

通信

证券研究报告

2024年02月24日

AI时代算力需求激涌，高速率光模块伴潮而行

投资评级

行业评级 强于大市(维持评级)

上次评级 强于大市

电信市场运营商资本开支温和增长，投入向算力网络倾斜。

2019-2023年三大运营商 CAPEX 年均复合增速为 4.61%，呈温和增长态势。2023 年运营商在算力网络的投入支出明显加快，2023 年中国移动预计算力网络资本开支 452 亿元；中国电信预计 2023 年算力投资实现 195 亿元；中国联通预计 2023 年算力网络资本开支将达到 149 亿元。

数通市场发展未来增速显著，带动光模块市场景气度。

数通市场需求扩增的动力来自两大方面：1) 2019-2022 北美 4 大云厂商资本开支增速高达 30.77%，数据中心建设投入持续加快。2) 叶脊网络架构大幅增加对光模块数量的需求，并助力光模块往更高速率方向迭代升级。

Chat GPT 打响 AI 军备竞赛，算力需求驱动光模块更新迭代进程提速。

模型升级迭代的背后是对海量数据的训练和推理，自 2012 年以来全球算力需求迎来快速增长，但 AI 的应用和发展存在“木桶效应”，即算力服务器集群之间的数据交换需要数据中心内部的网络互联做配套，从而激发出对高速率光模块的大量需求，推动光模块行业发展。

2024 年 800G 将步入规模化放量阶段，1.6T 时代逐步临近

受海外 AI 算力需求驱动，国内光模块龙头厂商 23 年 3 月 800G 需求起量，23 年下半年 400G 订单出货增加；随着交换芯片容量约每 2 年翻番，FiberMall 预计 1.6T 光模块需求有望于 2025 年实现。

LPO、CPO、硅光和薄膜铌酸锂方案成为光模块新技术发展方向

算力时代背景下，数据中心成为能耗大户，光模块技术的升级不仅仅是简单的速率翻倍，更需要解决速率提高所带来的功耗高、成本大等问题。LPO、CPO、硅光和薄膜铌酸锂方案有望成为解决该系列问题的突破口。

投资建议：1) 光模块及光器件：中际旭创、天孚通信、新易盛；2) 光芯片及硅光：源杰科技（与电子团队联合覆盖）、博创科技、仕佳光子；3) 交换机：中兴通讯、紫光股份（与计算机团队联合覆盖）、沪电股份（与电子团队联合覆盖）；4) 算力调度：润建股份、润泽科技（与机械团队联合覆盖）。

风险提示：1) AI 应用发展不及预期风险；2) 海外大厂投资不及预期风险；3) 中美摩擦升级风险；4) 行业竞争加剧，价格和盈利能力下降风险。

重点标的推荐

股票代码	股票名称	收盘价 2024-02-23	投资 评级	EPS(元)				P/E		
				2022A	2023E	2024E	2025E	2022A	2023E	2024E
300308.SZ	中际旭创	147.73	买入	1.52	2.62	5.65	7.55	97.19	56.39	26.15
300394.SZ	天孚通信	119.42	增持	1.02	1.76	3.16	4.19	24.61	67.85	37.79
300502.SZ	新易盛	59.96	买入	1.27	1.15	2.31	3.33	47.21	52.14	25.96
688498.SH	源杰科技	133.92	增持	1.18	0.26	1.48	2.40	113.49	515.08	90.49
300548.SZ	博创科技	25.03	增持	0.68	0.86	1.03	1.22	36.81	29.10	24.30
688313.SH	仕佳光子	9.88	增持	0.14	0.02	0.15	0.22	70.57	494.00	65.87
000063.SZ	中兴通讯	25.52	买入	1.69	2.10	2.63	3.09	15.10	12.15	9.70
000938.SZ	紫光股份	18.28	增持	0.75	0.76	0.95	1.16	24.37	24.05	19.24
002463.SZ	沪电股份	25.32	增持	0.71	0.79	1.15	1.44	35.66	32.05	22.02
002929.SZ	润建股份	34.10	买入	1.53	1.99	2.86	3.73	22.29	17.14	11.92
300442.SZ	润泽科技	25.64	买入	0.70	0.98	1.23	1.85	36.63	26.16	20.85

作者

唐海清 分析师
SAC 执业证书编号：S1110517030002
tanghaiqing@tfzq.com
王奕红 分析师
SAC 执业证书编号：S1110517090004
wangyihong@tfzq.com

行业走势图



资料来源：聚源数据

相关报告

- 《通信-行业研究周报:谷歌发布 Gemini Pro1.5 模型, OpenAI 视频生成“Sora”重磅登场》 2024-02-19
- 《通信-行业研究周报:海外云业务快速发展, AI 成为海外巨头重要投入方向》 2024-02-04
- 《通信-行业研究周报:通信行业高质量发展稳步推进, Open AI 大模型持续巩固优势》 2024-01-28

