

通信设备

光模块行业深度

领先大市-A(首次)

2023年6月28日

行业研究/行业专题报告

通信设备板块近一年市场表现



资料来源：最闻

分析师：

高宇洋

执业登记编码：S0760523050002

邮箱：gaoyuyang@sxzq.com

投资要点：

➢ 光模块是光通信产业链的重要组成部分，目前市场规模约百亿美元。光模块充分受益于下游电信与数通市场高速发展，行业增长确定性高。

➢ 从光模块下游市场来看：1) 电信市场是光模块的发源场景，需求保持稳中有增。一方面，电信用光模块主要集中在前传，对光模块速率要求低且产品迭代速率慢，三大运营商计划将资本开支逐渐倾斜于云计算等算力网络场景；另一方面，光纤入户规模扩大与速率向 10G PON 升级助推光模块需求；2) 数通市场是光模块主要市场且发展迅速。数据中心流量增加和网络架构由传统架构转变为叶脊网络架构推动光模块数量和速率快速提高。目前光模块正处于向 800G 速率发展的新周期。

➢ 以 ChatGPT 为首的 AI 军备竞赛开启，大幅拉动 800G 等高速光模块需求量，北美云厂商订单与业绩超预期催化光模块继续放量。英伟达超级计算机 DGX GH200 推出，单 GPU 与 800G 光模块配比量数倍增加；国内外云计算公司陆续推出大语言模型，训练侧模型架构复杂多元，参数量由千亿达到万亿，算力需求大幅增加；推理侧仍在流量增加时期，大模型军备竞赛进入稳态后随 AI 普及推理侧将成为光模块用量的主要阵地。

➢ 光模块速率提高与低功耗需求日益强烈，硅光、CPO 等技术趋势有望成为长期解决方案。硅光芯片采用“光互联”，叠加 CPO 技术，将光引擎与交换芯片共同封装，在速率提高的同时大大缩减功耗；此外，LPO，相干技术及薄膜铌酸锂等技术亦成为光模块优化主要新趋势，光模块科技属性增强，研发壁垒与客户认证壁垒增高，长期趋势进一步向好。

➢ 国内厂商主导光模块市场，新周期下龙头厂商技术与市场全面进展，行业内参与者均各有所长。1) 光模块全球市占前 10 中国占据 7 家，龙头企业在新技术与海外客户均布局完善，其他公司在技术与客户导入上各有侧重。2) 光芯片光在光模块中成本占比过半且主要以外购为主，我国相较海外仍落后，光模块厂商参股及收并购芯片厂商或进行自研芯片，一体化布局拥有更多自主权。

**投资建议：** 建议关注中际旭创，新易盛，天孚通信，华工科技，博创科技，光迅科技，联特科技，光库科技，剑桥科技，铭普光磁。

**风险提示：** 800G 需求不及预期； AI 产业发展不及预期； 公司海外客户合作及技术进步不及预期。



请务必阅读最后一页股票评级说明和免责声明

1



## 目录

1. 光模块：光通信的重要光电子器件.....	6
2. AI 拉动需求激增，国内厂商乘风而上.....	7
2.1 电信市场稳中有增，数通市场前景广阔.....	7
2.1.1 电信市场为光模块发源，运营商资本支出由 5G 倾向算力.....	7
2.1.2 数通市场推动演进，叶脊架构提升需求.....	9
2.2 AI：ChatGPT 开启技术革命，光模块市场星辰大海.....	10
2.2.1 AI 大模型打开算力需求天花板，云厂商业绩超预期催化后续支出.....	10
2.2.2 从英伟达计算机和大模型算力预测：800G 光模块需求量显著提升.....	12
2.3 格局：国内厂商主导市场，随技术迭代各有优势.....	17
2.3.1 产业东移，中国厂商逐渐占领高地.....	17
2.3.2 技术迭代：CPO 和硅光有望成为长期解决方案.....	18
2.3.3 厂商机遇：龙头升级换代，新秀弯道超车.....	23
2.3.4 向上布局芯片，一体化掌握话语权.....	24
3. 重点关注标的.....	26
3.1 从估值端看当前光模块市场.....	26
3.2 主要公司建议关注.....	26
4. 风险提示.....	32

## 图表目录

图 1：光模块示意图.....	6
图 2：光模块结构图.....	6
图 3：光通信产业链.....	6
图 4：全球电信侧光模块市场规模及预测（百万美元）.....	7
图 5：电信光传输网络建设对光收发模块的需求.....	7



图 6: 从左至右分别 (移动、电信、联通) 三大运营商资本支出结构变化 (单位: 亿元; %)	8
图 7: 国内互联网宽带接入用户数: FTTH/0	8
图 8: 按地区/国家划分的 PON 设备市场预测	8
图 9: 全球数据中心光模块市场规模及预测 (百万美元)	9
图 10: 通用数据中心使用光模块	9
图 11: 通用数据中心迭代: 叶脊网络架构	9
图 12: 可插拔光模块向 800G 演进	10
图 13: 5 大云厂商光模块销售情况	10
图 14: 神经网络模型持续扩大	11
图 15: Grace Hopper 芯片架构	11
图 16: 140 节点的 NVIDIA DGX SuperPOD 计算侧架构	13
图 17: 计算侧交换机及线缆数量	13
图 18: 140 节点的 NVIDIA DGX SuperPOD 存储侧架构	13
图 19: 存储侧交换机及线缆数量	13
图 20: 计算侧和存储侧的线缆全部采用 NVIDIA MF1S00-HxxxE	13
图 21: NVIDIA DGX GH200 存储侧架构	14
图 22: NVIDIA DGX GH200 结构	14
图 23: GPT-3 计算参数	14
图 24: NVIDIA A00 TENSOR CORE GPU 规格	15
图 25: A100 采用 FP16 精度训练大模型	15
图 26: FLOP 利用率 (GPT-3 为 21.3%)	15
图 27: NVIDIA H00 GPU 规格	16
图 28: 部分国内外大模型参数量节选	16



图 29: NVIDIA A30 Core GPU 规格.....	16
图 30: A10 技术规格和功能.....	16
图 31: 2021 年 光模块主要厂商市占率.....	18
图 32: 2021 年主要厂商市场份额变动情况.....	18
图 33: 不同速率光模块在以太网交换机中成本占比.....	18
图 34: CPO 封装示意图.....	18
图 35: 可插拔光模块发展历程.....	19
图 36: CPO 市场规模及增速.....	19
图 37: 400G ZR 光模块的功耗分解.....	19
图 38: 无需 DSP 的 交换机功耗可下降 25%左右.....	19
图 39: LPO 功耗可下降 50%左右.....	20
图 40: 传统光模块、LPO 与 CPO 各个维度的比较.....	20
图 41: 硅光模块的历史销售及预测.....	21
图 42: 硅光技术图解.....	21
图 43: 全球 ZR 光模块市场规模（单位：亿元）.....	21
图 44: 标准 400G 以太网接口的 PMD，特征和应用.....	21
图 45: 薄膜铌酸锂调制器与传统铌酸锂调制器对比.....	22
图 46: 高速率模块光芯片市场空间及预测（百万美元）.....	25
图 47: 2019-2024 年中国光芯片占全球光芯片比例预测.....	25
图 48: 光模块厂商参股或收购芯片供应商情况.....	25
图 49: 万得光模块板块 PE-Band.....	26
图 50: 中际旭创营收与归母净利润.....	27
图 51: 中际旭创营收按区域分.....	27



图 52: 中际旭创 PE-Band.....	27
图 53: 新易盛营收与归母净利润.....	28
图 54: 新易盛毛利率较高.....	28
图 55: 新易盛 PE-Band.....	28
图 56: 天孚通信营收与归母净利润.....	29
图 57: 天孚通信营收分产品.....	29
图 58: 天孚通信 PE-Band.....	29
图 59: 华工科技营收与归母净利润.....	30
图 60: 华工科技光电器件产品营收及同比.....	30
图 61: 华工科技 PE-Band.....	30
图 62: 博创科技营收与归母净利润.....	31
图 63: 博创科技营收分市场.....	31
图 64: 博创科技 PE-Band.....	31
表 1: 今年 AI 大模型军备竞赛不完全节选.....	11
表 2: 业绩超预期催化后续需求.....	12
表 3: 光模块全球市占率排名.....	17
表 4: 光模块厂商 800G 光模块进展.....	23
表 5: 主要厂商技术研发进展.....	23
表 6: 建议关注公司.....	31

## 1. 光模块：光通信的重要光电子器件

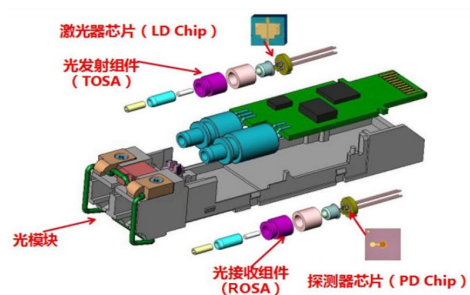
光模块是用于通信设备之间数据传输的光电子器件，主要作用是进行光电和电光转换。从结构上来看，光模块主要由光发射器件（TOSA，含激光器）、光接收器件（ROSA，含光探测器）、功能电路和光（电）接口等部分组成。其中，**光收发组件 TOSA/ROSA 是光模块的核心部分**，TOSA 负责将接收到的电信号转换为光信号，经光纤传送后，通过 ROSA 将光信号转换为电信号，并经前置放大器后输出。根据 OFWEEK 与前瞻产业研究院，光模块中光器件成本占比约 73% 左右，其中 TOSA 占到光器件总成本的 48%，即整个光模块成本的 35%；ROSA 占到了光器件总成本的 32%，即整个光模块成本的 23%。

图 1：光模块示意图



资料来源：中际旭创年报，山西证券研究所

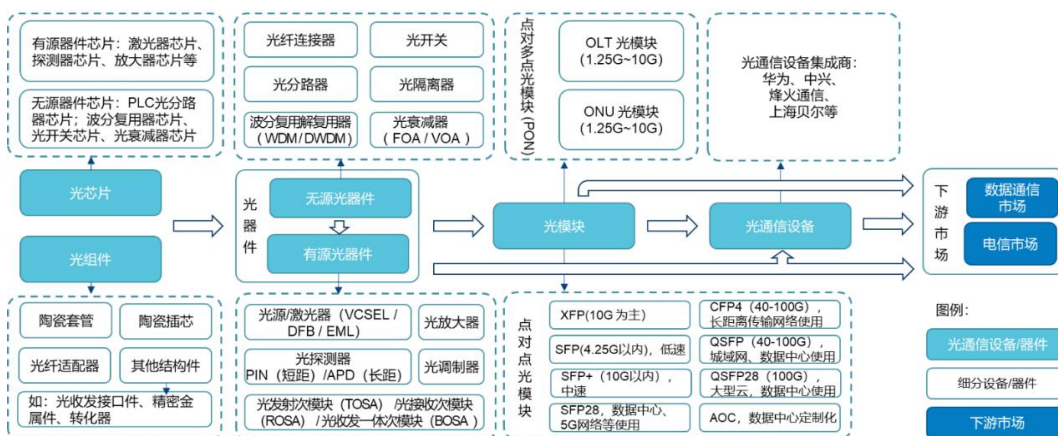
图 2：光模块结构图



资料来源：《5G 承载光模块》白皮书，山西证券研究所

从产业链来看，光模块在光通信产业链中位于中间位置，上游包括光芯片（分为有源和无源）和光组件，两者构成光器件（同样分为有源和无源），光器件与电芯片、PCB 等构成光模块，其中光器件占光模块成本最高，其中光芯片占光模块成本比超过 50%；光模块在产业链后续应用于光通信设备，并在终端应用于以中国移动、中国联通和中国电信三大运营商为主的**电信市场**，以及以谷歌、微软、亚马逊、Meta 四大云服务厂商为主的**数通市场**。

图 3：光通信产业链



资料来源：福睿研究，山西证券研究所

## 2. AI 拉动需求激增，国内厂商乘风而上

本部分我们将阐述当前光模块行业主要现状及趋势，光模块此前主要应用于电信与数通两大市场，其中电信市场份额逐渐减小，数通市场为主要市场；2022年后云厂商资本支出有所下行，但 ChatGPT 的出现引发 AI 科技革命，边际带动数通市场整体向上抬升，AI 成为光模块的主要增长点并有望长期拉动技术层面和需求层面的深刻变革；我国光模块厂商已主导光模块市场，在本轮迭代中各有优势，有望受益于 PE 拔升和业绩兑现迎来戴维斯双击。

### 2.1 电信市场稳中有增，数通市场前景广阔

#### 2.1.1 电信市场为光模块发源，运营商资本支出由 5G 倾向算力

电信市场是光模块的发源地。从电信网络传输需求来看，5G 传输网络由前传、中传和回传组成，分别将蜂窝基站、核心网络和数据中心连接起来，其中前传主要使用 10G、25G 光收发模块，中传主要使用 50G、100G、200G 光收发模块，回传主要使用 100G、200G、400G 光收发模块。前传子系统具有长距离高密度的特点，对光模块需求量最大，因此对光模块需求主要较低速率的光模块为主，根据 LightCounting，全球电信侧光模块市场 2020 年市场规模约 21.66 亿美元，预计到 2025 年将达到 33.54 亿美元，2023-2025 年增速分别 3.4%/11.8%/9.2%。同时根据我们判断的产业发展历史经验，电信市场光模块约 7-10 年速率迭代一次，相较数通市场相对较慢。

图 4：全球电信侧光模块市场规模及预测（百万美元）

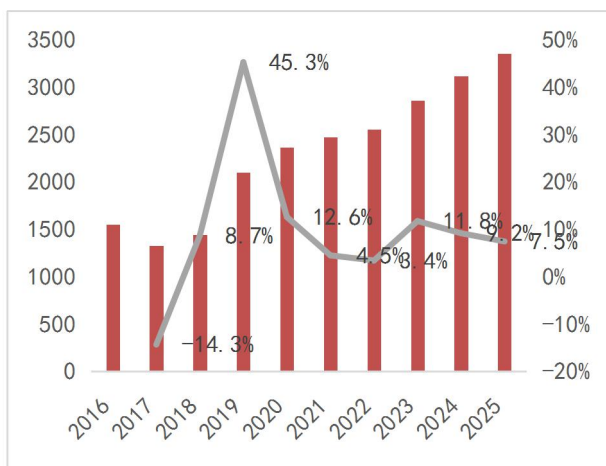
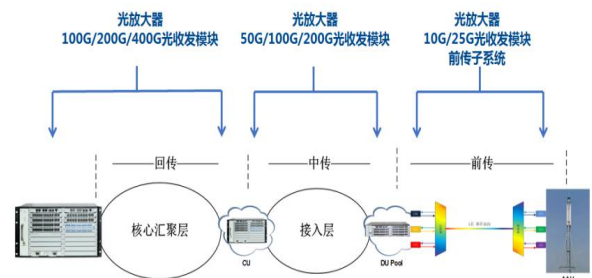


