

doi:10.13756/j.gtxyj.2024.240029.

专题:数据中心内光交换

刘璐,吴冰冰.片间光互连发展态势分析[J].光通信研究,2024(5):240029.

Liu L, Wu B B. Analysis of the Development Trend of Inter-Chip Optical Interconnection[J]. Study on Optical Communications, 2024(5):240029.

## 片间光互连发展态势分析(特邀)

刘璐,吴冰冰

(中国信息通信研究院 技术与标准研究所,北京 100191)

**摘要:**随着各行业数字化转型升级进度加快,大数据和云计算等新技术的迅速普及应用,尤其是近两年来,人工智能大模型热潮兴起,带动算力需求快速增长。当前,电子信息技术遭遇带宽和能耗挑战,摩尔定律面临失效危机。相比于电信号,光信号具有传输带宽大、传输损耗小、抗干扰能力强和可并行传输等诸多优势,利用光进行互连成为信息技术发展的重要方向。当前,“光进铜退”趋势持续,光互连的应用场景从机架、板卡进入芯片。在单个芯片封装中集成光子和电子器件,不仅可以提高集成度和端口密度,还可以实现低能耗以及低时延。如何发展片间光互连,成为当前的研究热点之一。文章面向数字时代需求,针对算力基础设施两大典型应用场景——数据中心以及计算中心,梳理了片间光互连的最新技术动态,分析了产业生态与标准化进展,并研判其发展趋势。文章从技术、产业和应用等角度进行了深入探究,明确光互连可以从带宽、能耗和时延等方面有效提升数据中心中交换芯片与外部以及计算中心中计算芯片与外部之间的互连性能。当前片间光互连已取得初步进展,相关技术与标准化研究热度持续上升,产业得到初步落地。相较于数据中心,计算中心片间光互连技术的性能要求更高,产业化标准化进度更为滞后。整体来说,光互连的短距化将使光子进一步发挥自身优势和挖掘应用潜力,支撑信息通信技术的持续进步。

**关键词:**数据中心;光电合封;光输入输出;光互连

**中图分类号:** TN929

**文献标志码:** A

## Analysis of the Development Trend of Inter-Chip Optical Interconnection

LIU Lu, WU Bingbing

(Institute of Technology and Standards, China Academy of Information and Communication Technology, Beijing 100191, China)

**Abstract:** With the acceleration of digital transformation and upgrading across various industries, the rapid popularization and application of new technologies such as big data and cloud computing have driven the exponential growth in demand for computing power, particularly with the emergence of artificial intelligence large models in recent years. Currently, electronic information technology faces challenges related to bandwidth limitations and energy consumption issues, while Moore's law is confronted with potential failure. In comparison to electrical signals, optical signals offer numerous advantages including larger transmission bandwidth, minimal transmission loss, robust anti-interference capabilities, and parallel transmission. Consequently, optical interconnection has become a crucial direction for information technology development. The ongoing trend of "optical advance and copper retreat" continues as optical interconnection applications transition from racks and boards to chips. Integrating photonic and electronic devices within a single chip package not only enhances integration levels and port density but also achieves lower power consumption rates along with reduced latency periods. Developing inter-chip optical interconnections has emerged as one of the current research topics. The paper analyzes the latest technological trends in optical interconnects between chips, focusing on two key application scenarios of computing infrastructure: data centers and computing centers. It also analyzes the progress of industrial ecology and standardization, while evaluating its development trajectory. This paper conducts a comprehensive study from the perspectives of technology, industry, and application, elucidating that optical interconnection can effectively enhance the interconnection performance between the switching chip and external devices in data centers. It can also improve the performance between the computing chips and external devices in computing centers, with regards to bandwidth, energy consumption, and delay. Currently, there has been initial progress in inter-chip optical interconnection, accompanied by increasing research interest in related technologies and standardization efforts within the industry. However, compared to data centers, inter-chip optical interconnection technology for computing centers has higher performance requirements while lagging behind in industrialization and standardization progress. Overall, short-distance optical interconnections will further leverage their inherent advantages to explore application potential and support continuous advancements in information and communication technology.

