

证券研究报告

2024年06月30日

行业报告：行业投资策略

通信

AI算力激涌，光通信乘风而上，通信+踏浪前行

作者：

分析师 唐海清 SAC执业证书编号：S1110517030002

分析师 王奕红 SAC执业证书编号：S1110517090004

分析师 康志毅 SAC执业证书编号：S1110522120002

分析师 林竑皓 SAC执业证书编号：S1110520040001

分析师 余芳沁 SAC执业证书编号：S1110521080006

请务必阅读正文之后的信息披露和免责声明



行业评级：强于大市（维持评级）  
上次评级：强于大市

随着22年11月OpenAI公司发布ChatGPT，海内外AI模型军备竞赛打响，全球对算力的需求激增，需要通信互联产业快速发展与之相匹配。持续重视AI核心标的的投资机会：海外线AI核心方向如光模块&光器件、液冷等领域值得重视，持续核心推荐；此外国产算力线如国产服务器，交换机，AIDC等方向核心标的建议积极关注。

**风险提示：1. AI应用发展不及预期风险：AI应用持续迭代，存在技术发展不及预期的可能性，导致整体行业发展进度放缓。2. 海外大厂投资不及预期风险：海外厂商资本开支不及预期，影响下游供应链需求。3. 中美摩擦升级风险：中美科技、金融、贸易等领域摩擦加剧。4. 行业竞争加剧，价格和盈利能力下降风险：进入行业的厂商增加，价格竞争使得盈利能力下降。5. 公司业绩不及预期：相关产品研发、推广以及市场拓展不及预期。**

### 重点标的推荐

| 股票<br>代码  | 股票<br>名称 | 收盘价<br>2024-06-28 | 投资<br>评级 | EPS(元) |       |       |       | P/E    |        |       |       |
|-----------|----------|-------------------|----------|--------|-------|-------|-------|--------|--------|-------|-------|
|           |          |                   |          | 2023A  | 2024E | 2025E | 2026E | 2023A  | 2024E  | 2025E | 2026E |
| 300308.SZ | 中际旭创     | 137.88            | 买入       | 2.71   | 6.75  | 9.68  | 11.42 | 50.88  | 20.43  | 14.24 | 12.07 |
| 300394.SZ | 天孚通信     | 88.42             | 增持       | 1.85   | 3.89  | 6.21  | 8.59  | 47.79  | 22.73  | 14.24 | 10.29 |
| 300502.SZ | 新易盛      | 105.55            | 买入       | 0.97   | 2.27  | 3.73  | 4.99  | 108.81 | 46.50  | 28.30 | 21.15 |
| 002463.SZ | 沪电股份     | 36.50             | 增持       | 0.79   | 1.24  | 1.66  | 2.06  | 46.20  | 29.44  | 21.99 | 17.72 |
| 600522.SH | 中天科技     | 15.85             | 买入       | 0.91   | 1.11  | 1.47  | 1.78  | 17.42  | 14.28  | 10.78 | 8.90  |
| 600487.SH | 亨通光电     | 15.77             | 买入       | 0.87   | 1.12  | 1.42  | 1.73  | 18.13  | 14.08  | 11.11 | 9.12  |
| 688498.SH | 源杰科技     | 130.99            | 增持       | 0.23   | 1.23  | 2.17  | 3.36  | 569.52 | 106.50 | 60.36 | 38.99 |
| 000063.SZ | 中兴通讯     | 27.97             | 买入       | 1.95   | 2.15  | 2.38  | 2.62  | 14.34  | 13.01  | 11.75 | 10.68 |
| 600941.SH | 中国移动     | 107.50            | 买入       | 6.16   | 6.66  | 7.15  | 7.68  | 17.45  | 16.14  | 15.03 | 14.00 |
| 300442.SZ | 润泽科技     | 23.95             | 买入       | 1.02   | 1.28  | 1.92  | 2.44  | 23.48  | 18.71  | 12.47 | 9.82  |
| 300638.SZ | 广和通      | 17.10             | 买入       | 0.74   | 0.96  | 1.22  | 1.51  | 23.11  | 17.81  | 14.02 | 11.32 |
| 002139.SZ | 拓邦股份     | 10.62             | 买入       | 0.41   | 0.57  | 0.71  | 0.88  | 25.90  | 18.63  | 14.96 | 12.07 |
| 688100.SH | 威胜信息     | 38.15             | 买入       | 1.05   | 1.31  | 1.68  | 2.12  | 36.33  | 29.12  | 22.71 | 18.00 |
| 300628.SZ | 亿联网络     | 36.77             | 买入       | 1.59   | 1.91  | 2.23  | 2.63  | 23.13  | 19.25  | 16.49 | 13.98 |
| 300667.SZ | 必创科技     | 13.53             | 增持       | 0.18   | 0.31  | 0.44  | 0.60  | 75.17  | 43.65  | 30.75 | 22.55 |
| 300627.SZ | 华测导航     | 29.85             | 增持       | 0.82   | 1.03  | 1.23  | 1.44  | 36.40  | 28.98  | 24.27 | 20.73 |