图 5：电信光传输网络建设对光收发模块的需求



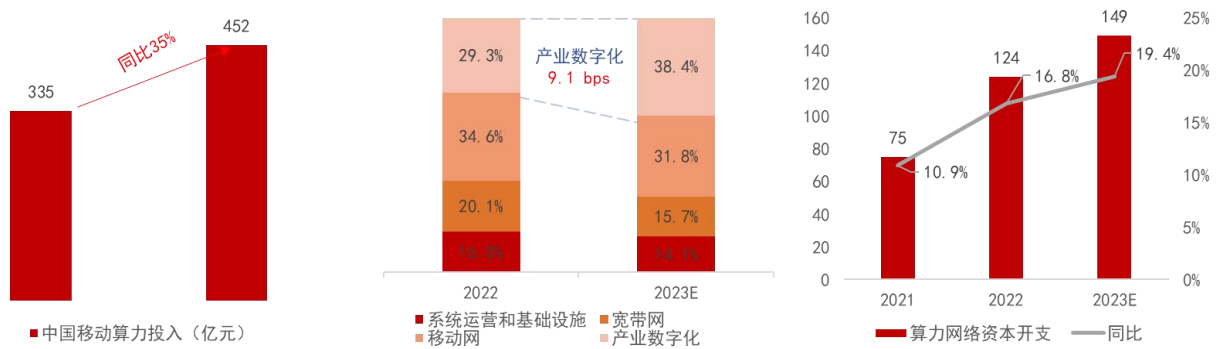
资料来源：源杰科技招股书援引 Lightcounting，山西证券研究所

资料来源：德科立招股书，山西证券研究所

(注：全球电信侧光模块市场规模及预测中不包括 FTTx 市场)

另一方面，我国 5G 投资逐步迈过高峰，增速趋于平稳。从资本支出结构上来看，三大运营商预计未来资本支出逐渐向云计算等算力方向倾斜：中国移动对 2023 年算力网络的资本开支预算提高到 452 亿元，同比提高近 35%；中国电信 2023 年计划在产业数字化方面资本支出占比提高 9.1%；中国联通计划 2023 年算力网络资本开支将达到 149 亿元，占总资本开支比例将达到 19.4%，同比增长超过 20%。

图 6：从左至右分别（移动、电信、联通）三大运营商资本支出结构变化（单位：亿元；%）

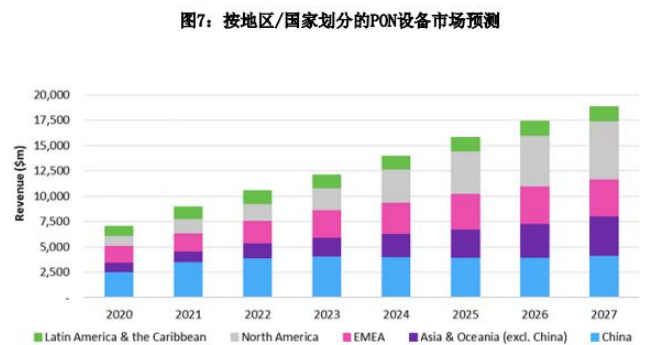
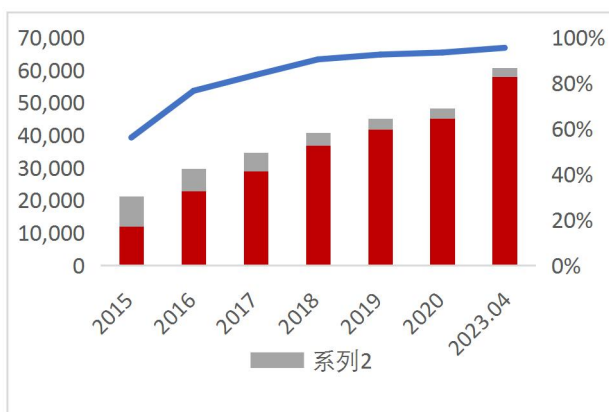


资料来源：C114 通信网援引自各公司财报，山西证券研究所

另一方面，光纤入户规模继续扩大，向 10G PON 升级已经是大势所趋，亚太运营商正引领全球接入网向 10G 速率升级。国内来看，截至 2023 年 4 月我国 FTTH/O 占比已超过 95%；截止 2022 年底，千兆及以上接入速率的固定宽带用户 9,175 万户，规模是上年末的 2.7 倍；建成具备千兆服务能力的 10G PON 端口数达 1,523 万个，较上年末接近翻倍。全球来看，根据 Omdia 数据显示，大多数国家的 FTTH 基础设施建设势头正在增强，预计到 2027 年，全球 FTTH 家庭渗透将超过 12 亿户；全球 PON 设备市场预计在 2027 年超过 180 亿美元。

图 7：国内互联网宽带接入用户数：FTTH/O

图 8：按地区/国家划分的 PON 设备市场预测





资料来源：wind，山西证券研究所

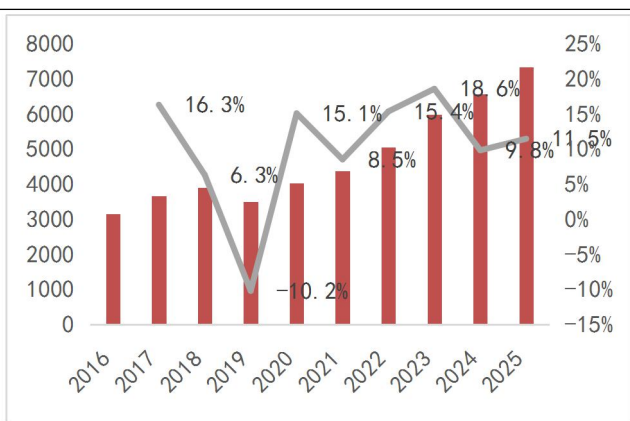
资料来源：中际旭创年报援引 Omdia，山西证券研究所

总结而言，电信市场方面，传统移动网络占资本开支比重有所减小，电信运营商加大算力网络布局作为第二增长曲线，叠加光纤入户规模扩大，10G PON 端口数量增加，整体市场对应光模块需求呈现稳中有增的发展状态。

### 2.1.2 数通市场推动演进，叶脊架构提升需求

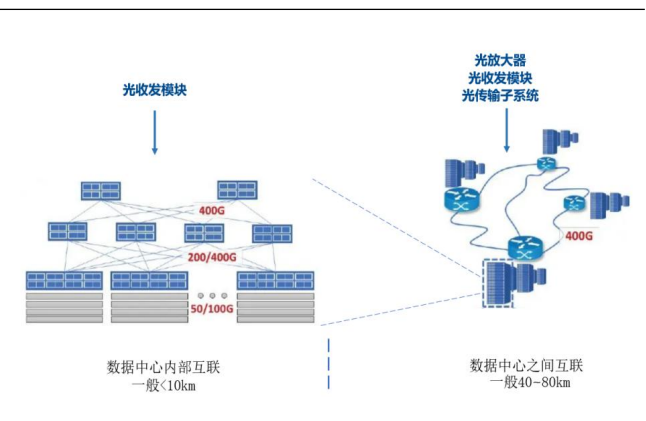
数通市场是目前光模块最大也是发展最快的市场，根据 Lightcounting 预测，数通市场将在几年迎来快速增长(未考虑 AI)。数通市场的快速发展主要得益于数据中心流量的大幅增加和数据中心网络架构的变化，具体来看，一方面，云计算、大数据、物联网等带来数据中心流量与交汇量的爆发，另一方面，数据中心架构升级，交换机之间连接数增加，两者共同推升光模块的用量。

图 9：全球数据中心光模块市场规模及预测（百万美元）



资料来源：源杰科技招股书援引 Lightcounting，山西证券研究所

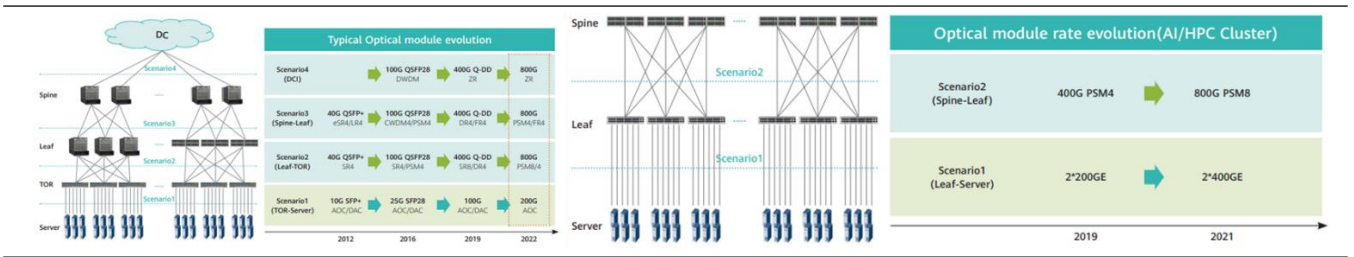
图 10：通用数据中心使用光模块



资料来源：德科立招股书，山西证券研究所

从数据中心架构来看，数据中心网络架构逐渐由传统的三层结构转向新型的叶脊（Spine-Leaf）结构，叶脊网络架构采用扁平化网络，每台脊交换机都与所有叶交换机相连，数据传输可以动态选择多条路径，能有效缓解宽带压力，提高数据传输效率，由于在叶脊网络架构下连接端口数倍数增加，因此光模块需求量随之提高。以通用服务器所需光模块进行测算，OR 交换机与机柜互联，一个机柜 6-10 台服务器，每一台都会和上层所有设备进行交叉互联。下行、上行端口收敛比一般是 3:1，由此测算，叶脊架构下 1 台服务器需要 4-6 只光模块，叶脊架构数据中心若需要 1 万台服务器，则需要约 6 万只光模块。

图 11：通用数据中心迭代：叶脊网络架构

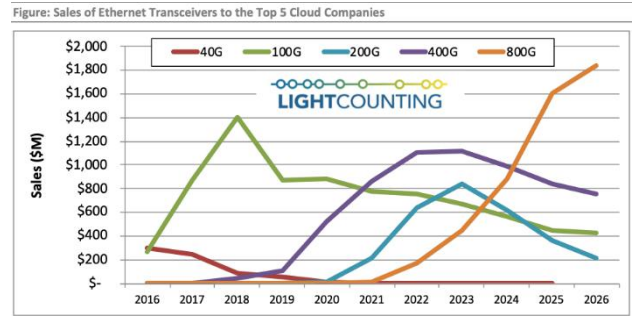


资料来源: ENABLING THE NEXT GENERATION OF CLOUD & AI USING 800GBS OPTICAL MODULESE 800G Pluggable MSA group, 山西证券研究所

此外，叶脊网络架构还提升了内部设备的连接密度、接口速率及交换容量，因此将推动光模块产品向高速率方向更新迭代。目前，光模块正处于由 400 G 向 800 G 进展的初期，800G 光模块有望在 2023 年后迎来快速放量。复盘来看，2018 年前光模块市场以 100G 为主，主要受电信市场 5G 建设驱动，2019 年后数通端资本支出快速提高，云计算推动光模块向 400G 迭代，根据 LightCounting 预测，2026 年前五大的云计算公司：阿里巴巴、亚马逊、Facebook、谷歌和 Microsoft 的光模块支出将增加到 30 亿美元以上，同时 800G 光模块将主导市场。

图 12：可插拔光模块向 800G 演进

图 13：5 大云厂商光模块销售情况



资料来源: ENABLING THE NEXT GENERATION OF CLOUD & AI USING 800GBS

资料来源: Lightcounting, 山西证券研究所

OPTICAL MODULESE 800G Pluggable MSA group, 山西证券研究所

## 2.2 AI: ChatGPT 开启技术革命，光模块市场星辰大海

本部分我们将从两部分阐述，首先，以 ChatGPT 为起始的 AI 革命开启，算力需求的爆发成为光模块市场的主要推动力；第二部分以英伟达计算机和 GPT-3 大模型为例进行测算。

### 2.2.1 AI 大模型打开算力需求天花板，云厂商业绩超预期催化后续支出

**ChatGPT 的出现引领 AI 浪潮：** ChatGPT 是美国初创公司 OpenAI 发布的人工智能 (AI) 对话聊天机器人软件 (模型)，是一种大型语言模型，ChatGPT 使用基于 GPT-3.5 架构的大型语言模型 (LLM)，能完成撰写邮件、代码、翻译等任务，以及实现更广泛的搜索服务等。随后 AI 军备竞赛开启，并大幅拉动算力

需求。目前国内已经出现多个 AI 大模型，腾讯混元大模型、百度文心大模型、鹏城实验室推出的盘古大模型、中国科学院自动化研究所研发的紫东太初大模型、阿里 AI 预训练模型 M6 等，未来大语言模型的数量有望继续增加。

表 1：今年 AI 大模型军备竞赛不完全节选

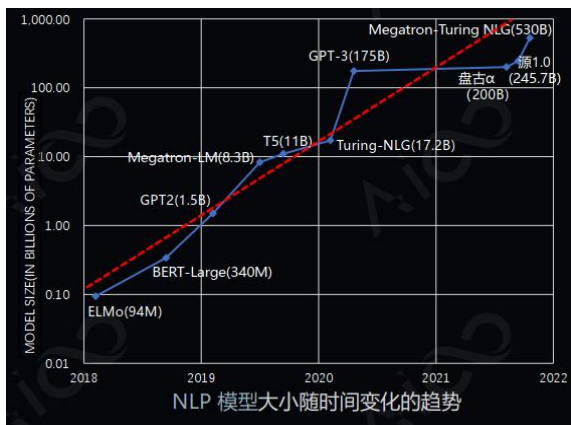
时间	公司	大模型/产品	事件
2月7日	谷歌	Bard	谷歌发布对话型 AI 系统 Bard 迎战 ChatGPT
2月8日	微软		微软将 ChatGPT 接入 Bing 搜索
3月15日	OpenAI	GPT-4	发布 GPT-4，支持多模态输入
3月16日	微软	Copilot	发布了植入 GPT-4 技术的 Copilot
3月16日	百度	文心一言	发布文心一言并启动内测
3月20日	阿里达摩院	文本生成 视频大模型	
3月24日	OpenAI	ChatGPT Plugin	支持第三方插件接入，并同时开源知识库检索插件源代码
3月27日	百度	文心千帆	面向客户提供企业级大语言模型服务
3月28日	腾讯		腾讯 AI Lab 发布自研 3D 游戏场景自动生成解决方案