**Key words:** data center; co-packaged optics; OIO; optical interconnection

收稿日期:2024-02-06; 修回日期:2024-03-12; 纸质出版日期:2024-10-10

基金项目:国家重点研发计划资助项目(2021YFB2800203)

作者简介:刘璐(1992—),女,安徽宿州人。高级工程师,博士,主要研究方向为光电子芯片器件。

通信作者:刘璐,高级工程师。E-mail:liulu@caict.ac.cn

© Editorial Office of *Study on Optical Communications*. This is an open access article under the CC BY-NC-ND license.

0 引言

人工智能应用和大模型训练等新应用快速发展,对算力提出更高要求。全球算力规模不断扩大,据《中国算力发展指数白皮书(2023 年)》,2022 年全球计算设备算力总规模达到 906 EFlops,增速达到 47%。算力的提升可以通过计算、存储和互连 3 方面的技术演进升级来实现。就互连方面,“光进铜退”的趋势从机柜、板卡延伸到芯片,光互连可利用光的低功耗、高带宽和低延迟等优势,部分替代传统的电互连方案<sup>[1]</sup>。包含片间光互连与片上光互连这两种形式的芯片级光互连正进入一个技术快速进步、产业快速发展的时期,当前行业巨头如英特尔和英伟达等已对交换和计算等场景中的芯片级光互连进行了布局,各大高校研究机构与国际标准组织也已开展大量工作,业界对此技术关注度持续上升<sup>[2-4]</sup>。不过当前芯片级光互连的发展现状、问题挑战和发展趋势等的研究尚不完善,有待进一步深入分析,以明确其应用潜力并找准发力点。

本文聚焦片间光互连,对算力基础设施两大典型应用场景—数据中心以及计算中心中的片间光互连技术分别展开研究,梳理其技术、产业和标准化进展,从带宽、能耗和时延等方面分析光对于互连性能的提升效果,明确这两种场景下片间光互连的具体差异,最后给出了发展片间光互连技术的 3 方面建议。

1 数据中心场景中的片间光互连技术发展态势

数据中心场景中的片间光互连技术聚焦光电合封(Co-Packaged Optics, CPO)技术。CPO 面向大型以太网网络交换机,将交换芯片与光芯片封装在一起,与外部其他芯片之间形成光互连。

传统可插拔光模块方式面临端口密度和能耗挑战,CPO 技术助力数据中心升级。近年来,随着高性能计算机和大型数据中心对数据传输带宽的要求越来越高,交换芯片的吞吐量将向 25.6 Tbit/s 甚至 51.2 Tbit/s 演进。传统技术采用可插拔光模块来实现交换机之间高带宽数据传输。而受交换机面板端口密度的限制,可插拔光模块已经无法以高密度的方式支撑 51.2 Tbit/s 数据交换。此外,当信道速率达到 112 Gbit/s 及以上时,可插拔光模块带来的较长信号传输距离导致交换芯片和光模块功耗和延迟增加,无法高效实现超大数据交换。CPO 技

术应运而生,根据 Broadcom 数据,可插拔光模块功耗为 15~20 pJ/bit,而 CPO 系统功耗可降低 50% 以上,达到 5~10 pJ/bit。有仿真表明,在全对全通信模式下,完成时间最多减少 40%<sup>[5]</sup>。与交换专用集成电路(Application Specific Integrated Circuit, ASIC)共同封装的光引擎可能会为大型数据中心中的可插拔光收发器提供替代方案。如图 1 所示,交换 ASIC 芯片和光引擎(光学器件)在同一高速主板上协同封装,光学引擎与主机 ASIC 之间距离缩短,可以大幅减少电学损耗,从而降低信号衰减、降低系统功耗、降低成本和实现高度集成<sup>[6]</sup>。有研究表明,若在交换机和服务器处使用 CPO 技术,可使网络容量增加一倍,同时减少 64% 的交换机数量<sup>[7]</sup>。