资料来源：Wind，天风证券研究所，注：PE=收盘价/EPS

内容目录

1. 光模块简介	5
1.1. 光模块的定义	5
1.2. 光模块的成本架构	5
1.3. 光模块产业链上下游	5
1.4. 光模块行业格局	8
2. 光模块行业概况	8
2.1. 电信市场运营商资本开支稳中见长，开支比重向算力网络倾斜	9
2.2. 资本开支高增、网络架构升级成为提升数通市场景气度的核心驱动力	10
2.3. AIGC 推动算力需求，“木桶效应”加快高速率光模块迭代升级	12
2.4. 2024 年 800G 将步入规模化放量阶段，1.6T 时代逐步临近	14
3. 封装与材料的革新成为光模块的未来发展方向	16
3.1. LPO	16
3.2. CPO	18
3.3. 硅光子技术	19
3.4. 薄膜铌酸锂	21
4. 重点关注标的	22
5. 风险提示	24

图表目录

图 1：光模块基本结构	5
图 2：光模块的工作原理	5
图 3：光模块成本占比	5
图 4：光芯片在不同速率光模块中的占比	5
图 5：光模块产业链全景图	6
图 6：光有源芯片分类概况	6
图 7：2023 年光芯片国产化率情况	6
图 8：各类光芯片竞争情况与主要供应商	7
图 9：电芯片分类概况	8
图 10：2017 年电芯片国产化率预测	8
图 11：2022 年 Top10 光模块厂商中国内厂商占 7 家	8
图 12：2015-2023 年光模块月度海关出口情况	8
图 13：2022-2028 年全球光模块市场 CAGR 将高达 12%	8
图 14：2024-2029 年中国光模块部署量占比全球 20%-25%	8
图 15：2017-2023 年全国移动通信基站数量	9
图 16：光模块在 5G 网络架构中的应用	9
图 17：2022 年三大运营商算力网络建设概况	9
图 18：2023 年全球云服务提供商资本开支占比	10
图 19：2020-2023Q3 年北美 4 大云厂商资本开支规模及增速	10

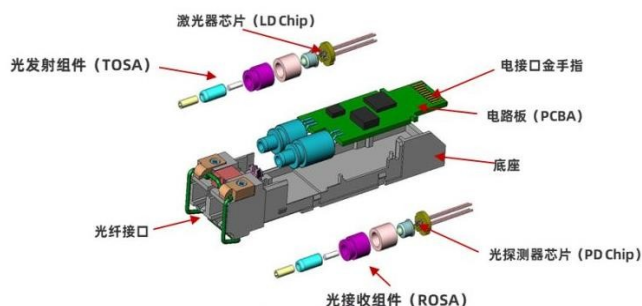
图 20 : 2022-2032 全球超大规模数据中心市场规模预测	10
图 21 : 2015-2021 全球超大规模数据中心数量	10
图 22 : 东西流量占比数据中心总流量的 85%	11
图 23 : 叶脊架构与传统三层架构的区别	11
图 24 : 2023 年国内外 AI 公司相继发布模型	12
图 25 : AlexNet 至 AlphaGo Zero6 年时间内算力需求翻 30 万倍	13
图 26 : 全球算力需求增速超过摩尔定律	13
图 27 : 不同组网方式下 GPU 与光模块用量配比	13
图 28 : 计算网络环境中 NVIDIA DGX H100 集群的网络架构	13
图 29 : 以太网交换芯片容量与不同速率光模块匹配时间线	14
图 30 : 800G 光模块在传统三层拓扑架构中的场景应用	14
图 31 : 800G 光模块在叶脊架构中的场景应用	14
图 32 : 三种 800G 光模块光接口架构	15
图 33 : 商用交换芯片容量大约每 2 年翻 1 倍	15
图 34 : 交换机密度每 2 年翻 1 倍	15
图 35 : 2010-2022 年光引擎功耗提升 26 倍	16
图 36 : 400G、800G、1.6T 光模块功耗对比	16
图 37 : LPO 采用线性直驱技术代替传统 DSP	17
图 38 : DSP 在 400G 光模块中的功耗占比约 49%	17
图 39 : LPO 技术优势包括低功耗、低延迟、低成本和可热拔插	17
图 40 : 国内光模块厂商 LPO 进展概况	17
图 41 : 传统可热拔插与 CPO 的区别	18
图 42 : 2022-2033 年数通市场 CPO 收入增长预期	18
图 43 : 2023 年 CPO 产业链企业分布情况	18
图 44 : 硅基光电子集成芯片概念图	19
图 45 : 光纤集成的演变发展	19
图 46 : 2022-2028 年硅基光电子芯片收入增长预测 (按应用划分)	19
图 47 : 2023 年硅光子产业链	20
图 48 : 2022 年电信、数通市场份额	20
图 49 : 国内上市公司硅光技术进展概况	20
图 50 : 三类电光调制器材料方案对比	21
图 51 : 调制器性能对比	21
图 52 : 薄膜铌酸锂调制器比传统铌酸锂尺寸减少 60%	22
图 53 : 传统铌酸锂与薄膜铌酸锂调制器各项指标对比	22
图 54 : 2022-2030 年全球铌酸锂调制器市场规模预测 (\$百万)	22
图 55 : 2019-2023Q3 中际旭创营收、归母净利润	23
图 56 : 2018-2023Q3 中际旭创毛利率、各项费用率	23
图 57 : 2018-2023Q3 天孚通信营收、归母净利润	23
图 58 : 2018-2023Q3 天孚通信毛利率、各项费用率	23
图 59 : 2018-2023Q3 新易盛营收、归母净利润	24
图 60 : 2018-2023Q3 新易盛毛利率、各项费用率	24