资料来源：公开资料整理，山西证券研究所

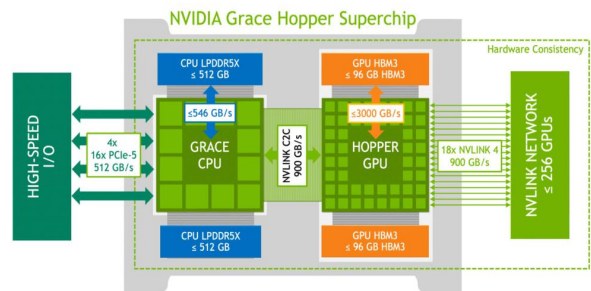
具体来看，AI 大模型分为**训练侧**和**推理侧**：**训练过程**：又称学习过程，是指通过大数据训练出一个复杂的神经网络模型，通过大量数据的训练确定网络中权重和偏置的值，使其能够适应特定的功能；**训练过程需要较高的计算性能、需要海量的数据**、训练出的网络具有一定通用性。**推理过程**：又称判断过程，是指利用训练好的模型，借助在训练中已确定参数的神经网络模型进行运算，使用新数据推理出各种结论。大模型训练过程的影响因素主要是**模型参数量**、**训练数据量**和**芯片算力**，在模型参数量方面，根据浪潮科技《AI 算力集群方案设计与优化》，过去 4 年大模型参数量从 94M 增加到 530B，增长了近 5600 倍；**算力芯片**目前主要由英伟达 GPU 提供，并在超级计算机中对应光模块的使用。英伟达 AI GPU 主要有 A100，H100 等，并在 5 月末新推出的超级计算机 DGX GH200 中搭载了其超级芯片 Grace Hopper，Grace Hopper 将 NVIDIA Hopper GPU 与 NVIDIA Grace CPU 结合在一起，所使用的 NVIDIA NVLINK-C2C 可提供 900Gb/s 的总带宽，大大提高在超级计算机中对应光模块的需求量。

图 14：神经网络模型持续扩大

图 15：Grace Hopper 芯片架构



资料来源：AI 算力集群方案设计与优化，山西证券研究所



资料来源：英伟达，山西证券研究所

此外，云厂商 2022 年业绩超出预期，英伟达数据中心营收创历史新高，主要受益于云商及企业客户对训练 AI 加大对 GPU 芯片的需求，进一步释放利好信号。随海内外云厂商陆续推出 AI 大模型及 AI 训练加码，高增长确定性将得到进一步延续。

表 2：业绩超预期催化后续需求

事件	主要厂商	厂商业绩
业绩超预期	英伟达	FY24Q1 实现营收 71.9 亿美元（YoY+13%），数据中心业务实现营收 42.8 亿美元（YoY+14%，QoQ+18%）。
	迈威尔	FY24Q1 实现营收 13.2 亿美元（YoY+9%）。
	Lumentum	FY2023Q3 实现营收 3.83 亿美元（YoY+3.0%）。

资料来源：wind，山西证券研究所

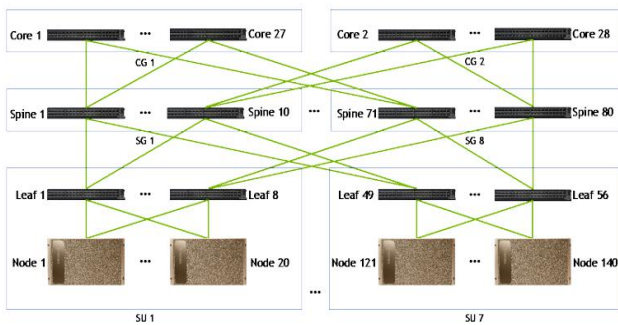
## 2.2.2 从英伟达计算机和大模型算力预测：800G 光模块需求量显著提升

首先我们以英伟达 DGX SuperPOD 为例测算单颗 GPU 与光模块的比例。DGX A100 和 DGX H100 网络集群主要使用 InfiniBand 和以太网两类，根据网络功能分为计算、存储、InBand 和 Out-of-Band 四种，其中计算和存储使用 IB 网络，InBand 和 Out-of-Band 使用以太网。IB 网络涉及 AI 运算和存储，是光模块的主要需求来源，而以太网则需求较小，因此下面测算仅考虑 IB 网络对应的光模块需求。

以 140 个节点的 DGX A100 集群为例，每台 DGX A100 服务器需要 8 颗 A100 GPU 芯片，共需  $140 \times 8 = 1120$  颗芯片；集群中每 20 台 DGX A100 服务器组成一个 SU（scalable units，可扩展单元），共 7 个 SU；计算侧网络采用三层交换机的胖树(fat-tree)拓扑结构，一到三层分别称为 Leaf（叶）交换机、Spine（脊）交换机、Core（核心）交换机。IB 网络架构下实现了完整的胖树（Fat-Tree）拓扑结构，胖树架构在训练场景中最优的配置是上下行端口数一样，即无阻塞的网络，具体来看：（1）第一层：每个 SU 中配置 8 台叶交换机，共

有 56 台叶交换机；根据图 15 所示，服务器与叶交换机共需 1120 条线缆；（2）第二层：每 10 台脊交换机组成一个交换机组 Spine Group（SG），每个 SU 的第一个叶交换机连接到 SG1 中的每个交换机，每个 SU 的第二个叶交换机连接到 SG2 中的每个交换机，共有 80 个脊交换机，根据图 15 所示，叶交换机与脊交换机共需 1124 条线缆；（3）第三层：每 14 台核心交换机组成一个交换机组 Core Group（CG），对应共 28 个核心交换机，根据图 15 所示，脊交换机与核心交换机共需 1120 条线缆。

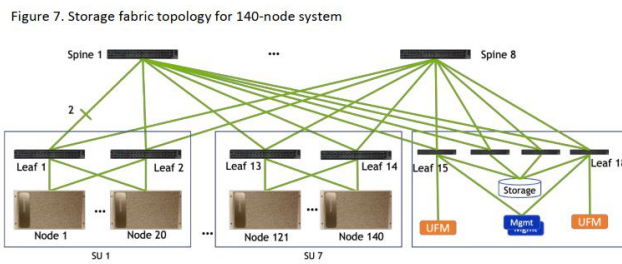
图 16：140 节点的 NVIDIA DGX SuperPOD 计算侧架构



资料来源：英伟达，山西证券研究所

存储侧采用两层结构，即只有叶交换机和脊交换机，根据图 17 所示存储侧叶和脊分别有 18 和 8 共 26 台交换机，对应共 660 条线缆。

图 18：140 节点的 NVIDIA DGX SuperPOD 存储侧架构



资料来源：英伟达，山西证券研究所

考虑使用 A100 系统的 NVIDIA DGX SuperPOD 白皮书中对应计算侧和存储侧线缆全部采用 NVIDIA MF1S00-HxxxE，为 AOC 有源光缆，因此每个端口对应一个光模块，即每条光缆对应 2 个光模块，因此计算侧和存储侧共需要  $(1120+1124+1120) * 2 + (280+92+288) * 2 = 8048$  个光模块，即单颗 GPU 对应所需 200G 光模块数量约为 1:7.2。

图 20：计算侧和存储侧的线缆全部采用 NVIDIA MF1S00-HxxxE

图 17：计算侧交换机及线缆数量

Table 2. Compute fabric switch and cable counts

Nodes	SUs	QM8790 Switches			Cables		
		Leaf	Spine	Core	Leaf	Spine <sup>1</sup>	Core
20 (Single SU)	1	8	5		160	164	
40	2	16	10		320	324	
60	3	24	20		480	484	
80	4	32	20		640	644	
120	6	48	80	24	960	964	960
140 (DGX SuperPOD)	7	56	80	28	1120	1124	1120

1. UFM Appliance is connected to two different spine switches.

The compute fabric uses NVIDIA Quantum QM8790 switches (Figure 6).

资料来源：英伟达，山西证券研究所

图 19：存储侧交换机及线缆数量

Table 3. Storage fabric counts

Nodes	SUs	Storage Ports	QM8790 Switches		Cables		
			Leaf	Spine	To-Node	To-Storage <sup>1</sup>	Spine
20	1	24	4	2	40	36	64
40	2	40	6	4	80	52	96
60	3	40	8	4	120	52	128
80	4	56	12	8	160	68	192
120	6	80	16	8	240	92	256
140	7	80	18	8	280	92	288

1. Includes connection to management servers and UFM.

资料来源：英伟达，山西证券研究所

Compute InfiniBand Cables <sup>1</sup>		
3374	200 Gbps QSFP56	DGX A100 systems, spine, and core NVIDIA MF1S00-HxxxE
Storage InfiniBand Cables <sup>1</sup>		
568		DGX A100 systems and spine
56 <sup>2</sup>	200 Gbps QSFP56	Storage NVIDIA MF1S00-HxxxE
16		Management nodes

资料来源：英伟达，山西证券研究所

5月29日英伟达最新推出 DGX GH200 超级计算机，搭载 256 个英伟达 GH200 Grace Hopper 超级芯片，每个超级芯片可视为一个服务器，通过 NVLINK SWITCH 进行互联。从结构上来看，DGX GH200 采用两层的胖树拓扑结构，第一、二层分别使用 96 个、36 个交换机，每台 NVLINK SWITCH 交换机拥有 32 个速率为 800 G 的端口；此外 DGX GH200 还配备了 24 个 NVIDIA Quantum-2 QM9700 IB 交换机用于 IB 网络。按照端口进行估算，L1 层距离较近因此假定使用铜缆进行连接不涉及光模块，L2 层 36 个交换机在无收敛胖树架构下，第二层交换机端口向下与 L1 层交换机上端口互联，因此共需要  $36 \times 32 \times 2 = 1152$  个 800G 光模块；IB 网络架构下，24 台交换机需要  $24 \times 32 = 768$  个 800G 光模块；则 DGX GH200 超级计算机共需要  $1152 + 768 = 1920$  个 800G 光模块，对应每颗 GH200 芯片 12 个 800G 光模块，相较于 DGX SuperPOD 在光模块数量和速率上均有显著提升。

图 21: NVIDIA DGX GH200 存储侧架构

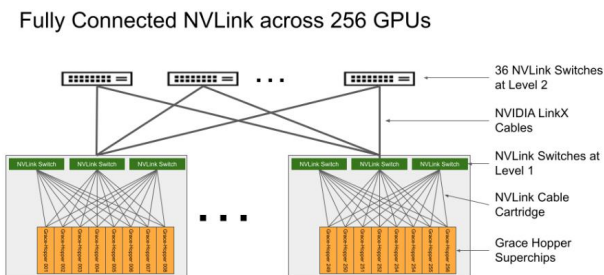


Figure 2. Topology of a fully connected NVIDIA NVLink Switch System across NVIDIA DGX GH200 consisting of 256 GPUs

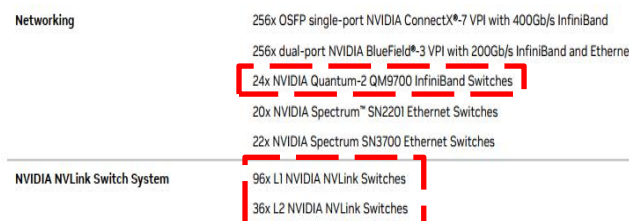
资料来源：英伟达，山西证券研究所

下面我们再从大模型训练侧和推理侧分别计算大语言模型所带来的光模块数量增量。

1) 训练侧：以 GPT-3 为例，首先，用来衡量计算量的单位为 FLOPs（浮点运算次数），根据 OPEN AI 的论文《Language Models are Few-Shot Learners》，GPT-3 中最大的模型（1746 亿参数）的训练大约需要  $3.14 \times 10^{23}$  次浮点运算（FLOPs）。

图 23: GPT-3 计算参数

图 22: NVIDIA DGX GH200 结构



资料来源：英伟达，山西证券研究所

Model	Total train compute (PF-days)	Total train compute (flops)	Params (M)	Training tokens (billions)	Flops per param per token	Mult for bwd pass	Fwd-pass flops per active param per token	Frac of params active for each token
T5-Small	2.08E+00	1.80E+20	60	1,000	3	3	1	0.5
T5-Base	7.64E+00	6.60E+20	220	1,000	3	3	1	0.5
T5-Large	2.67E+01	2.31E+21	770	1,000	3	3	1	0.5
T5-3B	1.04E+02	9.00E+21	3,000	1,000	3	3	1	0.5
T5-11B	3.82E+02	3.30E+22	11,000	1,000	3	3	1	0.5
BERT-Base	1.89E+00	1.64E+20	109	250	6	3	2	1.0
BERT-Large	6.16E+00	5.33E+20	355	250	6	3	2	1.0
RoBERTa-Base	1.74E+01	1.50E+21	125	2,000	6	3	2	1.0
RoBERTa-Large	4.93E+01	4.26E+21	355	2,000	6	3	2	1.0
GPT-3 Small	2.60E+00	2.25E+20	125	300	6	3	2	1.0
GPT-3 Medium	7.42E+00	6.41E+20	356	300	6	3	2	1.0
GPT-3 Large	1.58E+01	1.37E+21	760	300	6	3	2	1.0
GPT-3 XL	2.75E+01	2.38E+21	1,320	300	6	3	2	1.0
GPT-3 2.7B	5.52E+01	4.77E+21	2,650	300	6	3	2	1.0
GPT-3 6.7B	1.39E+02	1.20E+22	6,660	300	6	3	2	1.0
GPT-3 13B	2.68E+02	2.31E+22	12,850	300	6	3	2	1.0
<b>GPT-3 175B</b>	<b>3.64E+03</b>	<b>3.14E+23</b>	<b>174,600</b>	<b>300</b>	<b>6</b>	<b>3</b>	<b>2</b>	<b>1.0</b>