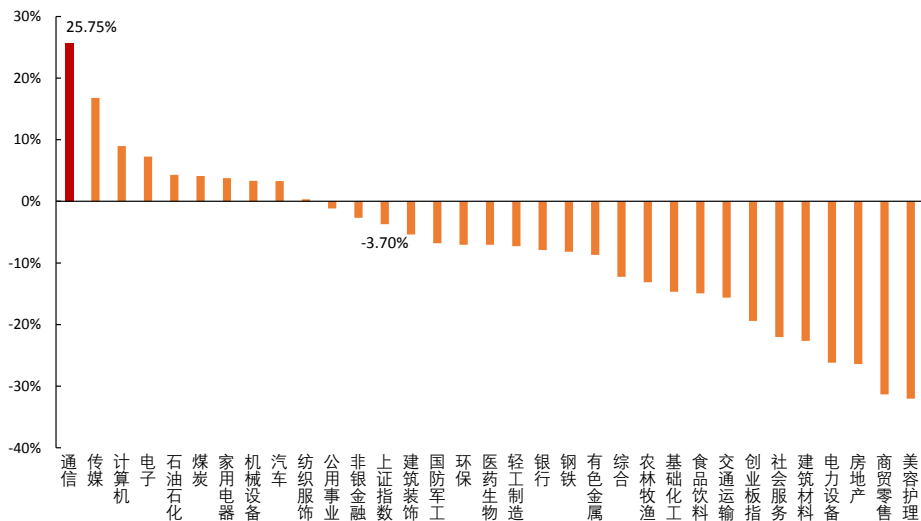


# 2023年通信行业行情回顾

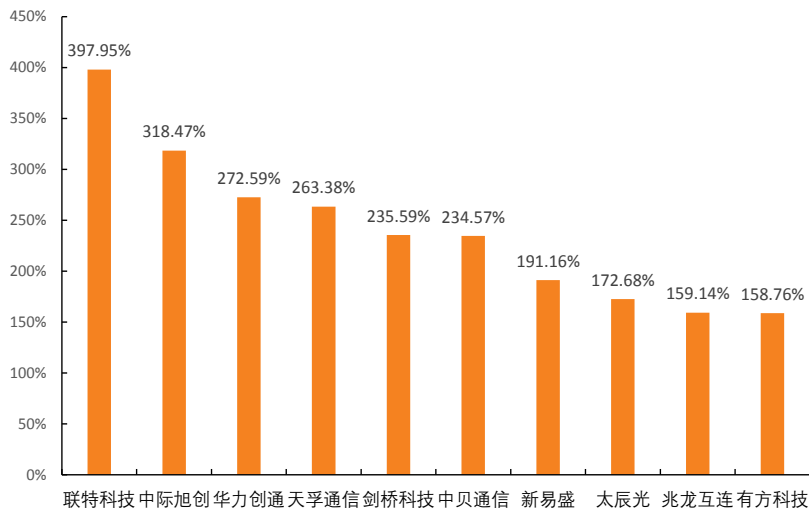
# 2023年通信行业行情回顾：通信指数大幅跑赢，AIGC产业大变革大机遇

- ▶ 截至2023年12月29日，通信（申万）指数大幅跑赢上证指数、创业板指数，在31个行业板块中位列第一名。
- ▶ 从细分方向来看，2023年涨幅表现领先标的联特科技、中际旭创、天孚通信、剑桥科技均为【光模块】相关，中贝通信为【算力租赁】板块，华力创通【卫星互联网】板块，兆龙互联为【服务器线缆】板块。整体来看，AI带动了通信板块2023年核心受益标的强势表现。

图：2023年各行业涨跌幅（截止2023年12月29日）



图：2023年相关个股涨幅排名

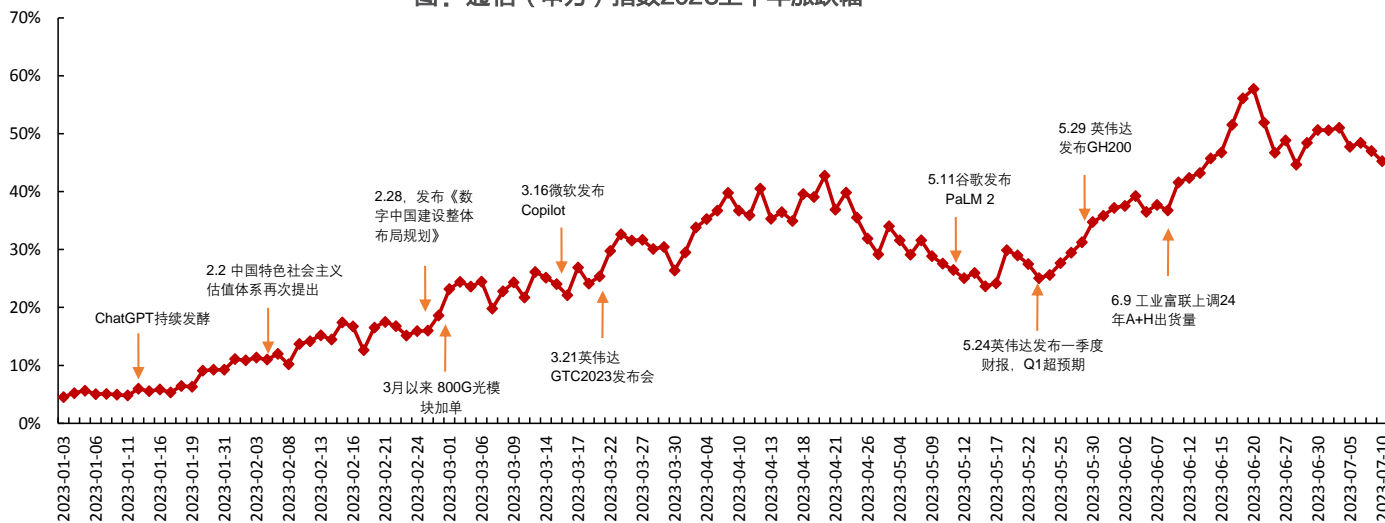


# 通信行业行情回顾：上半年受光模块加单及芯片产能催化

## ► 上半年复盘十大事件

- 1) 2022年11月30日，ChatGPT发布，随后不断发酵，春节GPT海外发酵，春节回来加速；
- 2) 2月2日，中国证监会召开的2023年系统工作会议上，易主席再次提及中国特色估值体系；
- 3) 2月28日，中共中央、国务院印发《数字中国建设整体布局规划》，明确提到数字中国建设按照“2522”的整体框架进行布局；
- 4) 3月以来北美厂商多次追加800G光模块订单；
- 5) 3月16日微软Power Platform宣布在Power Apps、Power Automate和Power Virtual Agents中加入Copilot；
- 6) 3月21日GTC2023发布会，英伟达发布NVIDIA DGX Cloud等产品；
- 7) 5月11日谷歌发布基础大语言模型PaLM 2；
- 8) 5月24日英伟达一季报超预期；
- 9) 5月29日英伟达推出GH200，对应800G光模块用量增长；
- 10) 6月9日工业富联上修24年A+H卡出货量。