1. 光模块简介

1.1. 光模块的定义

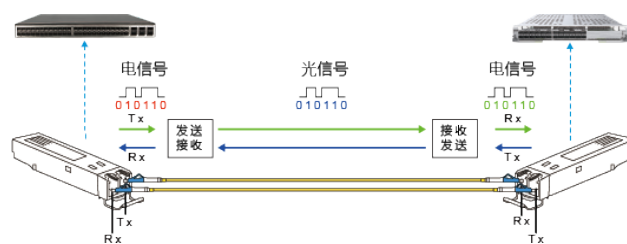
光模块是一种用于光通信的设备，它包括发光器和接收器，可以将电信号转换为光信号并发送到远程设备，也可以将收到的光信号转换为电信号并传递给计算机、路由器或交换机等设备。光模块通常用于数据中心、企业网络、电信运营商网络等领域，用于实现高速、远距离、高可靠性的通信。

图 1：光模块基本结构



资料来源：H3C 官网，天风证券研究所

图 2：光模块的工作原理



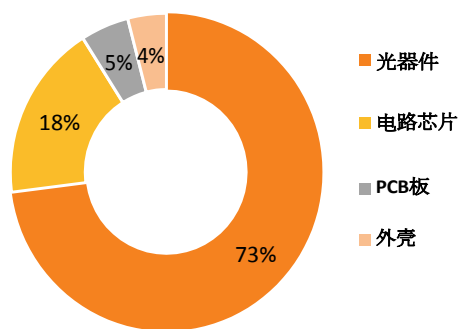
资料来源：华为官网，天风证券研究所

1.2. 光模块的成本架构

光模块中实现光电转换作用的核心部件是光电芯片，光芯片是光模块中完成光电信号转换的直接芯片，而电芯片是实现对光芯片工作的配套支撑，两者都是光模块的核心部件。从成本占比来看，光芯片通常占光模块成本的 40%-60%，电芯片占 10%-30% 之间。

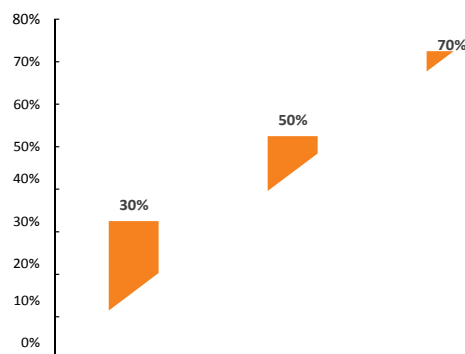
光模块的主要升级在速率，光通信芯片的成本随着光模块速率的不断升高而提高。作为最主要的成本构成，芯片的差异也成为了衡量光器件高低端的主要标准。越高速、越高端的光模块，光芯片和电芯片的成本占比就越高。