资料来源：《Language Models are Few-Shot Learners》，山西证券研究所

在使用英伟达 A100 GPU 对大模型进行训练的情况下，参照英伟达官网，可使用 FP16 的精度进行训练，对应 312 TFLOPS（312 万亿次浮点运算能力）；根据总计算量=GPU 数量\*GPU 算力\*运算时间，则 GPU 数量=总计算量/（GPU 算力\*运算时间），则在训练时间为 1 天的情况下 GPT-3 需要 GPU 数量为  $3.14 * 10^{23}$  FLOPs / ( $312 * 10^{12}$  FLOPs/s \* 86400s)，约为 **11648 张**，在训练十天的情况下需要  $3.14 * 10^{23}$  FLOPs / ( $312 * 10^{12}$  FLOPs/s \* 86400s\*10)=1165 张 GPU。根据前面的测算，假定在系统中每张 A100 GPU 需要 7 颗 200G 光模块，则 GPT-3 单日完成训练需要 8 万颗以上 200G 光模块。如果考虑到 FLOPS 的利用率约为 20%-30%，则需要光模块的数量将约为 35 万颗。

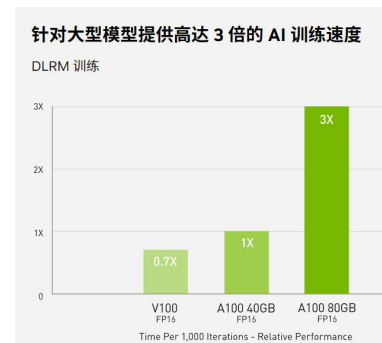
图 24：NVIDIA A00 TENSOR CORE GPU 规格

NVIDIA A100 TENSOR CORE GPU 规格 (SXM4 和 PCIe 外形规格)

	A100 80GB PCIe	A100 80GB SXM
FP64	9.7 TFLOPS	
FP64 Tensor Core	19.5 TFLOPS	
FP32	19.5 TFLOPS	
Tensor Float 32 (TF32)	156 TFLOPS   312 TFLOPS*	
BFLOAT16 Tensor Core	312 TFLOPS   624 TFLOPS*	
FP16 Tensor Core	312 TFLOPS   624 TFLOPS*	

资料来源：英伟达，山西证券研究所

图 25：A100 采用 FP16 精度训练大模型



资料来源：英伟达，山西证券研究所

图 26：FLOP 利用率（GPT-3 为 21.3%）

Model	# of Parameters (in billions)	Accelerator chips	Model FLOPS utilization
GPT-3	175B	V100	21.3%
Gopher	280B	4096 TPU v3	32.5%
Megatron-Turing NLG	530B	2240 A100	30.2%
PaLM	540B	6144 TPU v4	46.2%

资料来源：《PaLM: Scaling Language Modeling with Pathways》，山西证券研究所

目前大模型数量快速增加，据不完全统计，截至目前，参数在 10 亿规模以上的大模型国内已发布 79 个；同时参数快速提高，最高可达万亿以上，对算力及光模块的需求同样将大幅上涨。根据 GPT-3 的参数，我们假定一个在 GPT-3 框架下的万亿规模大模型，根据上图可知参数量与训练侧所需算力大致呈线性同比，计算得用 H100 训练该模型大致需要 1.5-2 万张 H100 GPU，考虑 FLOPS 利用率的情况下需要 7.5-10 万颗 GPU，对应每张 H100 GPU 大致需要 2.5 颗 800 G 光模块，则在 GPT-3 框架下的万亿规模大模型约需要 20 万颗 800 G 光模块。

图 27：NVIDIA H00 GPU 规格

外形规格	H100 SXM	H100 PCIe
FP64	34 teraFLOPS	26 teraFLOPS
FP64 Tensor Core	67 teraFLOPS	51 teraFLOPS
FP32	67 teraFLOPS	51 teraFLOPS
TF32 Tensor Core	989 teraFLOPS*	756teraFLOPS*
BFLOAT16 Tensor Core	1979 teraFLOPS*	1513 teraFLOPS*
TF16 Tensor Core	1979 teraFLOPS*	1513 teraFLOPS*
FP8 Tensor Core	3958 teraFLOPS*	3026 teraFLOPS*
INT8 Tensor Core	3958 TOPS*	3026 TOPS*
GPU 显存	80GB	80GB
GPU 显存带宽	3.35TB/s	2TB/s
解码器	7 NVDEC 7 JPEG	7 NVDEC 7 JPEG

资料来源：英伟达，山西证券研究所

图 28：部分国内外大模型参数量节选

大模型	开发者	参数量(亿)
GPT-3	OPEN AI	1750
MT-NLG	微软、英伟达	5300
Switch Transformer	谷歌	1.6万
INTERN	商汤	100
盘古NLP	华为云、循环智能	1000
盘古 α	华为云、北京大学	2000
源 1.0	浪潮信息	2500
ERNIE 3.0	百度	2600
八卦炉 (BAGUALU)	清华大学、阿里达摩院等	174万
PaLM 2	谷歌	3400

资料来源：电子发烧友，山西证券研究所

2) 推理侧：目前大模型处于快速发展阶段，以训练侧用量为主；而当大模型成熟后光模块的主要需求将转向推理侧。推理侧光模块用量可通过用户流量来计算：乐观预期下，假设在稳定状态时 ChatGPT 日均访问量为 10 亿次，每次 ChatGPT 回答 1000 字，粗略按照 token 与汉字比例为 1: 1 估算约 1000 个 token，共计 10000 亿个 token；按照与推理侧保持一致用 GPT-3 作为预测的大模型，根据《Scaling Laws for Neural Language Models》，训练阶段 FLOPS/每个参数/每个 token 的比值为 6，推理阶段约为 2，根据则一次访问需要的算力为 2\*10000 亿\*1750 亿共计 3.5\*10<sup>23</sup>FLOPS；由于推理阶段算力需求相对训练较低，使用 A100 芯片存在成本和显存上的不足，因此选用 A30 甚至 A10 芯片作为推理芯片。以 A30 算力测算，则需求量计算方法同理，约为 2.5 万张 GPU，考虑到 FLOPS 利用率约为 20%-30%，则约为 13 万张 GPU；同理，我们假设一个在 GPT-3 框架下的万亿规模大模型，以 A30 算力测算得到推理侧大约需要 80 万张 GPU；对应所需光模块数量上百万颗。同理在中性预期下稳定状态时 ChatGPT 日均访问量为 5 亿次，悲观预期下稳定时 ChatGPT 日均访问量为 1 亿次，则对应所需 GPU 数量分别 40 万张/8 万张，中性预期下对应光模块需求量超百万颗，悲观预期下对应光模块数量接近百万颗。

图 29：NVIDIA A30 Core GPU 规格

图 30：A10 技术规格和功能



### A30 Tensor Core GPU 规格

FP64	5.2 teraFLOPS
FP64 Tensor Core	10.3 teraFLOPS
FP32	10.3 teraFLOPS
TF32 Tensor Core	82 teraFLOPS   165 teraFLOPS*
BFLOAT16 Tensor Core	165 teraFLOPS   330 teraFLOPS*
FP16 Tensor Core	165 teraFLOPS   330 teraFLOPS*
INT8 Tensor Core	330 TOPS   661 TOPS*
INT4 Tensor Core	661 TOPS   1321 TOPS*
媒体引擎	1个光流加速器 (OFA) 1个 JPEG 解码器 (NVJPEG) 4个视频解码器 (NVDEC)
GPU 显存	24GB HBM2
GPU 显存带宽	933GB/s
互联	PCIe Gen4: 64GB/s 第三代 NVLINK: 200 GB/s**
外形规格	双插槽、全高、全长 (FHFL)

资料来源：英伟达，山西证券研究所

### A10 技术规格和功能

FP32	312 兆次浮点运算
TF32 Tensor 核心	62.5 兆次浮点运算   125 兆次浮点运算*
BFLOAT16 Tensor 核心	125 兆次浮点运算   250 兆次浮点运算*
FP16 Tensor 核心	125 兆次浮点运算   250 兆次浮点运算*
INT8 Tensor 核心	250 兆次浮点运算   500 兆次浮点运算*
INT4 Tensor 核心	500 兆次浮点运算   1,000 兆次浮点运算*
RT 核心	72 个 RT 核心
编码/解码	1 个编码器 2 个译码器 (+AV1 解码)
GPU 内存	24GB GDDR6
GPU 内存带宽	每秒 600 GB
互联技术	第四代 PCIe 64GB/秒
尺寸规格	单插槽、全高全长尺寸 (FHFL)

资料来源：英伟达，山西证券研究所

## 2.3 格局：国内厂商主导市场，随技术迭代各有优势

### 2.3.1 产业东移，中国厂商逐渐占领高地

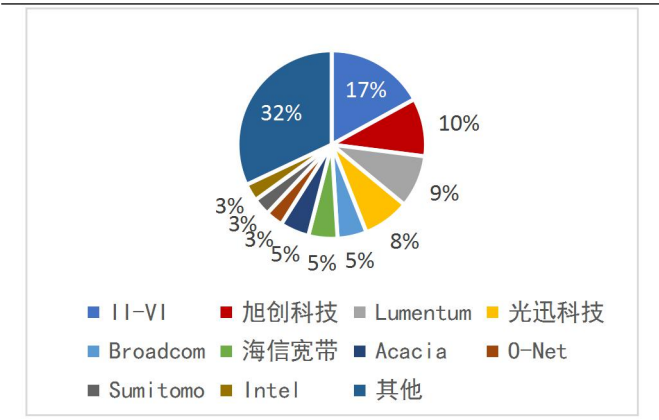
从市场份额上来看，根据 LightCounting，在过去的十年里，中国的光器件和模块供应商凭借成本优势逐渐在全球市场上获得份额，目前在全球光模块市场占主导地位。2015 年前，全球前十大光模块厂商仅光迅科技一家中国企业；2021 年，中际旭创等五家供应商进入全球前十，主要厂商相较前一年市场份额均有提升；2022 年，光模块全球市前 10 名中国占据 7 家，其中中际旭创、Coherent、思科、华为四家厂商占据全球光模块市场份额超过 50%，中际旭创和 Coherent 分别获得近 14 亿美元的收入。

表 3：光模块全球市占率排名

2010	2016	2018		2021	2022
Finisar	Finisar	Finisar	1	中际旭创 &II-VI	中际旭创 &Coherent
Opnext	海信宽带	中际旭创	2		
Sumitomo	光迅科技	海信宽带	3	华为海思	Cisco (Acacia)
Avago	Acacia	光迅科技	4	Cisco (Acacia)	华为海思
Source Photonic	FOIT (Avago)	FOIT (Avago)	5	海信宽带	光迅科技
Fujitsu	Oclaro	Lumentum	6	Broadcom	海信宽带
JDSU	中际旭创	Acacia	7	新易盛	新易盛
Emcore	Sumitomo	Intel	8	光迅科技	华工正源
WTD	Lumentum	AOI	9	Molex	Intel
NeoPhotonics	Source Photonic	Sumitomo	10	Intel	Source Photonic

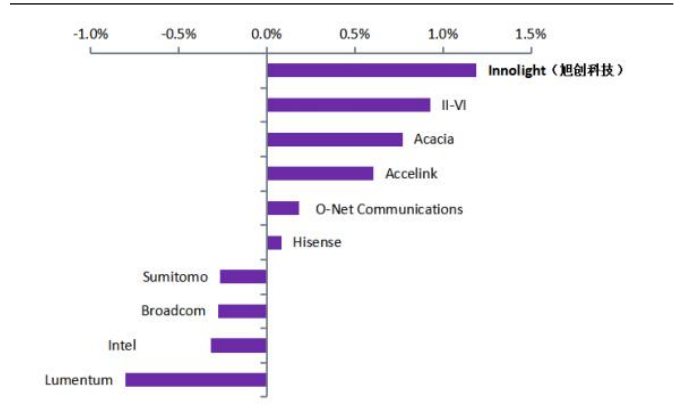
资料来源：中际旭创年报，Lightcounting，山西证券研究所

图 31：2021 年 光模块主要厂商市占率



资料来源：中际旭创年报援引 Omdia，山西证券研究所

图 32：2021 年主要厂商市场份额变动情况

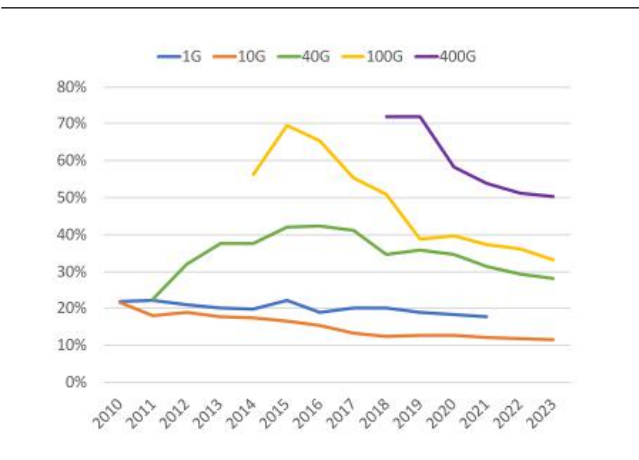


资料来源：中际旭创年报援引 Omdia，山西证券研究所

### 2.3.2 技术迭代：CPO 和硅光有望成为长期解决方案

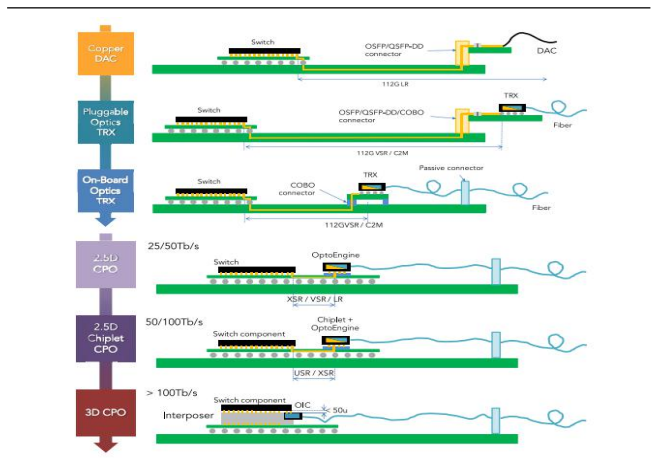
从光模块发展进程来看，随光模块速率增高，光模块在以太网交换机中成本占比快速提高，同时从 1.6T 开始，传统可插拔速率升级或达到极限，CPO 共同封装光子是业界公认未来高速率产品形态。CPO (Co-Packaged Optics, 共封装光学) 指把光引擎和交换芯片共同封装在一起的光电共封装，最终取代光模块。相较于传统可插拔形式，CPO 缩短交换芯片和光引擎间的布线距离，进而降低电信号驱动功耗。根据 Cisco 官网，把 51.2T 系统中的可插拔光模块替换为 CPO 后，将交换 ASIC 与光引擎连接所需的功率可减少 50%，整机系统总功率减少 25-30%。