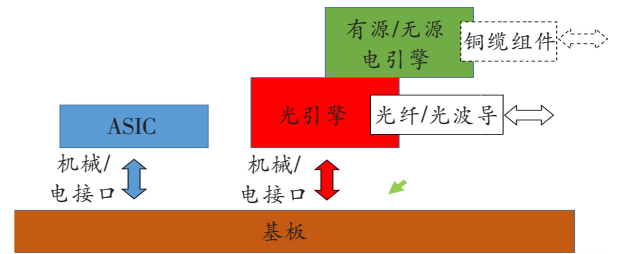


图 1 CPO 结构

Figure 1 The structure of CPO

系统级技术向更高集成度演进,硅光集成方案成为主流。在 CPO 系统中,光引擎围绕在交换芯片四周。光引擎由电芯片、光芯片以及其他控制单元构成。可插拔模块时代,光源、调制器和探测器通常分立封装,但传统的分立封装架构中,调制器和探测器异质异构设计无法满足高通量同时小型化的光引擎设计需求,系统级技术向更高集成度演进成为必然趋势。目前报道的技术方案主要为垂直腔面发射激光器(Vertical Cavity Surface Emitting Laser, VCSEL)阵列方案和硅光集成方案。基于硅光集成的单模方案具有无需气密封装、高带宽和易集成等优势,为当前的热点和主流方案,相关研究持续开展<sup>[8-10]</sup>。不过 VCSEL 方案在短距离信号互连中的成本和功耗优势明显,其研究也在持续进行<sup>[11]</sup>。英特尔在 2024 年的 ISSCC 会议上发布了其最新 CPO 进展,可实现 4×64 Gbit/s 的信号传输,系统功耗仅 1.3 pJ/bit。

集成光子技术方案中的光源集成一直是核心技术挑战之一,外置光源成为趋势。使用外置光源的原因来自于材料的本征特性:硅、氮化硅和二氧化硅都是间接带隙材料,单靠自身材料实现电致激射十分具有挑战性,因此需要通过其他的方式引入光源。

关于将光源与调制器连接,技术方向有两种,即光源与调制器进行片上集成,称为“光源内置”或将光源独立封装;光源耦合到光通路后再耦合调制器,称为“光源外置”。外置光源降低了光收发单元的热量,从而降低了光模块的功耗,是目前各家单位较为倾向的一种放置方式。

近 3 年 CPO 样机相继发布,产业各方积极推动。2020 年,在美国光纤通讯展会上,英特尔公司发布了 CPO 的首款样机;2021 年,Ranovus 公司发布了 CPO 2.0 架构,降低了 40% 的功耗和成本;2022 年,Marvell 公司展示了其带宽为 1.6 Tbit/s 的 CPO 样机,未来将支持其 51.2 Tbit/s 的交换机芯片;Ranovus 公司与 AMD 公司面向数据中心人工智能应用研发了新的 CPO 系统并进行了共同展示;2022 年,Broadcom 和腾讯达成战略合作伙伴关系,将提供 25.6 Tbit/s CPO 交换机以加速 CPO 在云基础设施中的应用;2023 年,在美国光纤通讯展会上,Broadcom 和 Marvell 均发布了 51.2 Tbit/s 的交换芯片;2021 年 1 月,亨通洛克利凭借与英国洛克利光子合作的优势,开发出了基于硅光技术的 3.2 Tbit/s CPO 工作样机,属于国内首台;旭创、海信、光迅和新易盛等公司表示在 CPO 领域有技术储备。总体来看,绝大多数涉及 CPO 的厂商在美国,国内产业发展略有滞后。

CPO 远期市场景气,产业链条初步形成。据咨询机构 Yole 预测,CPO 市场将从 2022 年的 600 万美元增长到 2033 年的 2.87 亿美元。Marvell 认为,CPO 是应对未来高速互连需求的唯一选择。Broadcom 表示,CPO 应用的时间节点可能比 5 年更早,对于一个包含 256 K 图形处理器(Graphics Processing Unit,GPU)的计算系统,CPO 相对于线性驱动可插拔的功率可以减少约 1 MW。CPO 与硬件捆绑,需要很强的硬件平台和工艺设计能力,从可插拔到 CPO 的转换对企业研发实力提出了较高要求,当前 CPO 产业链包含设计、光引擎供应商、激光器供应商、交换机厂商、硅光代工厂和设备商等,预计不会形成封闭垄断的生态系统。首先在 CPO 时代当前已至少有 4 个玩家:英特尔、Broadcom、英伟达和 Cisco,其次终端用户考虑成本,提倡白盒、解耦,倾向于开放的生态系统,预计后期参与厂商会进一步增多。总体来看,CPO 当前研究聚焦 51.2 Tbit/s 交换机,预计大量商用的时间点应该在未来 3~5 年。