图 3：光模块成本占比



资料来源：易天光通信官网，天风证券研究所

图 4：光芯片在不同速率光模块中的占比

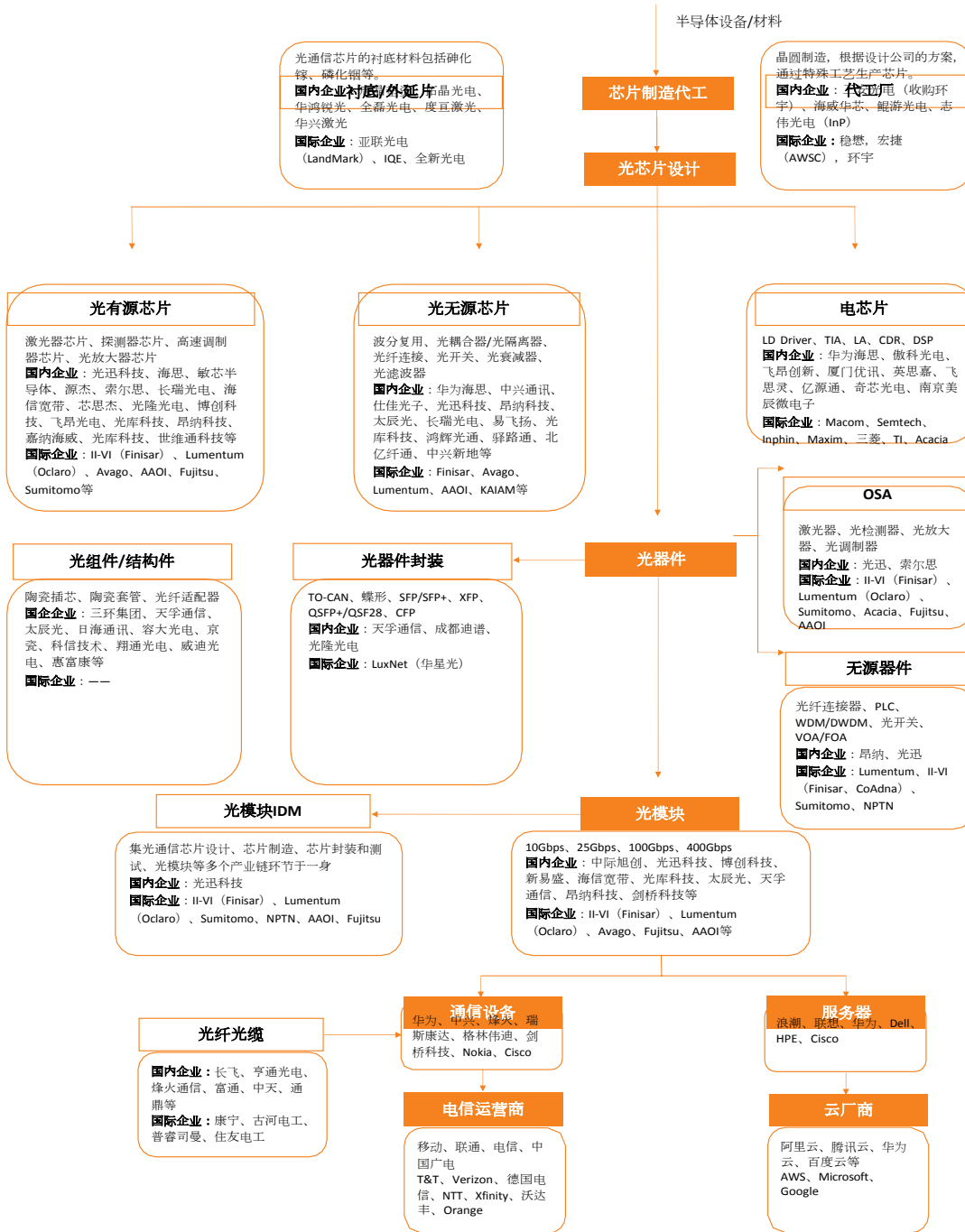


资料来源：易天光通信官网，天风证券研究所

1.3. 光模块产业链上下游

上游主要包括光芯片、电芯片、光器件等供应商，其中光器件供应商较多，国产化率较高，但芯片工艺技术壁垒高，研发成本大，国外大厂占据高端光芯片、电芯片领域市场大部分份额。光模块身处中游，属于技术壁垒相对较低的封装环节。下游包括互联网及云计算企业、电信运营商、数据通信和光通信设备商等。其中互联网及云计算企业、电信运营商为光模块最终用户。

图 5：光模块产业链全景图



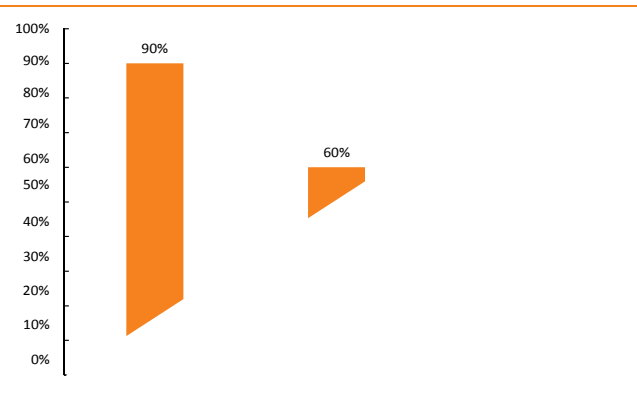
资料来源：创道硬科技研究院，天风证券研究所

光芯片国内厂商竞争领域集中于 25Gb/s 以下产品，25Gb/s 及以上产品国产化率尚有不足。按速率划分，光芯片一般分为 2.5Gb/s、10Gb/s、25Gb/s 及以上各种调制速率。按功能划分，光芯片可分为激光器芯片和探测器芯片，激光器芯片用于发射信号，将电信号转化为光信号，按出光结构进一步分为面发射芯片和边发射芯片，主要包括 VCSEL、FP、DFB、EML；探测器芯片用于接收信号，将光信号转化为电信号，主要包括 PIN 和 APD。2022 年 EML 大规模商用的最高速率已达到 100Gb/s，DFB 和 VCSEL 激光器芯片大规模商用的最高速率已达到 50Gb/s。

图 6：光有源芯片分类概况

图 7：2023 年光芯片国产化率情况

芯片类型	产品类型	工作波长	传输速率	传输距离	产品特性	应用领域	材
激光器芯片	VCSEL (面发射)	800-900nm	155M-25G	短距500m	线宽窄、功耗低、500m以内的短距离传输, 如调制速率高、耦合效率高、传输距离短、线性度差	数据中心机柜内部传输、电子领域 (3D感测)	
	FP (边发射)	1310-1550nm	155M-10G	中距40km	调制速率高、成本低、耦合效率低、线性度差	主要应用	
	DFB (边发射)	1270-1610nm	2.5G-40G	长距80km	谱		
	EML (边发射)	1270-1610nm	1				
探测器芯片	PIN						



资料来源：源杰科技招股说明书、创道硬科技研究院，天风证券研究所

资料来源：ICC，中商产业研究院，天风证券研究所

欧美日国家光芯片厂商具有技术经验先发优势，逐步实现产业闭环，并建立起较高的行业壁垒，拥有可量产 25Gb/s 速率以上光芯片的技术。国内厂商在芯片制造中对外延技术的掌握尚未成熟，因此高端外延片主要依赖进口，导致发展受限。分速率市场来看，国内厂商目前能够规模量产 2.5Gb/s、10Gb/s 激光器芯片，25Gb/s 激光器芯片仅少部分厂商实现批量发货，50Gb/s、硅光方案大部分厂商仍处于验证试产阶段。