图 33：不同速率光模块在以太网交换机中成本占比



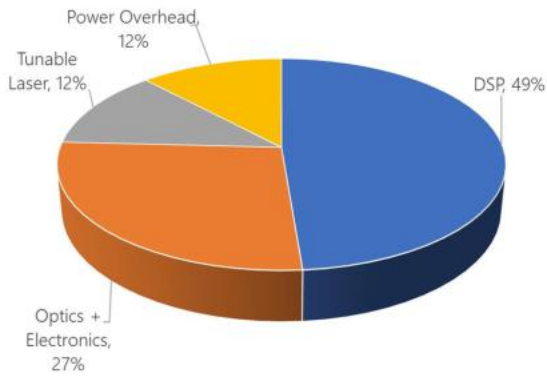
资料来源：Co-packaged datacenter optics Opportunities and

图 34：CPO 封装示意图

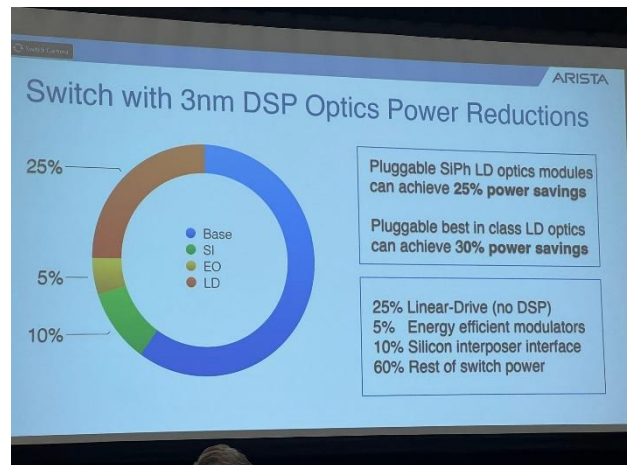


资料来源：Co-packaged datacenter optics Opportunities and





资料来源：R. Nagarajan, et.al., "Low Power DSP-Based Transceivers for Data Center Optical Fiber Communications", Journal of Lightwave Technology 39, 5221(2021),山西证券研究所

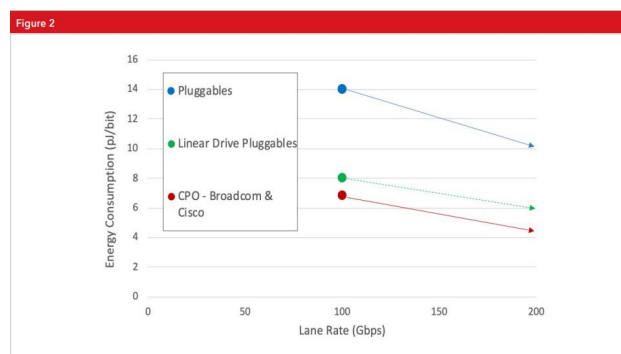


资料来源：fabricatedknowledge 援引 Arista, 山西证券研究所

此外，由于不再采用 DSP，系统的延迟和成本降低，可应用到对延迟要求比较高的场景，例如高性能计算中心(HPC)中 GPU 之间的互联；相对于 CPO，由于 LPO 仍然采用可插拔模块的形式，未进行较大的封装形式革新，可以利用成熟的光模块供应链，因此被认为是 800G 光模块时代最具潜力的技术路线。

图 39：LPO 功耗可下降 50%左右

图 40：传统光模块、LPO 与 CPO 各个维度的比较



资料来源：2023 Post-Show Report, 山西证券研究所

	Retimed Pluggable	Linear-Drive Pluggable	Linear-Drive CPO
Power	●	●	●
Cost	●	●	●
Latency	●	●	●
Product Maturity	●	●	●
Serviceability	●	●	●
Late Binding Commitment	●	●	●
Link Performance	●	●	●
Link Accountability	●	●	●
Interoperable Ecosystem	●	●	●

资料来源：SILICON PHOTONICS, LINEAR DRIVE PLUGGABLE AND CO-PACKAGED OPTICS, 山西证券研究所

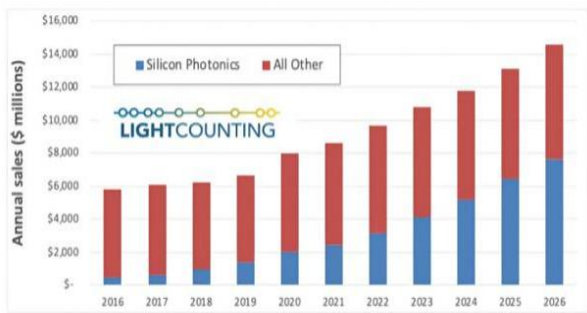
硅光子技术是基于硅和硅基衬底材料，利用现有 CMOS 工艺进行光器件开发和集成的新一代技术。传统光模块主要采用高速电芯片、光芯片和光学组件等器件封装而成，本质上属于“电互联”；硅光技术的核心是采用激光束代替电子信号传输数据，本质上是“光互联”。随着光模块向 400G、800G 甚至 1.6T 等高速率演进，以 Tb/s 的光纤传输速度或将成为光通信传输速率瓶颈，而硅光子集成技术具备的超高传输速率能打破这一瓶颈：根据摩尔定律，相同尺寸芯片中能容纳的晶体管数量，因为制程技术推进，每 18~24 个月会增加一倍；由于芯片是电讯号，传输会有讯号损失，即使单位面积晶体管数量渐增，仍无法避免电耗

损的问题，硅光技术以光讯号代替电讯号，或成为突破摩尔定律的关键。

此外，硅光方案集成度高，且在峰值速度、能耗、成本等方面均具有良好表现。一方面，当前主流的光集成技术以稀有材料磷化铟作为主要材料，材料成本昂贵，难以实现大规模集成；硅材料价格低廉且已经成熟应用于电子集成电路，适合规模化生产。另一方面，以磷化铟为材料的光集成技术只负责数据的交换，不涉及数据的存储与处理，不利于通信信息安全。而以硅为材料的光集成技术兼具数据的交换、存储以及处理，是下一代光通信的技术趋势。

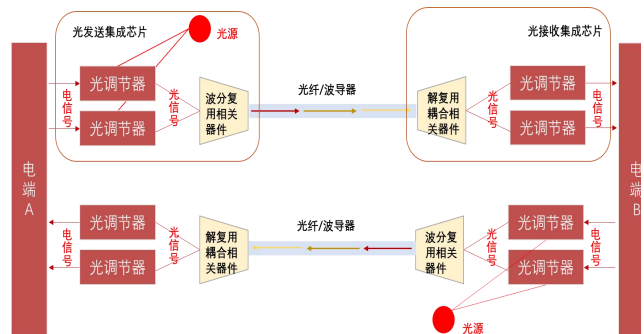
硅光技术既可用在传统可插拔光模块中，也可以用在 CPO 方案中，硅光子模块和超大规模 CMOS 芯片以共封装形式（CPO）封装在一起将成为未来高速率传导的主要方案。根据 Lightcounting 的预测。硅光将在 2021-2026 年继续获得市场份额，全球硅光模块市场将在 2026 年达到近 80 亿美元，有望占到一半的市场份额。

图 41：硅光模块的历史销售及预测



资料来源：中际旭创年报援引 Lightcounting，山西证券研究所

图 42：硅光技术图解

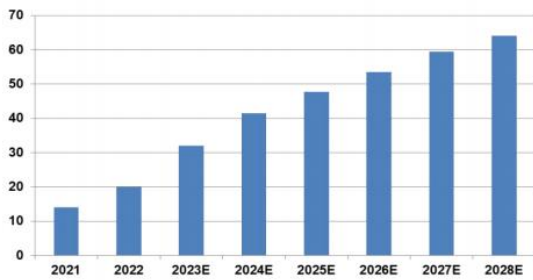


资料来源：中际旭创年报，山西证券研究所

数据中心光互联方案可根据其传输距离来选择两种支撑技术，一种是直接探测技术，另一种是相干探测技术，其中直接探测的应用场景更适合相对短距离互联。相干技术是利用要传输的信号来改变光载波的频率、相位和振幅，凭借着高容量、高信噪比等优势在城域网内的长距离 DCI 互联中得到广泛应用。随着单通道传输速率的提高，现代光通信领域越来越多的应用场景开始用到相干光传输技术，相干技术也从过去的骨干网下沉到城域甚至边缘接入网。相干技术预计在 2025 年后可以开始商用，Omdia 预计 2025 年相干将达到 250 万支规模；2022-2025 年，400G 相干光模块年复合增长率将超 40%。根据讯石预测，2023 年-2028 年，ZR 光模块市场规模将从约 30 亿元，增长至超过 60 亿元。

图 43：全球 ZR 光模块市场规模（单位：亿元）

图 44：标准 400G 以太网接口的 PMD, 特征和应用



资料来源：中际旭创年报援引讯石，山西证券研究所

PMD	距离	应用	技术
DAC	2-3m	机架内和服务器	无源铜缆, 50G PAM4
SR8	100m	企业	并行多模, 50G/λ - PAM4
DR4	500m	数据中心和企业	并行单模100G/λ - PAM4, 还可以支持4x100G分支跳线
FR4	2km	大型数据中心	单模, 100G/λ, PAM4
LR8	10km	电信范围	单模, 50G/λ, PAM4
ZR	80km	城域和数据中心互连	单模/相干, PAM4

资料来源：可插拔相干光模块白皮书，山西证券研究所

超高速光通信调制器芯片与模块是用于长途相干光传输和超高速数据中心的核⼼光器件，有望跟随光网络设备市场持续保持增长。目前行业内光调制技术主要有三种：基于硅光、磷化铟和铌酸锂材料平台的电光调制器。其中，硅光调制器主要是应用在短程的数据通信用收发模块中，磷化铟调制器主要用在中距和长距光通信网络收发模块，铌酸锂电光调制器主要用在 100Gbps 以上的长距骨干网相干通讯和单波 100/200Gbps 的超高速数据中⼼中。传统铌酸锂电光调制器为体材料铌酸锂调制器：体材料铌酸锂调制器具有带宽高、稳定性好、信噪比高、传输损耗小、工艺成熟等优点，但在传输速率需求不断提升的形势下，体材料铌酸锂调制器也在一些性能上遭遇瓶颈，而且体积较大，不利于集成。**新一代薄膜铌酸锂调制器芯片技术将解决这些问题。**具有“光学硅”之称的铌酸锂材料通过最新的微纳工艺，制备出的薄膜铌酸锂调制器具有高性能、低成本、小尺寸、可批量化生产且与 CMOS 工艺兼容等优点，是未来高速光互连极具竞争力的解决方案。

图 45：薄膜铌酸锂调制器与传统铌酸锂调制器对比

图例	传统铌酸锂调制器	薄膜铌酸锂调制器
折射率对比度 Δn	~0.1, 弱光学限制	~0.7, 强光学限制
调制效率	~10 V·cm	~2 V·cm
器件长度	~10 cm	0.5-2 cm
电光带宽	~ 35 GHz	> 67 GHz
集成度	难以集成	可大规模集成

资料来源：光电汇 OESHOW，山西证券研究所

### 2.3.3 厂商机遇：龙头升级换代，新秀弯道超车

光模块高速率低功耗高集成的趋势带来诸多新机遇，龙头厂商在高速率和多种技术方案下全面推进，其他厂商在高速率方面亦有基于不同技术方案的研究，或成为弯道超车，提高市占率的有效途径。具体来看，800G 方面，光模块厂商在传统 800 光模块和基于硅光方案的 800G 均有不同程度进展。

表 4：光模块厂商 800G 光模块进展

光模块厂商	800G 光模块产品
中际旭创	2023OFC 推出了其基于 5nm DSP 和先进硅光子技术的第二代 800G 模块，同时拥有功耗低于 14W 的 800G OSFP DR8+和 2xFR4 光通信模块。800G 光模块已批量出货。
新易盛	2023OFC 现场演示基于薄膜铌酸锂（TFLN）调制器技术的 800G OSFP DR8 光模块产品，搭配 5nm DSP 芯片，功耗 11.2W。同时推出 LPO 光模块，包含 EML、TFLN 和 SiPh 三种方案。
华工科技	公司的 800G SR8 已经在国内市场头部厂商送样测试，DR8 和 FR8 产品在微软和英伟达同样，预计 23 年下半年出货量将快速增长。
剑桥科技	公司的 800G 光模块基于传统 EML 和 硅光两种方案，对于薄膜铌酸锂方案也在积极关注。公司也推出了线性驱动的 800G 光模块产品。
光迅科技	2023OFC 上展示了 800G QSFP-DD800 的 SR8 光模块产品，同时公司拥有 800G QSFP-DD 2x400G FR4 和 DR8 光模块；800G 光模块已有出货，是否做大批量生产取决于客户订单。
博创科技	公司具备 800G 硅光模块的产品能力，同时也在研发 CPO 相关产品

资料来源：各公司年报、wind、投资者调研纪要、投资者问答、官网，山西证券研究所

新兴技术方面，主要厂商在各项技术上的布局都有不同进度和侧重。从厂商角度来看，龙头中际旭创和新易盛布局较为全面且进展较快，高速率硅光芯片均已研发成功，CPO 技术研发亦有序推进；其他厂商在技术拓展则各有侧重，华工科技在硅光，剑桥科技在 LPO，天孚和联特在 CPO 方向较为突出；从技术端来看，硅光集成技术的研发及产业化主要集中于硅光芯片制造，同时厂商多基于自身的硅光芯片配合下游客户研发 CPO 方案。

表 5：主要厂商技术研发进展

光模块厂商	CPO	LPO	硅光	相干技术
中际旭创	CPO（光电共封）技术和 3D 封装技术持续研发进程中。	有 LPO 的方案与技术	在 OFC 2022 现场展示基于自主设计硅光芯片 800G 可插拔 OSFP2*FR4 和 QSFP-DD800DR8+ 硅光光模块，2023 OFC 推出了其基于 5nm DSP 和先进硅光子技术的第二代 800G 模块。公司 400G 和 800G 硅光模块全部采用自研的硅光芯片，1.6T 光模块和 800G 硅光模块	推出业界最高输出功率 400G ZR / OpenZR+ QSFP-DD 相干模块，输出光功率可以达到+5dBm，模块功耗业界最低。相干系列产品等已实现批量出货