大型国际标准组织作出系列部署,CPO 标准体

系初步建立。光互连论坛(Optical Internetworking Forum,OIF)于 2020 年 11 月批准启动 CPO 框架实施协议伞形项目;2023 年 3 月发布了用于数据中心内部交换应用的 3.2 Tbit/s CPO 模块实施协议;2023 年 8 月发布了外置光源小型可插拔模块实施协议支持 CPO 应用。此外,OIF 在 2023 年的欧洲光通讯展会上进行了 CPO 外置光源小型可插拔模块展示以及外置光源与盲配合连接器的互连展示;2020 年 12 月,板载光学联盟成立 CPO 工作组,2022 年发布 CPO 交换机设计白皮书;联合开发基金会发布了 3.2 Tbit/s CPO 的产品需求、CPO 外置光源和 CPO 组件等 3 份白皮书;国际光电委员会已于 2022 年 1 月启动研究项目,正在起草 100 Tbit/s + 研究报告;中国计算机互连技术联盟于 2023 年 7 月发布了团标《半导体集成电路光互连接口技术要求》。整体来看,当前对于 CPO 框架与外置光源的研究已基本完成,预计下一步的工作重点在于测试规范。

## 2 计算中心场景中的片间光互连技术发展态势

计算中心场景中的片间光互连技术聚焦光输入输出(Optical Input Output,OIO)技术。OIO 面向分布式计算系统,将计算芯片与光芯片封装在一起,与外部其他芯片之间形成光互连。

传统方式无法满足高性能计算互连需求,OIO 技术成为下一代计算中心的重要路径。高性能计算工作负载不断推动现有系统架构极限,随着系统规模的扩大,需要计算芯片之间建立更高密度带宽的连接,传统互连方式已力不从心。一方面,传统可插拔光学器件的物理尺寸较大、能耗较高,无法有效满足生成式人工智能的需求;另一方面,传统可插拔光学的成本在每 Gbit/s 1~2 美元,而对于生成式人工智能来说,需要降至约 1/10 才能具有成本效益。根据 Lockheed Martin 公司数据,相比于传统商业解决方案,OIO 可将数据传输带宽提升 7 倍,功耗降低为 1/5,尺寸降低为 1/12,大幅提升互连性能,满足高性能计算场景需求,同时也为资源池化提供了性能保证。

芯粒为片间互连提供可行的产品形态,OIO 需要结合芯粒与硅光技术。芯粒现已成为芯片设计行业的主流技术路径,其为片间互连设备提供了合理的产品形态,同时也解决了资源池化架构下的带宽、延迟和能耗等问题。芯粒需要先进的封装集成技



术,其中,晶圆级封装技术使用高密度的细间距微球垂直互连多个芯片,可大幅提升带宽密度;三维封装技术利用硅通孔等可进一步增加芯片间的互连密度。而硅光技术能够将光收发器、调制器与控制逻辑集成在同一封装中。通过这些技术的创新结合,OIO 可将数百万个晶体管与数百个光子器件集成在一起,实现以低于 5 pJ/bit 的功耗驱动数公里传输距离的数十 Tbit/s 带宽,支撑创建跨机架扩展的逻辑上连接、物理上分布的计算架构。