图 8：各类光芯片竞争情况与主要供应商

产品速率	产品类型	竞争情况	主要供应商
2.5G	1310nm DFB激光器芯片	应用于光纤接入PON (GPON) 数据上传光模块，技术相对成熟，市场竞争较为激烈	源杰科技、武汉敏芯、三安光电、中科光芯、雷光科技、光安伦
	1490nm DFB激光器芯片	应用于光纤接入PON (GPON) 数据下传光模块，产品性能、可靠性要求高，实现批量供货厂商较少，公司等国内厂商市场份额较大，国产化率较高	三菱电机、源杰科技、海信宽带
	1270nm DFB激光器芯片	应用于光纤接入10G-PON (GX-PON) 数据上传光模块，产品难度较2.5G 1310nm DFB激光器芯片更高，但供应商逐步增多，市场竞争逐步加剧	三菱电机、源杰科技、武汉敏芯、海信宽带、光迅科技
	1550nm DFB激光器芯片	应用于40km/80km长距离传输光模块，产品性能、可靠性要求高，实现批量供货厂较少	三菱电机、源杰科技、海信宽带、光迅科技
10G	1270nm DFB激光器芯片	应用于光纤接入10G-PON (XGS-PON) 数据上传光模块，产品性能、可靠性要求高，实现批量供货厂商较少，公司等国内厂商市场份额较集中	三菱电机、MACOM、源杰科技、武汉敏芯、海信宽带
	1310 nm FP激光器芯片	应用于4G移动通信网络光模块，技术相对成熟，市场竞争较为激烈	三菱电机、源杰科技、云岭光电、武汉敏芯、海信宽带
	1310nm DFB激光器芯片		
25G	CWDM 6 波段DFB激光器芯片	应用于4G/5G移动通信网络光模块，技术相对成熟，国内厂商逐渐扩大市场份额	MACOM、Lumentum、源杰科技、武汉敏芯
	LWDM 12 波段DFB激光器芯片	站方案，国外厂商发货产品较少，该产品公司等国内光芯片厂商在2020年市场大批量发货	科技、武汉敏芯
	MWDM 12 波段DFB激光器芯片	应用于100G数据中心光模块，产品难度大，国内部分厂商实现产品突破	Avago、MACOM、Lumentum、源杰科技、武汉敏芯
	CWDM 4 波段DFB激光器芯片		
50G 硅光直流光源	PAM4 CWDM 4 波段DFB激光器芯片	应用于100G/200G/400G数据中心光模块，技术难度高，国内部分厂商进行产品布局，尚未实现批量发货；公司50G PAM4 DFB激光器处于设计验证测试阶段，工业级大功率硅光激光器处于工程验证测试阶段	Avago、Lumentum
	1270/1290/1310/1330nm 大功率 25/50/70mW激光器芯片		

资料来源：中商情报网，天风证券研究所

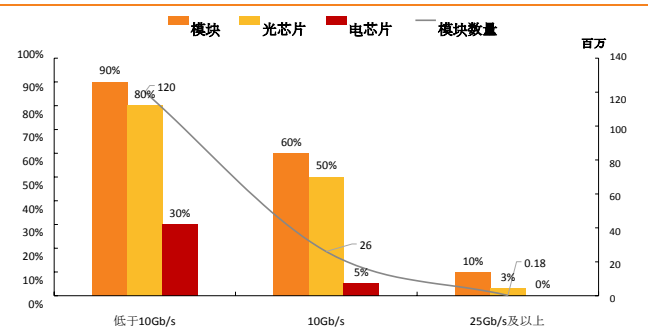
电芯片国产化率偏低，核心供应厂商仍以海外企业为主。电芯片包括激光驱动芯片 LD Driver、跨阻放大器芯片 TIA、限幅放大芯片 LA、DSP 电信号处理芯片等，但国内只有少数供应商涉足 25Gb/s 及以下速率的电芯片产品，25Gb/s 以上基本依赖进口。

图 9：电芯片分类概况

芯片类型	典型产品	功能描述
电芯片	激光驱动芯片 (LD Driver)	应用于直调激光器, LDD 驱动激光器发射强度不同的光信号, 构成数字 0101010...
	跨阻放大器芯片 (TIA)	置于光接收探测器之后, 将探测器接收并转换成的电流中能够处理的电压信号。
	限幅放大芯片 (LA)	将探测器接收并转换成的电流信号, 转换为电压信号。
	时钟数据恢复芯片 (CDR)	从接收的数据信号中提取时钟
	高速调制器驱动芯片	驱动高速调制
	DSP 电信号处理芯片 串行/解串 SerDes	
	温度	

资料来源：创道硬科技研究院，天风证券研究所

图 10：2017 年电芯片国产化率预测



资料来源：中国光电子器件产业技术发展路线图（2018-2022 年），天风证券研究所

1.1. 光模块行业格局

近年来国内光模块厂商高速发展，行业内呈现“西退东进”趋势。2010 年至今以来国内光模块厂商飞速发展，全球排名逐步靠前，2010 年国内仅有一家武汉电信器件有限公司（WTD）入围全球 top10 榜单，LightCounting 公布的 2022 年全球 Top10 榜单中国内厂商强者愈强，共计 7 家入围，其中中际旭创与 Coherent 并列高居榜首。