			已开发成功并进入送测阶段	
新易盛	已有 CPO 技术布局。	已推出 LPO 方案的 800G 光模块产品，已推出多款 LPO 方案产品。	已成功推出基于硅光解决方案的 800G、400G 光模块产品；子公司 Alpine 在硅光模块及硅光子芯片技术提供可靠保证	已成功推出 400G ZR/ZR+相干光模块产品
华工科技			已具备从硅光芯片到硅光模块的全自研设计能力，800G 硅光模块于 2022 Q3 推出市场；	海外导入阶段，正在进行客户验证。
剑桥科技	CPO 方面主要研发外置光源。	收购海外主推 LPO 的芯片厂商；7 月和 9 月送样测试各种类型的 LPO 光模块	完成 400G DR4/DR4+硅光产品的版本迭代，实现小批量发货；硅光 800G 光模块在北美主要网络设备供应商测试顺利。	
天孚通信	开发 CPO 用光引擎，为客户提供一站式解决方案，CPO 光引擎目前进展快。			
联特科技	基于 SIP（硅光）和 TFLN（薄膜铌酸锂）调制技术的 800G 光模块产品会相继推出；基于数据中心的 800G/1.6T 研发 CPO 项目进行中		“基于硅光技术的下一代光模块和光器件开发”项目进行中	
博创科技	CPO 产品正在研发		一直在硅光方面有布局，基于硅光子技术的 400G-DR4 硅光模块已实现量产出货，800G 硅光模块开发中	
光迅科技	CPO 技术布局中		2023 年在 OFC2023 上推出 1.6T QSFP-XD DR8 高速光模块，400G QSFP-DD ZR+相干光模块等产品；200G/400G/800G 硅光芯片已具备量产能力。	400G 相干光模块产品技术行业领先。

资料来源：各公司年报、投资者问答、投资者调研纪要、wind、官网，山西证券研究所

### 2.3.4 向上布局芯片，一体化掌握话语权

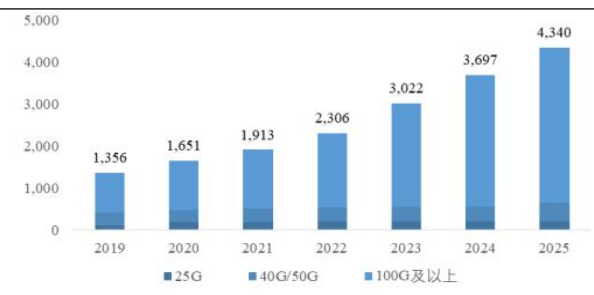
根据中际旭创披露的 2016 年 1-3 月光模块成本构成，光芯片及组件占光模块成本比超过 50%，是光模



块中成本占比最高的部分。我国光芯片企业已基本掌握 2.5G 和 10G 光芯片的核心技术，25G 及以上高速率光芯片方面，我国技术水平相对落后：根据 ICC 统计，25G 光芯片的国产化率约 20%，但 25G 以上光芯片的国产化率仍较低约 5%，光模块或光器件厂商仍然以外购为主。

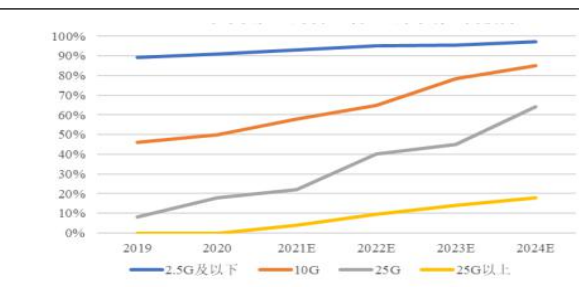
在对高速传输需求不断提升背景下，25G 及以上高速率光芯片市场增长迅速，且我国自主研发进步较快。根据 Omdia 对数据中心和电信场景激光器芯片的预测，2019 年至 2025 年，25G 以上速率光模块所使用的光芯片占比逐渐扩大，整体市场空间将从 13.56 亿美元增长至 43.40 亿美元，年均复合增长率将达到 21.40%。同时，我国光芯片厂商的全球份额将进一步提升，根据 ICC 预测，2019-2024 年，中国光芯片厂商销售规模占全球光芯片市场的比例将不断提升，其中高速率光芯片增长更快。

图 46：高速率模块光芯片市场空间及预测（百万美元）



资料来源：源杰科技招股书援引 Omdia，山西证券研究所

图 47：2019-2024 年中国光芯片占全球光芯片比例预测



资料来源：源杰科技招股书援引 ICC，山西证券研究所

对光模块厂商而言，长期来看，拥有芯片能力的企业将拥有更多自主权和议价权，中际旭创等厂商已纷纷通过参股或收购等形式向上游进行一体化布局。其中，源杰科技为我国芯片龙头，高速率光芯片领域的技术实力已可以对标世界领先水平，并同时拥有华为作为主要股东，虽然由于下游云厂商对芯片供应商要求较为严格，因此处于可信度等原因，目前光模块厂商难以立即更换芯片供应商，但随着 800G 光模块放量，芯片需求量快速提高，源杰科技的光芯片未来有望逐步得到云厂商认可，提高在海外的渗透率。此外，新易盛于 2022 年完成对 Alpine 的收购，将其硅光芯片用于硅光产品；光迅科技、华工科技等均可实现一定程度的光芯片自供。

图 48：光模块厂商参股或收购芯片供应商情况

芯片供给及采购情况	光模块厂商	形式	光芯片厂商
①苏州旭创的光器件、集成电路芯片等核心零部件主要采购自美国、日本、香港、中国台湾等国家和地区； ②成都博翰的 2.5G 速率产品的光芯片主要由国内厂商提供，10G 及以上速率的光器件、集成电路芯片等核心零部件主要采购自美国、日本、台湾等国家和地区。	中际旭创	参股	源杰科技
收购 IPX 公司和法国阿尔玛伊公司，具备 DWDM AWG 芯片从晶圆、芯片到模块、系统的生产能力。实现了 10G 及以下速率光芯片批量供货，25G 光芯片规模出货，其中 25G VCSEL 基本可以自供，25G DFB 60% 可以自供。	光迅科技		光迅科技
云岭光电在 10G 系列光芯片自供率占比近 80%。	华工科技	参股	云岭光电 长光华芯
公司高速光模块所需芯片主要源于海外供应商。硅光产品使用了子公司 Alpine 自研的硅光芯片产品。	新易盛	收购	Alpine

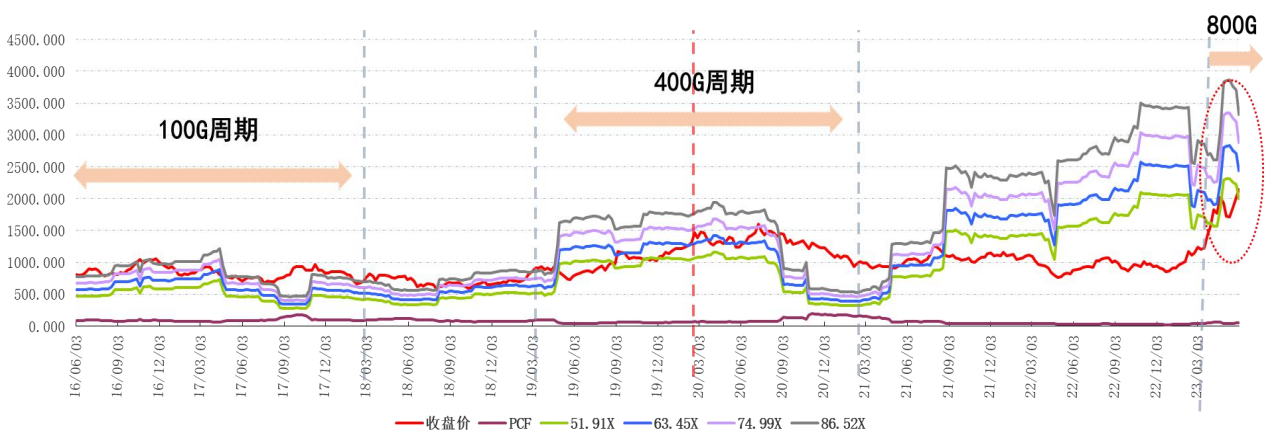
资料来源：源杰科技 8-1-2 发行人及保荐机构第二轮回复意见，源杰科技招股书，新易盛 2022 年年报，wind，山西证券研究所

### 3. 重点关注标的

#### 3.1 从估值端看当前光模块市场

历史复盘来看，在 100G，400G 周期期间，往往估值端先于业绩启动，主要逻辑在于预期带来估值拉升，而从下游给出订单指引到业绩反应在财务报表需要几个月的时间，因此业绩端兑现有所滞后。2023 年起 800G 新周期已在估值端反映，下游云厂商资本支出与订单指引不断提高，进一步拉高市场预期和景气度，光模块厂商凭借各自技术优势及海外客户有序拓展放量在即。

图 49：万得光模块板块 PE-Band



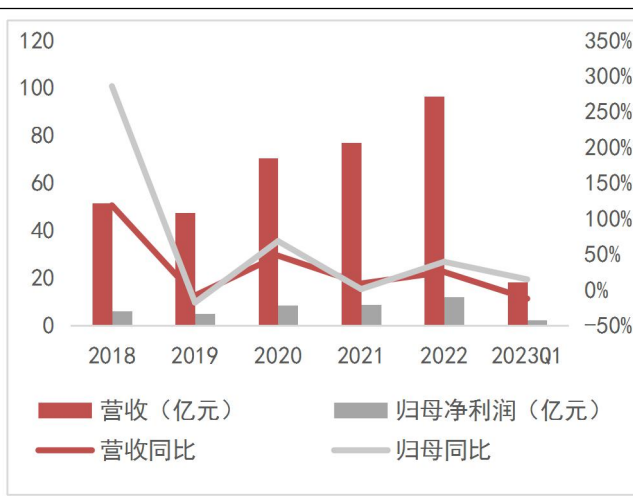
资料来源：wind，山西证券研究所

#### 3.2 主要公司建议关注

**中际旭创：国内光模块翘楚**，2022 年全球光模块市占与 IV-II 并列第一，光模块产品与技术全面布局且技术领先，下游与谷歌，亚马逊，中兴等多家云计算巨头合作紧密，上游持股国内光芯片龙头源杰科技。公司营收与业绩稳健增长，2020-2023Q1 营收分别 70.50/76.95/96.42/18.37 亿元，同比分别 48.17%/9.16%/25.29%/-12.04%，归母净利润分别 8.65/8.77/12.24/2.50 亿元，同比分别 68.55%/1.33%/39.57%/14.95%；按区域分，公司以海外为主，2020-2022 海外客户营收分别占比 69.31%/75.08%/86.88%。

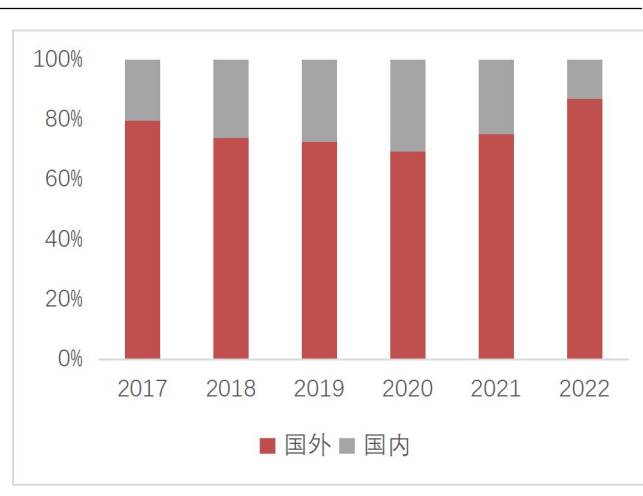
风险提示：客户合作不及预期；研发进展不及预期；下游需求不及预期等。

图 50：中际旭创营收与归母净利润



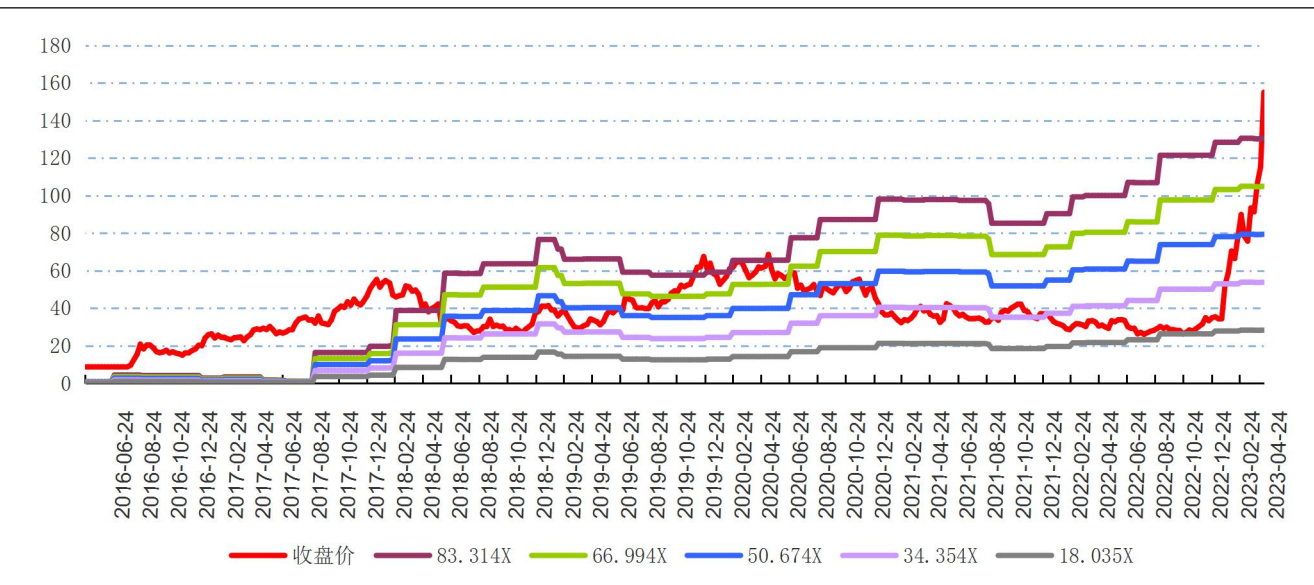
资料来源：wind，山西证券研究所

图 51：中际旭创营收按区域分



资料来源：wind，山西证券研究所

图 52：中际旭创 PE-Band

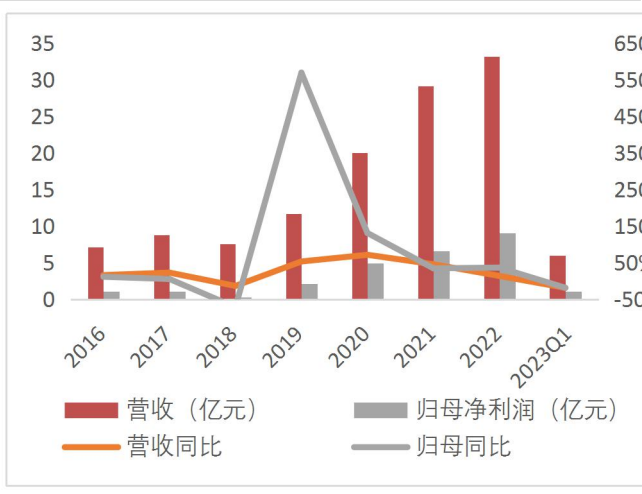


资料来源：wind，山西证券研究所

**新易盛：400G 光模块放量期间实现突破**，2022 年公司完成对 Alpine 的收购，同时拥有 CPO，LPO，硅光芯片和相干光技术。公司营收与业绩稳健增长，2020-2023Q1 营收分别 19.98/29.08/33.11/6.00 亿元，同比分别 71.52%/45.57%/13.83%/-18.73%，归母净利润分别 4.92/6.62/9.04/1.08 亿元，同比分别 131.03%/34.60%/36.51%/-18.57%；公司毛利率在同行业靠前，2020-2023Q1 毛利率分别 36.86%/32.17%/36.66%/34.04%。