计算领域各大巨头均已进行布局,OIO 相关产品初具雏形。Ayar Labs 在 OIO 方面具有较强实力,2022 年首次验证了采用 OIO 技术的 5.12 Tbit/s 现场可编程门阵列(Field Programmable Gate Array,FPGA),该系统集成了 1 个 FPGA 芯片和 5 个 OIO 芯片<sup>[12]</sup>;2023 年,Ayar Labs 与英特尔 FPGA 进行了集成 OIO 解决方案的演示,实现了双向 4 Tbit/s 数据传输;英特尔基于其长期的硅光积累,提出了一种适合集成于可扩展处理单元中的 256 Gbit/s 三维集成硅光接收机<sup>[13]</sup>;AMD 在 2023 年国际固态电路会议上展示了一种使用硅中介层将 GPU 与高带宽存储器集成的方案,并探讨了将计算与内存堆叠的前景;英伟达与 Ayar Labs 具有合作关系,其内部也正在开发硅光技术以实现未来 GPU 系统高通量、高能效的光学 NVLink 连接;三星在 2023 年开放计算全球峰会上提出了在高带宽存储器与逻辑芯片间采用 OIO 技术进行数据互连的构想;曦智科技在 2023 年全球闪存峰会上发布了其首款适用外设组件互连快速总线(Peripheral Component Interconnect Express,PCIe)和计算快速连接(Compute Express Link,CXL)协议的数据中心光互连硬件产品,并现场演示了内存扩展光互连解决方案。

OIO 远期市场景气,产业链条初步形成。据咨询机构 Yole 预测,OIO 市场将从 2022 年的 500 万美元增长到 2033 年的 23 亿美元。据咨询机构 Hyperion Research 数据,业界普遍认为光互连能力将在未来两年内迅速提升至 10 Tbit/s 以上,比当前电互连提升 10 倍。当前 OIO 产业链包含激光器供应商、计算芯片公司、硅光代工厂和服务器厂商等。当前已有产品小批量出货,预计其商业生态完全爆发时间在 5 年以后。

OIO 标准研制尚处初期,物理层软件层研究同步开展。随着 OIO 的成熟,行业标准对生态系统发展的牵引地位愈发重要。连续波波分复用多源协议

(Continuous-Wave Wavelength Division Multiplexing Multi-Source Agreement,CW-WDM MSA)旨在对密集带宽的光波长进行标准化,其定义了一组 O 波段的波长网格,符合规范的激光器之间可实现互操作<sup>[14]</sup>。外设组件互连特别兴趣组(Peripheral Component Interconnect Special Interest Group,PCI-SIG)于 2023 年 8 月宣布成立 PCIe 光学工作组,研究为 PCIe 规范引入光学传输接口的可能性,或将开发适用于光学接口的新外形尺寸。通用芯粒互连快速总线(Universal Chiplet Interconnect Express,UCIe)规范对物理层、适配层、协议层以及晶粒到晶粒接口进行了定义,制造商可以基于 UCIe 规范实现芯粒之间的光互连。CXL 规范规定了内存缓存一致性,允许中央处理器(Central Processing Unit,CPU)和加速器之间共享内存资源,有效支撑计算资源和存储资源的池化和解耦。相比以太网协议,CXL 协议提供了高效的数据同步,点对点的传输延迟可以从以太网的 10  $\mu$ s 量级减少到 100 ns 量级。OIF 已于 2023 年启动节能接口项目,对 CW-WDM 技术以及 PCIe、UCIe 和 CXL 等的光学接口启动研究。CW-WDM MSA、PCIe、UCIe 以及 CXL 等协议对物理层和软件层的标准化将助力实现不同硬件和软件框架之间的无缝集成。随着这些关键标准的发展,OIO 将提供高级工作负载所需的可扩展的高效互连,从而有望成为下一代生成式人工智能计算架构的重要推动者。

OIO 比 CPO 的性能要求更高,产业化标准化进度更为滞后。同样能效情况下,OIO 的边带宽密度比 CPO 的大一个数量级;相同误码率情况下,OIO 的延迟是 CPO 的 1/30~1/20。相对于 CPO,OIO 的产品样机更少,标准研制的启动时间更晚,且主要基于计算领域标准向光学方向拓展,尚未建立自己的标准体系框架。