图：通信（申万）指数2023上半年涨跌幅

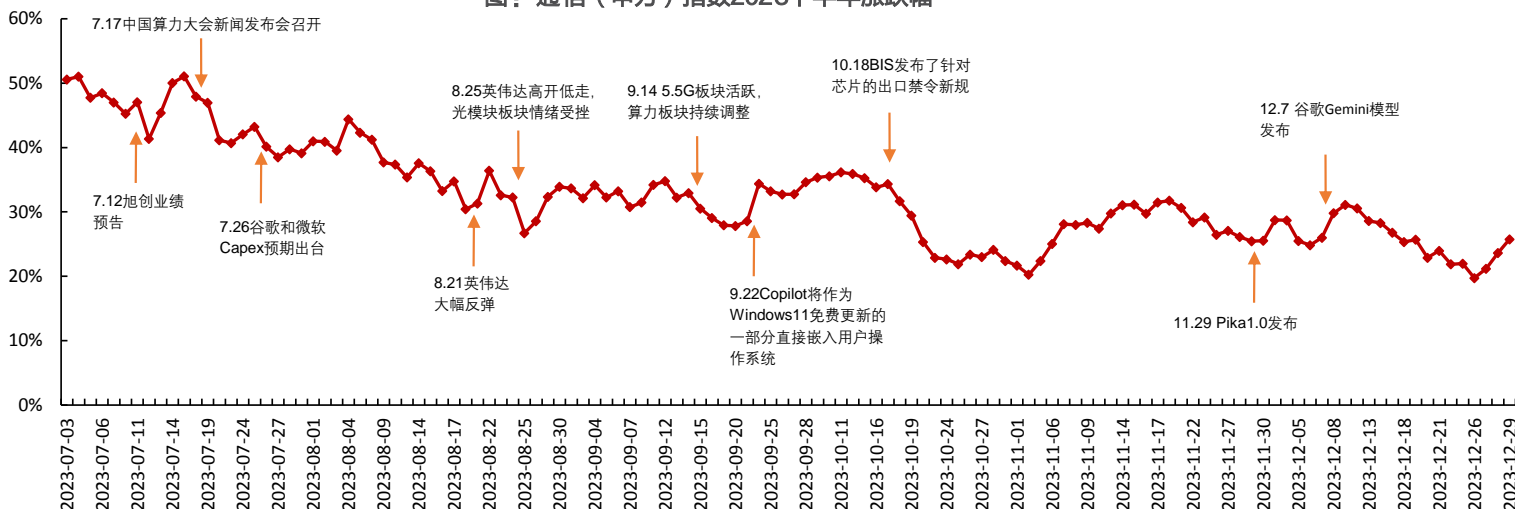


# 通信行业行情回顾：下半年受筹码结构及边际催化减弱影响，板块调整

## 下半年复盘十大事件

- 1) 7月12日，中际旭创业绩预告，业绩高增长；
- 2) 7月17日，2023中国算力大会新闻发布会召开，工信部将从①政策接续，强化顶层设计、②加强技术创新，培育良好生态（加强CPU、GPU和服务器等重点产品研发，软硬件生态体系建设）、③建优算力网络，促进应用落地，三个方面推进我国算力发展；
- 3) 7月26日，谷歌和微软Capex预期出台；
- 4) 8月21日，英伟达大幅反弹，随后8月24日英伟达公布业绩；
- 5) 8月25日，英伟达公布业绩后股价高开低走，光模块板块情绪受挫；
- 6) 9月14日，5.5G板块活跃，华为推动5G-A演进，带动相关产业链包括灿勤、大富、凡谷等上涨，AI算力相关标的持续调整；
- 7) 9月22日，节前反弹，Copilot将作为Windows11免费更新的一部分直接嵌入用户操作系统；
- 8) 10月18日，美国商务部工业和安全局（BIS）发布了针对芯片的出口禁令新规，国内算力产业链受较大情绪影响；
- 9) 11月29日，Pika Labs发布AI视频生成器Pika1.0版本；
- 10) 12月7日，谷歌Gemini模型发布，AMD上调加速器市场规模预测。