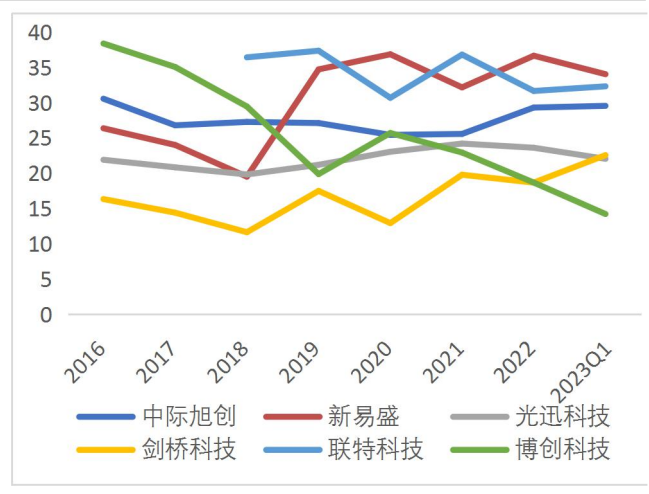
风险提示：客户合作不及预期；研发进展不及预期；下游需求不及预期等。

图 53：新易盛营收与归母净利润



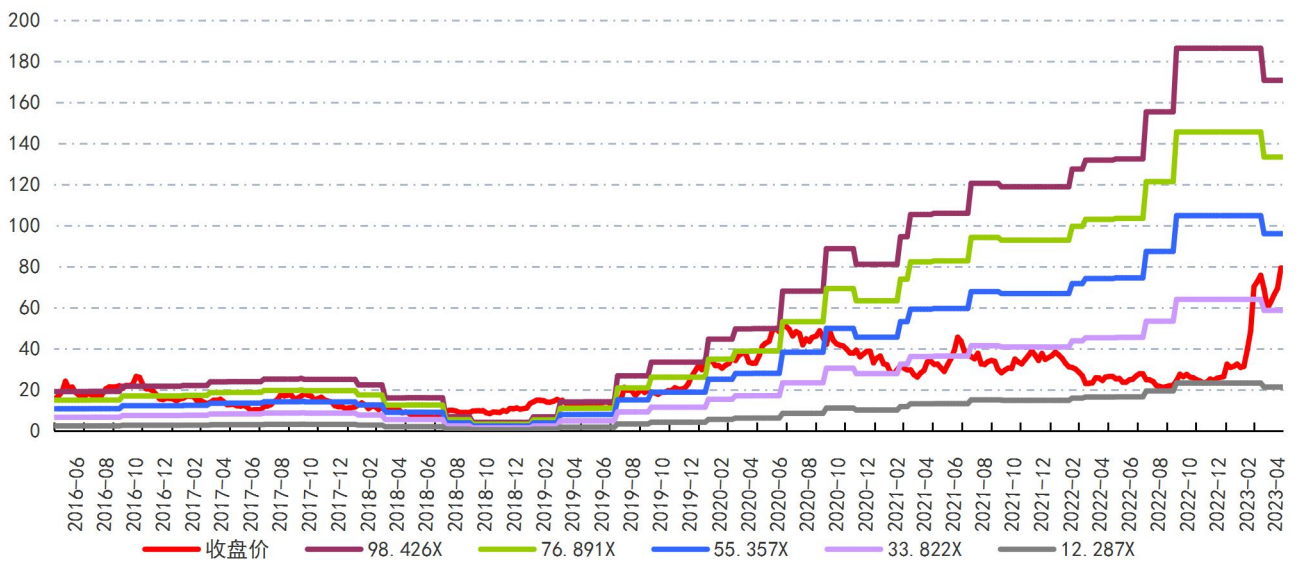
资料来源：wind，山西证券研究所

图 54：新易盛毛利率较高



资料来源：wind，山西证券研究所

图 55：新易盛 PE-Band

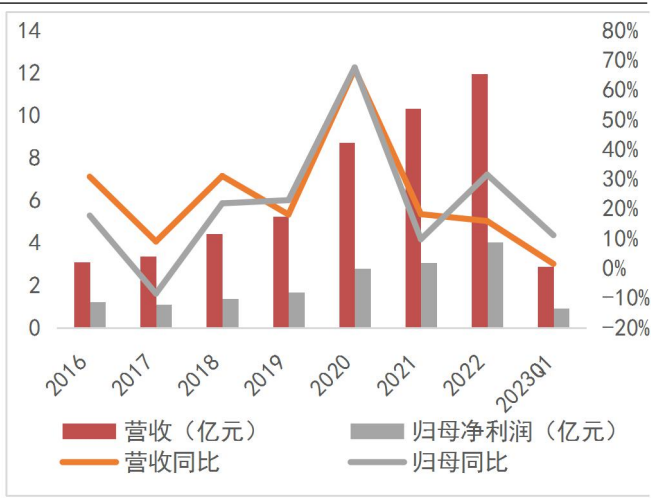


资料来源：wind，山西证券研究所

天孚通信：无源器件起家先发优势与海外客户建立合作，目前横向拓展至有源业务，光引擎产品在 CPO 方案下为刚需，在行业内进展领先。公司营收与业绩稳健增长，2020-2023Q1 营收分别 8.73/10.32/11.96/2.87 亿元，同比分别 67.03%/18.20%/15.89%/1.50%，归母净利润分别 2.79/3.06/4.03/0.92 亿元，同比分别 67.55%/9.77%/31.51%/11.14%；公司有源产品占比提高，2020-2022 有源产品占比分别 13.34%/8.23%/18.20%。

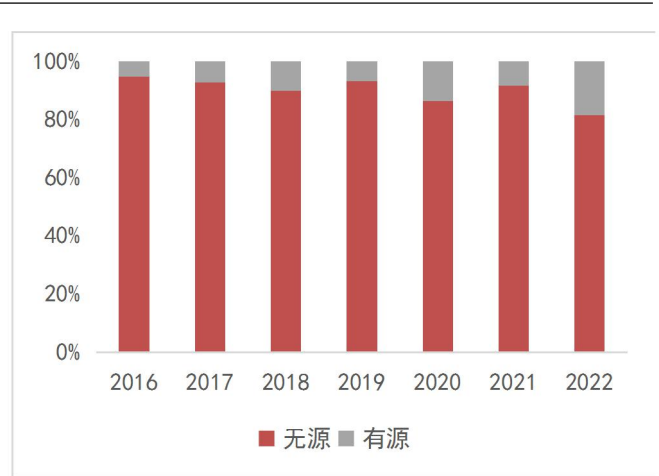
风险提示：客户合作不及预期；研发进展不及预期；下游需求不及预期等。

图 56：天孚通信营收与归母净利润



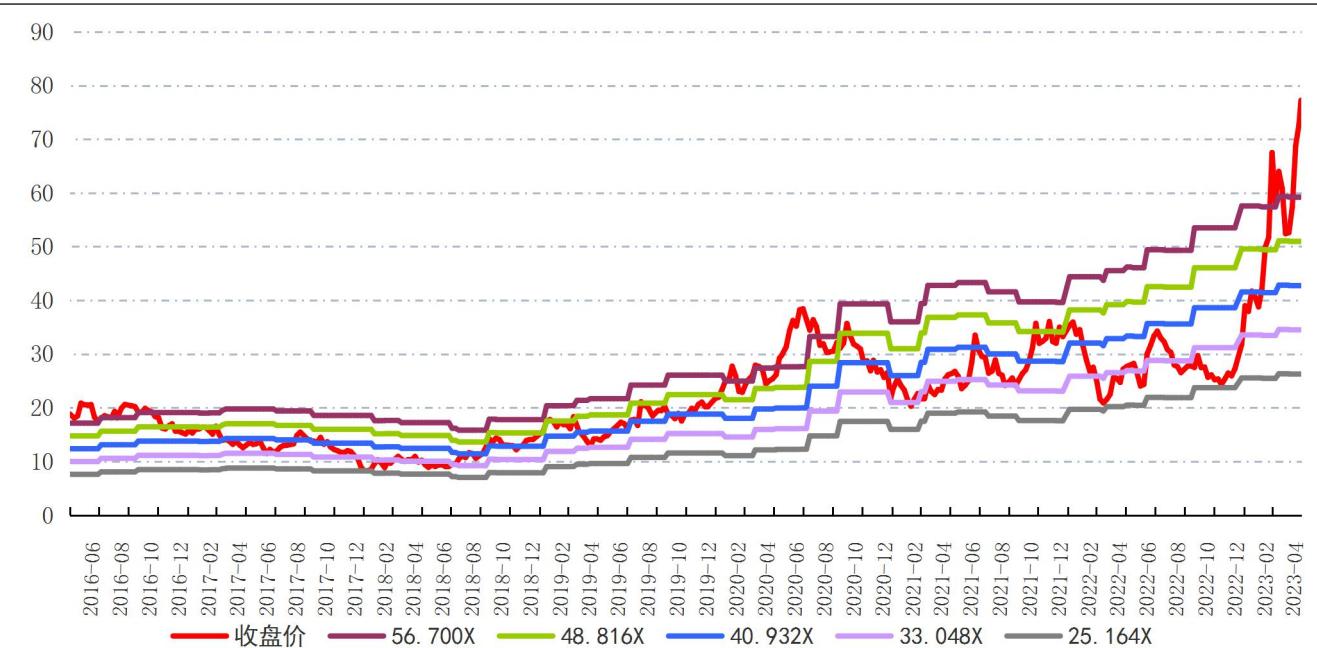
资料来源：wind，山西证券研究所

图 57：天孚通信营收分产品



资料来源：wind，山西证券研究所

图 58：天孚通信 PE-Band

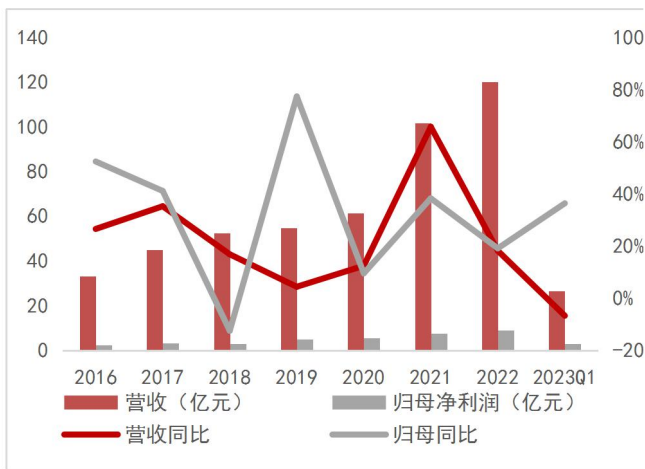


资料来源：wind，山西证券研究所

华工科技：背靠高校多项技术齐头并进，国内华为等云厂商主要供应，海外客户积极开拓。公司营收与业绩稳健增长，2020-2023Q1 营收分别 61.38/101.67/120.11/26.65 亿元，同比分别 12.40%/65.65%/18.14%/-6.76%，归母净利润分别 5.50/7.61/9.06/3.08 亿元，同比分别 9.49%/38.24%/19.07%/36.34；

