成送审稿评审。

2024 年 9 月，ODCC 发布 AI 网络光交换机技术报告，该研究项目由美团牵头，中国信息通信研究院和华为公司共同参与贡献。研究报告深入研究和探讨了光交换技术在数据中心网络中的应用前景和挑战。通过对光交换机技术的基本原理、发展历程和应用场景的详细介绍，以及与其他交换技术的比较，帮助读者全面了解光交换机技术的优势和局限性。2025 年 6 月 ODCC 夏季全会上，华为公司专家宣讲了名为“ETH-X DC-OXC 使能 Scale UP 保护资源池化”的特邀报告，详细介绍了超节点内，基于 DC-OXC 的保护资源池化方案

的细节，及相较于现有冗余备份方案的主要优势。

5.2.2 光交换产业生态进展

2025 年 7 月，OCP 产业组织宣布成立 OCS 子项目，旨在促进开放 OCS 技术的协作，以满足人工智能等数据密集型应用中对高带宽、低延迟和能效日益增长的连接需求。该项目涵盖目前国际所有主流 OCS 设备厂家及北美客户，项目由 OCP 成员 iPrionics 和 Lumentum 共同领导，初始参与者包括 Coherent、Google、Lumotive、Microsoft、nEye、NVIDIA、Oriole Networks 以及 POLATIS (HUBER+SUHNER)。项目主要聚焦 OCS 配套光网络软件定义。在 2025 年 8 月的 OCP 亚太峰会 (OCP APAC) 上，OCS 子项目作为四大进展之一，首次亮相，后续将以月度例会形式持续推动进展。

6. 发展趋势与发展建议

6.1 发展趋势

在可插拔光模块产品方面，LP0 技术主要聚焦光模块内部架构革新，通过去除 DSP 芯片或仅保留部分重定时功能，降低功耗与延迟。该项技术目前已实现了初步商业化，并在 112G SerDes 场景下完成部署。在光电混合与共封装技术方面，NP0/CP0/OIO 技术通过不断缩短光引擎与算力/交换芯片的电路距离来实现对能耗限制的突破。NP0 作为可插拔光模块与 CP0 之间的过渡方案，目前较为热门，是一种比较稳妥的技术路线选择。OIO 技术落地需重构芯片架构，因此与 GPU/CPU 厂商深度绑定，预计待 2028 年后硅光芯片良率突破以及 2.5D/3D 封装技术成熟后，将逐步得到发展。长期来看，OIO 技术终将取代电互连成为算力芯片的“神经束”，尤其在 3D 堆叠 GPU、存算一体架构中，光替代电进行传输是突破“功耗墙”与“带宽墙”的唯一路径。中国企业需在 CP0 阶段积累硅光集成能力，同步布局 OIO 芯片级光互连技术，方能在未来占据优势。

此外，作为下一代智算中心互连技术的核心，在全光交换领域，其发展趋势主要集中在软硬件协同控制，标准化建设和多技术路径竞争等方面。在多技术路径层面，MEMS 微镜阵列正在成为当前的主流方案，目前该技术主要还存在机械稳定性与良率、插入损耗等方面的问题。在软硬件协同控制层面，全光交换与软件定义网络的深度融合，在大型科技企业的引领下，已发展到全新的高度。

6.2 产业发展建议

在 LP0 研发方面，短期来看，基于 112G SerDes 的 LP0 技术已

经通过了实际验证，并已开始逐步商用化部署，技术上具有可行性，并受到各大厂商的关注。长期来看，224G SerDes 是下一代网络升级的方向，尚有诸多的问题亟需突破。在智算场景中，建议 LP0 模块部署于网络交换设备间以满足低功耗需求。在 LP0 的标准化方面建议针对互操作性、信号质量、诊断功能等制定强制性规范，推动 LP0 MSA 标准完善，同时建议由云服务厂商或者运营商牵头建立多厂商联合实验室，促进相关的生态体系建设。

对于 NP0/CP0/OI0 光电共封装技术，建议加速制定 CP0/OI0 的通用接口、光引擎封装尺寸及散热管理标准，解决多厂商互操作性问题，以智算中心场景为核心，推动 NP0/CP0 的应用落地，解决发热、芯片良率等问题，同步验证 OI0 在 GPU 互连领域的应用效果。

对于光交换领域，在 OCS 方向，目前已经实现初步商业化。光交换作为支撑超大规模 Scale-Out 架构的关键基础设施，在智能中心中的应用前景十分乐观，短期来看，OCS 最有可能进行小范围的试点应用，长期来看，光分组交换将具有更大的潜力。

附录:

常见缩略语

缩略语	英文全名	中文解释
AI	Artificial Intelligence	人工智能
AIGC	Generative artificial intelligence	生成人工智能
ASIC	Application-Specific Integrated Circuit	专用集成电路
CPO	Co-Packaged Optics	光电共封装
DC-OXC	Data Center-Optical Cross-Connect	数据中心全光交叉
DLBS	Direct Light Beam Steering	直接光束偏转
DLC	Digital Liquid Crystal	数字液晶
DSP	Digital Signal Processor	数字信号处理器
EFLOPS	Exa Floating-point Operations Per Second	百亿亿次浮点运算 每秒
GPU	Graphics Processing Unit	图形处理器
IPEC	International Photonics & Electronics Committee	国际光电委员会
LGA	Land Grid Array	栅格阵列封装
LPO	Linear-drive Pluggable Optics	线性驱动可插拔光 模块
LRO	Linear Receive Optics	线性接收光学
MEMS	Micro-Electro-Mechanical System	微机电系统
MSA	Multi-Source Agreement	多源协议
NPO	Near-packaged Optics	近封装光学
OBO	On-Board Optics	板上光学
OBS	Optical Burst Switching	光突发交换
OCP	Open Compute Project	开放计算项目

OCS	Optical Circuit Switch	光线路交换
ODCC	Open Data Center Committee	开放数据中心委员会
OIF	Optical Internetworking Forum	光互连论坛
OIO	Optical Input/Output	光纤输入/输出
OPS	Optical Packet Switching	光分组交换
PCB	Printed Circuit Board	印刷电路板
POD	Point of Delivery	分发点
RLM	Remote Light Module	远端光源模块
SerDes	Serializer/Deserializer	串行器/解串器
TIA	Trans-Impedance Amplifier	跨阻放大器
TPU	Tensor Processing Units	张量处理单元
UCIe	Universal Chiplet Interconnect Express	通用小芯片互连通道
XPU	eXtended Processing Unit	泛指各种处理器