2015-2023 年光模块月度出口规模震荡上行，2023 年 11 月、12 月连续创下新高。根据海关出口数据，2015-2023 年我国光模块出口年化复合增速为 12.26%，上升趋势明显。2023 年呈现前低后高的走势，主要原因为国内头部厂商高速率产品海外订单 Q2、Q3 逐步起量，拉动出口需求增长。

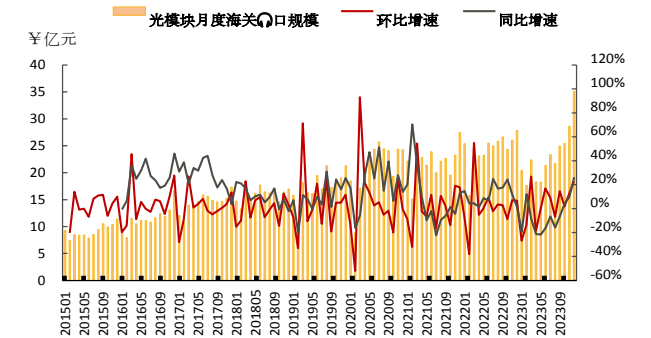
图 11：2022 年 Top10 光模块厂商中国内厂商占 7 家

Ranking of Top 10 Transceiver Suppliers				
2010	2016	2018	2022	
Finisar	Finisar	1	Finisar	InnoLight & Coherent (tie)
Opnext	Hisense	2	InnoLight	
Sumitomo	Accelink	3	Hisense	Cisco (Acacia)
Avago	Acacia	4	Accelink	Huawei (HiSilicon)
Source Photonics	FOIT (Avago)	5	FOIT (Avago)	Accelink
Fujitsu	Oclaro	6	Lumentum/Oclaro	Hisense
JDSU	InnoLight	7	Acacia	Eoptolink
Emcore	Sumitomo	8	Intel	HGG
WTD	Lumentum	9	AOi	Intel
NeoPhotonics	Source Photonics	10	Sumitomo	Source Photonics

Source: LightCounting

资料来源：中国电子元件行业协会官网、LightCounting，天风证券研究所

图 12：2015-2023 年光模块月度海关出口情况



资料来源：海关统计数据在线查询平台，天风证券研究所

2. 光模块行业概况

2022-2028 年全球光模块市场规模 CAGR 将实现 12%，2028 年全球规模有望突破 200 亿美元。根据 Yole Intelligence 数据，2022 年全球光模块市场规模为 110 亿美元，在大型云服务运营商对 800G 高数据速率模块的高需求和国家电信对增加光纤网络容量的要求推动下，预计 2028 年将增至 223 亿美元，2022-2028 年化复合增速将达到 12%。

光模块国内市场需求规模为 20-30 亿美元，占比全球约 1/4。根据 LightCounting 预测，2018-2023 年中国光模块部署量占比全球 25%-35%，2024-2029 年占比 20%-25%，略有下降。北美云服务商在 AI 集群中部署 800G 的计划较为激进，这将成为未来 2-3 年中国份额下降的主要因素。

图 13：2022-2028 年全球光模块市场 CAGR 将高达 12%

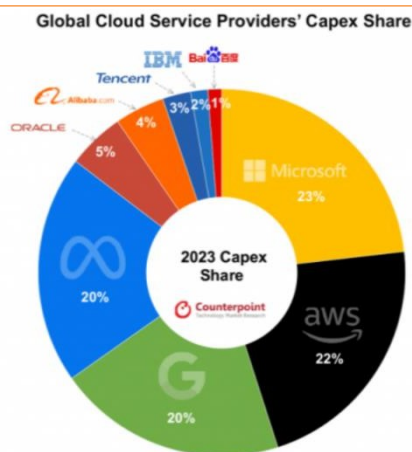
图 14：2024-2029 年中国光模块部署量占比全球 20%-25%

资料来源：北京通信信息协会官网，天风证券研究所

1.4. 资本开支高增、网络架构升级成为提升数通市场景气度的核心驱动力

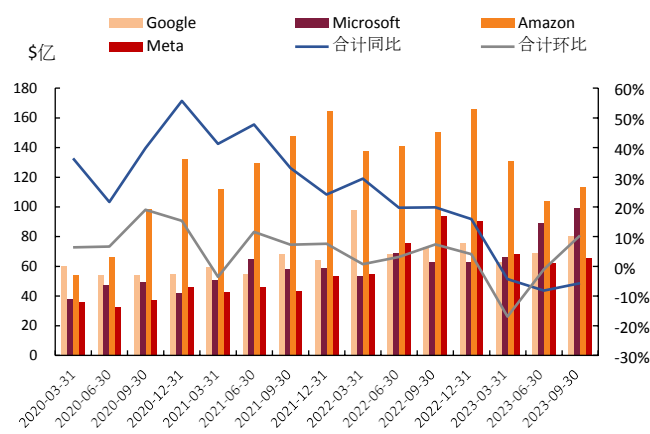
北美 4 大云服务商资本开支占比全球约 85% 份额，2019-2022 年 CAPEX 年化复合增速高达 30.77%。根据 Counterpoint Research 数据预测，2023 年北美 4 大云服务商 Microsoft、Amazon、Google 和 Meta 将占全球总资本支出的 85%。其中 2022 年整体资本开支有所放缓，背后或可归结于疫情期间的全球经济下行，但自 23Q2 开始逐步回暖。