### 3 结束语

高带宽、低延迟和高能效的要求驱动了互连技术的进步,片间光互连已取得初步进展。本文对数据中心以及计算中心场景中的片间光互连技术、产业和标准等方面进行了梳理与分析。片间光互连仍有许多关键问题亟待解决,需要技术攻关协同产业标准发展。技术层面,需要进一步完善设计工具软件,推动先进封装与集成技术发展。产业层面,行业巨头应进一步发挥强牵引作用,带动片间光互连在数据中心、高性能计算、人工智能和航空航天等多领

域的应用。同时标准研制应同步推进,积极完善 CPO 测试方法标准,推动提高 PCIe、UCIe 和 CXL 等规范与光互连的适配性。

#### 参考文献:

- [1] 张新全,余少华. 光电融合破解带宽、能耗难题[J]. 光通信研究,2021(5):1—14.  
Zhang X Q, Yü S H. Address the Challenges of Bandwidth and Power Consumption through Photonics-electronics Convergence[J]. Study on Optical Communications, 2021(5):1—14.
- [2] 刘璐,吴冰冰. 面向下一代数据中心的光电合封技术分析[J]. 通信世界,2023(10):45—46.  
Liu L, Wu B B. Analysis of Co-packaged Optics Technology for Next Generation Data Center[J]. Communications World,2023(10):45—46.
- [3] 刘璐,吴冰冰. 片上与片间光互连技术与产业分析[J]. 通信世界,2023(18):43—44.  
Liu L, Wu B B. On-chip and Inter-chip Optical Interconnection Technology and Industry Analysis [J]. Communications World,2023(18):43—44.
- [4] 杜特,马汉斯,姜鑫鹏,等. 片上光互连器件的智能化设计研究进展[J]. 物理学报,2023,72(18):182—210.  
Du T, Ma H S, Jiang X P, et al. Research Progress of Intelligent Design of On-chip Optical Interconnection Devices[J]. Acta Physica Sinica,2023,72(18):182—210.
- [5] Maniotis P, Schares L, Kuchta D M. How Data Center Networks can Improve Through Co-packaged Optics[C]//2023 Optical Fiber Communications Conference and Exhibition (OFC). San Diego, CA, USA: IEEE,2023:Th1D. 3.
- [6] OIF- Co-Packaging-FD-01. 0-2022, Co-Packaging Framework Document[S].
- [7] Maniotis P, Kuchta D M. Exploring the Benefits of Using Co-packaged Optics in Data Center and AI Supercomputer Networks: a Simulation-based Analysis[J]. Journal of Optical Communications and Networking, 2024,16(2):A143—A156.
- [8] Tan M, Xu J, Liu S Y, et al. Co-packaged Optics (CPO):Status, Challenges, and Solutions[J]. Frontiers of Optoelectronics,2023,16(1):1.
- [9] Muth K,Raghuraman V. Key Technology Enablers for Co-packaged Optics[J]. Optical Interconnects XXIII, 2023,PC12427:257434675.
- [10] Uemura H, Matsui N, Motoji R, et al. A Silicon-photonics Optical Transmitter for 12-port 1.6 Tbps Co-packaged Optics Modules [J]. Optical Interconnects XXIII, 12427,2023:258112671.
- [11] Kuchta D M. Developments of VCSEL-based Transceivers for Co-packaging[C]//2023 Optical Fiber Communications Conference and Exhibition (OFC). San Diego,CA,USA:IEEE,2023:M4E. 6.
- [12] Hosseini K,Kok E,Shumarayev S Y,et al. 5.12 Tbps Co-packaged FPGA and Silicon Photonics Interconnect I/O[C]//2022 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits). Honolulu, HI, USA:IEEE, 2022:9830221.
- [13] Xuan Z, Balamurugan G, Huang D, et al. A 256 Gbps Heterogeneously Integrated Silicon Photonic Microring-based DWDM Receiver Suitable for In-package Optical I/O[C]//2023 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits). Kyoto, Japan:IEEE, 2023: 10185280.
- [14] CW-WDM MSA Promoter Members. , CW-WDM MSA Technical Specifications Rev 1.0 [DB/OL]. (2021-07-04) [2024-02-06]. <https://cw-wdm.org/?wpdmdl=2092>.