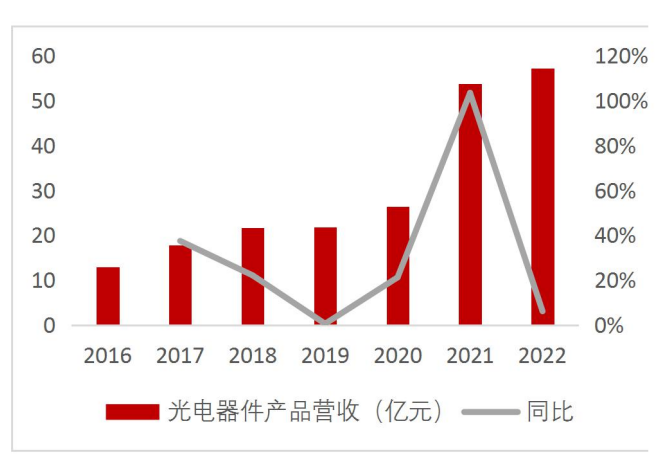
风险提示：客户合作不及预期；研发进展不及预期；下游需求不及预期等。

图 59：华工科技营收与归母净利润



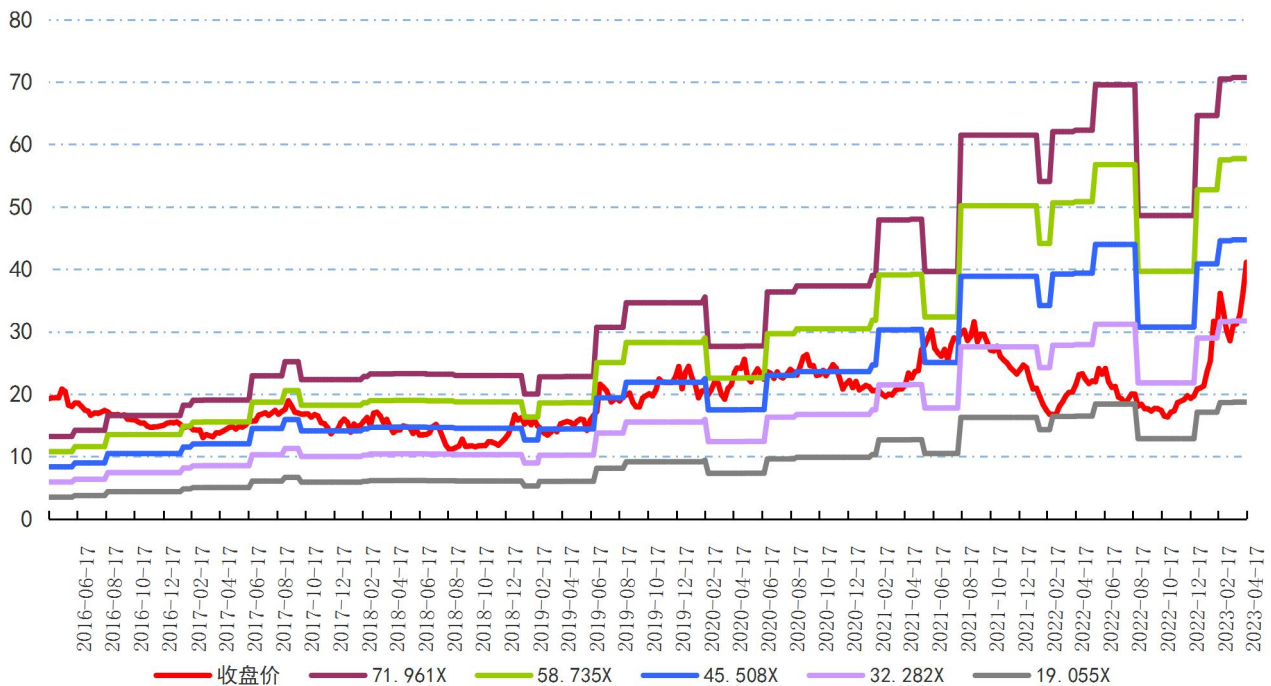
资料来源：wind，山西证券研究所

图 60：华工科技光电器件产品营收及同比



资料来源：wind，山西证券研究所

图 61：华工科技 PE-Band

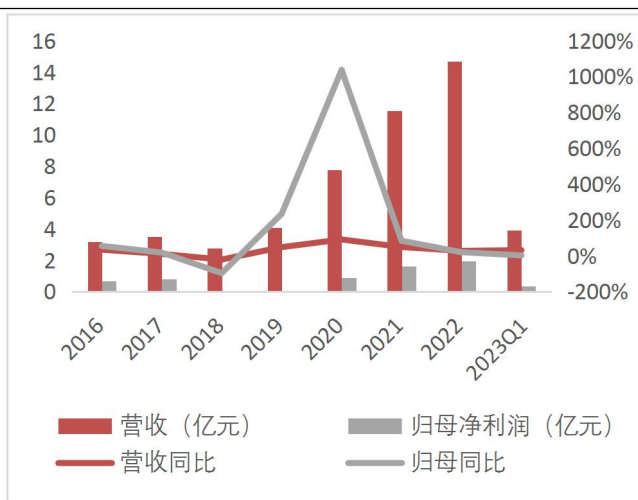


资料来源：wind，山西证券研究所

**博创科技：电信市场为主，硅光产品前瞻布局，助力数通市场加速发展。**公司营收与业绩稳健增长，2020-2023Q1 营收分别 7.77/11.54/14.67/3.93 亿元，同比分别 90.76%/48.59%/27.08%/31.11，归母净利润分别 0.88/1.62/1.94/0.37 亿元，同比分别 1036.48% /83.60% /19.59%/1.90%；

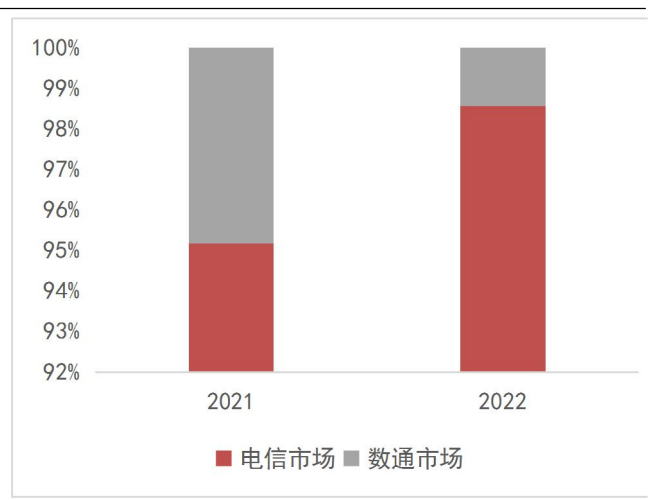
风险提示：客户合作不及预期；研发进展不及预期；下游需求不及预期等。

图 62：博创科技营收与归母净利润



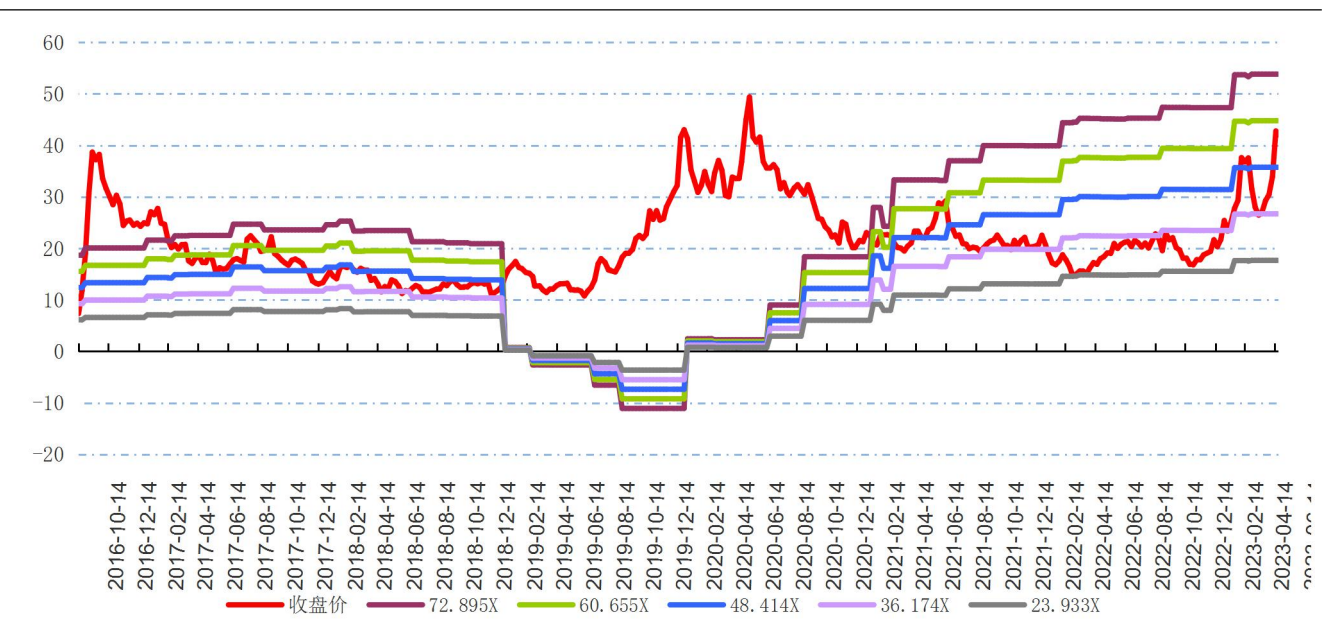
资料来源：wind，山西证券研究所

图 63：博创科技营收分市场



资料来源：wind，山西证券研究所

图 64：博创科技 PE-Band



资料来源：wind，山西证券研究所

此外，行业内其他公司凸显优势，建议关注光迅科技，联特科技，光库科技，剑桥科技，铭普光磁。

表 6：建议关注公司

	营收（百万元）				归母净利润（百万元）				PE			
	2022A	2023E	2024E	2025E	2022A	2023E	2024E	2025E	2022A	2023E	2024E	2025E
中际旭创	9,641.79	11,332.50	14,635.10	18,134.42	1,223.99	1,510.16	2,020.13	2,553.01	17.69	85.01	63.55	50.28
新易盛	3,310.57	3,757.75	5,023.41	6,485.71	903.58	906.43	1,258.84	1,657.56	13.33	55.17	39.72	30.17
天孚通信	1,196.39	1,609.36	2,168.33	2,908.10	402.94	505.51	652.70	853.01	24.79	84.10	65.14	49.84

华工科技	12,011.03	15,213.37	18,851.30	23,156.52	906.08	1,213.67	1,558.41	1,937.54	18.21	31.26	24.42	19.64
博创科技	1,466.72	1,934.42	2,478.95	3,068.36	194.23	254.28	318.59	385.02	22.79	39.48	31.51	26.08
光迅科技	6,911.88	7,616.83	8,587.33	9,907.80	608.41	687.00	801.67	923.20	18.04	46.76	40.07	34.80
联特科技	824.71	1,032.50	1,483.50	2,133.00	113.26	128.50	197.50	293.50	24.11	158.18	102.92	69.26
光库科技	642.44	863.60	1,177.60	1,543.60	117.80	143.20	199.40	259.00	52.66	116.31	83.53	64.31
剑桥科技	3,785.61	-	-	-	171.47	-	-	-	16.75	-	-	-
铭普光磁	2,323.40	-	-	-	68.88	-	-	-	38.60	-	-	-

资料来源：wind，山西证券研究所（数据截止 2023.06.21 收盘，采用 wind 一致预期）

## 4. 风险提示

**800G 需求不及预期：** 若 800G 整体需求不及预期则可能导致公司业绩不及预期。

**AI 产业发展不及预期：** 光模块行业主要增长点和预期来自于 AI 训练，若 AI 技术进步不及预期则将会影响光模块行业需求。

**公司海外客户合作及技术进步不及预期：** 光模块厂商需求目前以海外云厂商为主，若与海外云厂商合作不及预期或技术未能达到要求，则将会影响行业景气度。



### 分析师承诺：

本人已在中国证券业协会登记为证券分析师，本人承诺，以勤勉的职业态度，独立、客观地出具本报告。本报告清晰准确地反映本人的研究观点。本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点直接或间接受到任何形式的补偿。本人承诺不利用自己的身份、地位或执业过程中所掌握的信息为自己或他人谋取私利。

### 投资评级的说明：

以报告发布日后的 6--12 个月内公司股价（或行业指数）相对同期基准指数的涨跌幅为基准。其中：A 股以沪深 300 指数为基准；新三板以三板成指或三板做市指数为基准；港股以恒生指数为基准；美股以纳斯达克综合指数或标普 500 指数为基准。

无评级：因无法获取必要的资料，或者公司面临无法预见的结果的重大不确定事件，或者其他原因，致使无法给出明确的投资评级。（新股覆盖、新三板覆盖报告及转债报告默认无评级）

### 评级体系：

#### ——公司评级

- 买入： 预计涨幅领先相对基准指数 15%以上；
- 增持： 预计涨幅领先相对基准指数介于 5%-15%之间；
- 中性： 预计涨幅领先相对基准指数介于-5%-5%之间；
- 减持： 预计涨幅落后相对基准指数介于-5%- -15%之间；
- 卖出： 预计涨幅落后相对基准指数-15%以上。

#### ——行业评级

- 领先大市： 预计涨幅超越相对基准指数 10%以上；
- 同步大市： 预计涨幅相对基准指数介于-10%-10%之间；
- 落后大市： 预计涨幅落后相对基准指数-10%以上。

#### ——风险评级

- A： 预计波动率小于等于相对基准指数；
- B： 预计波动率大于相对基准指数。

### 免责声明：

山西证券股份有限公司(以下简称“公司”)具备证券投资咨询业务资格。本报告是基于公司认为可靠的已公开信息，但公司不保证该等信息的准确性和完整性。入市有风险，投资需谨慎。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议。在任何情况下，公司不对任何人因使用本报告中的任何内容引致的损失负任何责任。本报告所载的资料、意见及推测仅反映发布当日的判断。在不同时期，公司可发出与本报告所载资料、意见及推测不一致的报告。公司或其关联机构在法律许可的情况下可能持有或交易本报告中提到的上市公司发行的证券或投资标的，还可能为或争取为这些公司提供投资银行或财务顾问服务。客户应当考虑到公司可能存在可能影响本报告客观性的利益冲突。公司在知晓范围内履行披露义务。本报告版权归公司所有。公司对本报告保留一切权利。未经公司事先书面授权，本报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯公司版权的其他方式使用。否则，公司将保留随时追究其法律责任的权利。

依据《发布证券研究报告执业规范》规定特此声明，禁止公司员工将公司证券研究报告私自提供给未经公司授权的任何媒体或机构；禁止任何媒体或机构未经授权私自刊载或转发公司证券研究报告。刊载或转发公司证券研究报告的授权必须通过签署协议约定，且明确由被授权机构承担相关刊载或者转发责任。

依据《发布证券研究报告执业规范》规定特此提示公司证券研究业务客户不得将公司证券研究报告转发给他人，提示公司证券研究业务客户及公众投资者慎重使用公众媒体刊载的证券研究报告。

依据《证券期货经营机构及其工作人员廉洁从业规定》和《证券经营机构及其工作人员廉洁从业实施细则》规定特此告知公司证券研究业务客户遵守廉洁从业规定。

### 山西证券研究所：

#### 上海

上海市浦东新区滨江大道 5159 号陆家嘴滨江中心 N5 座 3 楼

#### 太原

太原市府西街 69 号国贸中心 A 座 28 层  
电话：0351-8686981  
<http://www.i618.com.cn>

#### 深圳

广东省深圳市福田区林创路新一代产业园 5 栋 17 层

#### 北京

北京市西城区平安里西大街 28 号中海国际中心七层  
电话：010-83496336