图 18：2023 年全球云服务提供商资本开支占比



资料来源：Counterpoint Research，天风证券研究所

图 19：2020-2023Q3 年北美 4 大云厂商资本开支规模及增速



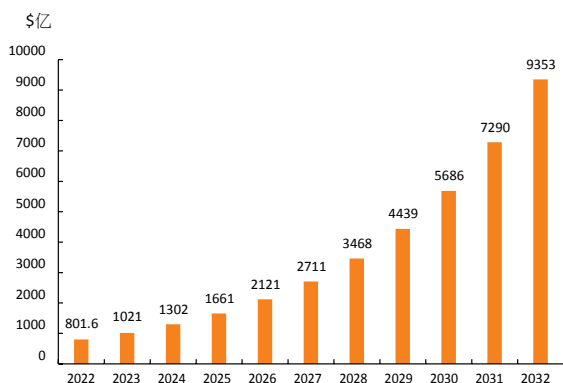
资料来源：Wind，天风证券研究所

全球数据规模扩张使得对数据中心的建设投入加快，数据中心内部设备需要网络互联，从而对应产生光模块的大量需求。随着通信、互联网应用的不断发展，个人、企业对算力和数据储存的需求逐步往“云”上迁移。IDC 最新数据显示全球数据规模将从 2022 年的 103.66ZB，提升至 2027 年的 284.3ZB，2022-2027 年化复合增速将达到 22.36%，我们认为数据规模的增加需要更多的服务器、交换机，这些设备需要信息交换和网络互联，也就对应产生光模块的大量需求。

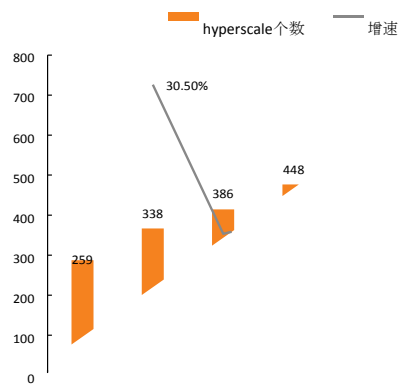
Statista 测算全球 hyperscale 数量已经从 2015 年的 259 个，提升至 2021 年的 700 个，且 Synergy Research Group 最新数据显示 2023 年该数量已接近 900 个，容量占比全球所有数据中心的 37%。按照 Precedence Research 数据来看，2022 年全球 hyperscale 市场规模为 801.6 亿美元，预计 2032 年将提升至 9353 亿美元，2023-2032 年年化复合增速将达 27.9%。

图 20：2022-2032 全球超大规模数据中心市场规模预测

图 21：2015-2021 全球超大规模数据中心数量



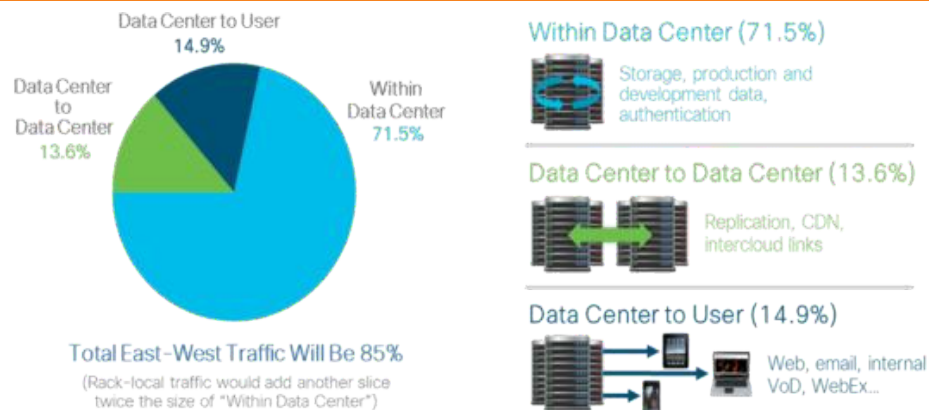
资料来源：Precedence Research，天风证券研究所



资料来源：Statista，天风证券研究所

叶脊网络架构大幅增加对光模块数量的需求，助力光模块往更高速率方向迭代升级。数据中心的流量分为两种：1) 南北流量：数据中心外的客户端与数据中心内的服务器之间通信的流量；2) 东西流量：数据中心内不同服务器之间通信的流量。Cisco2021 年数据显示，数据中心流量中以东西流量为主，其占比高达 85%。

图 22：东西流量占比数据中心总流量的 85%

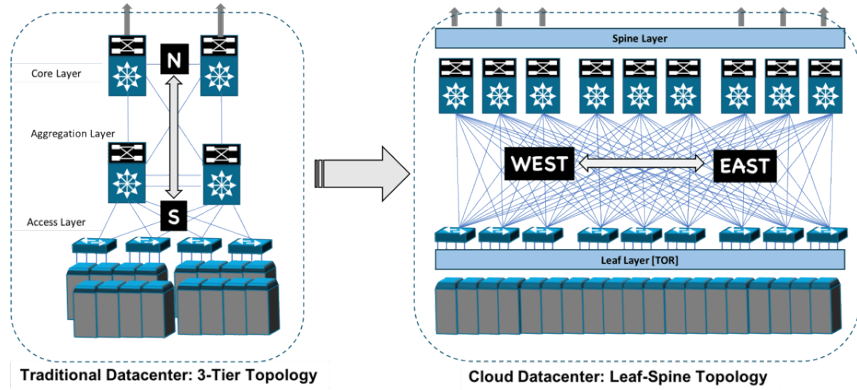


Source: Cisco Global Cloud Index, 2016-2021

资料来源：Cisco，天风证券研究所

传统三层拓扑网络架构中的东西向流量必须经过汇聚层和核心层的设备转发，经过了许多非必要节点，导致最终用户获得的响应时间变慢。相比而言，叶脊架构 (Spine-Leaf) 的扁平化设计能将东西流量分布在多条路径上，更适于在东西流量较大的场景中使用。进一步讨论，脊交换机和叶交换机之间的互联需要合理带宽比例进行匹配，其上行链路应始终比下行链路运行更快，以免端口链路阻塞，因此叶脊架构在部署更多光模块数量的同时，也对光模块速率提出更高要求。

图 23：叶脊架构与传统三层架构的区别



资料来源：Semiconductor Engineering，天风证券研究所

1.5. AIGC 推动算力需求，“木桶效应”加快高速率光模块迭代升级

2022 年 11 月 30 日 OpenAI 公司旗下模型 ChatGPT 的发布打响全球型军备竞赛。2022 年 11 月 30 日 OpenAI 公司发布大语言模型 ChatGPT，公司 CEO Sam Altman 宣布 ChatGPT 用户仅用时 5 天已经突破 100 万人，且在 2023 年 1 月末，ChatGPT 用户数量突破 1 亿，成为史上用户增长数最快的消费者应用。ChatGPT 的兴起引领了全球多家 AI 公司相继发布自家模型，海外 Google 推出 PaLM2 模型，Meta 发布 LLaMA-13B，Microsoft 基于 ChatGPT 打造 New Bing；国内百度率先发布文心一言。

图 24：2023 年国内外 AI 公司相继发布模型

时间	公司	模型	概述
2023年2月24日	Meta	LLaMA	LLaMA模型旨在帮助研究人员推进工作，在生成文本、对话、总结书面材料、证明数学定理或预测蛋白质结构等更复杂的任务方面有很大的前景。LLaMA目前包含70亿、130亿、330亿和650亿这4种参数规模的模型。其中，参数规模最小的LLaMA 7B也经过了超1万亿个tokens的训练。Meta表示，在大多数基准测试中，参数仅为十分之一的LLaMA-13B的性能优于OpenAI推出的GPT3(175B)，也即支持ChatGPT的GPT3.5的前身。
2023年3月16日	百度	文心一言	文心一言有五大能力：文学创作、商业文案创作、数理逻辑推算、中文理解、多模态生成。
2023年3月14日	OpenAI	GPT-4	相比于上一代GPT-3，GPT-4的优势包括1) 接受图像和文本输入，生成文本；2) 更准确地解决难题，具有更广泛的常识和解决问题的能力；3) 更具创造性和协作性；4) 能够处理超过25000个单词的文本，允许长文内容创建、扩展对话以及文档搜索和分析等用例。
2023年5月6日	Microsoft	New Bing	多模态内容和文成图；历史记录功能；开放插件接入；
2023年5月10日	Google	PaLM 2 (Pathways Language 2)	PaLM 2模型提供了不同规模的四个版本，其中轻量级的Gecko模型可以在移动设备上运行，速度非常快，不联网也能在设备上运行。谷歌还推出了两个专业领域大模型，其中，Med-PaLM 2能回答各种医学问题，是首个在美国医疗执照考试中达到专家水平的大语言模型。
2023年11月7日	OpenAI	GPT-4 Turbo	具有 128k 上下文的 Chat GPT-4 模型；开放新的 API：DALL-E 3、Whisper V3 和 GPT-4 Turbo 等；自定义版本的 ChatGPT；GPT Store 即将上线
2023年11月29日	Pika Labs	Pika1.0	Pika 1.0能够生成和编辑3D动画、动漫、卡通和电影。此外，用户还能够通过Pika实现画布延展、局部修改、视频时长拓展等编辑需求。
2023年12月1日	Amazon	Titan	三款“Titan”系列生成式 AI 模型，其中包含亚马逊旗下首个图像生成模型“Titan Image Generator”、文字生成模型“Amazon Titan Text Express”及“Titan Text Lite”。其中Titan Image Generator 具备“图片编辑”及“隐藏水印”等功能，允许用户以英语输入提示词句，以生成“专业等级”的图像。
2023年12月6日	Google	Gemini	谷歌的Gemini产品分为三种尺寸：Ultra、Pro 和 Nano。其中，Ultra是适用于高度复杂任务的最高且最有能力的模型；Pro则是跨多种任务进行扩展的最佳模型；而Nano是专为设备上任务设计的最高效模型。

资料来源：央视网、澎湃新闻、凤凰网、中国经济网、IT之家、中关村在线，天风证券研究所

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/166211050052010105>