图：通信（申万）指数2023下半年涨跌幅



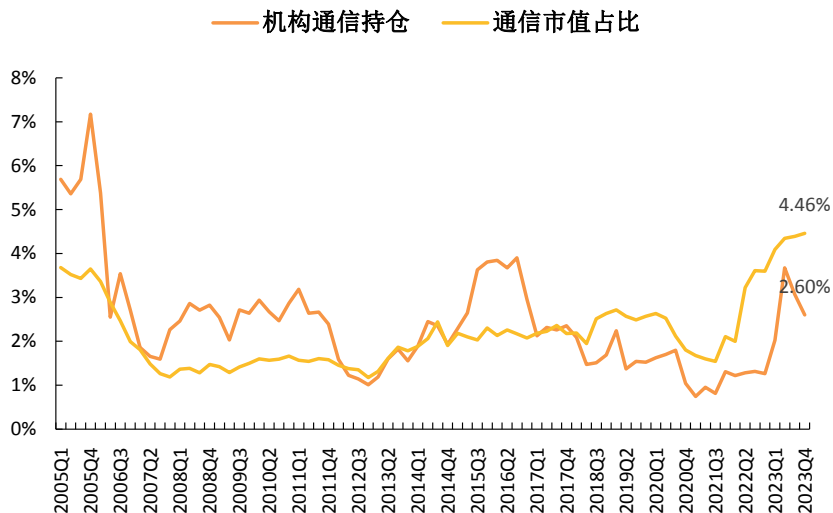
## 2023年通信行业行情回顾：机构持仓中等/历史估值低位

- 从机构持仓比例来看：通信板块基金持仓仍属于低配，截止23年四季度末比例约2.60%。
- 从通信估值水平来看：通信板块估值处于历史估值相对低位，截止23年四季度末约11.18%分位。

图：通信（申万）指数2023下半年涨跌幅



图：通信（申万）指数2023下半年涨跌幅





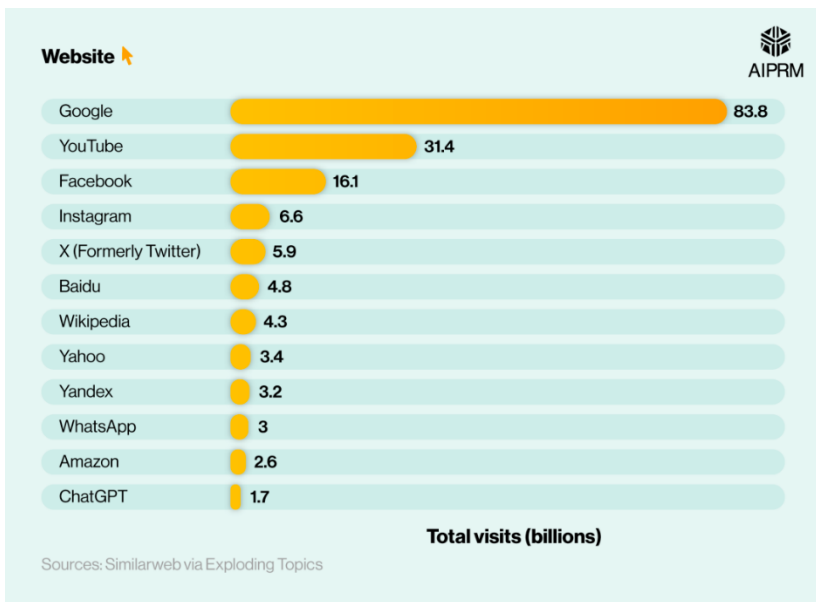
# 1. AI算力：光模块、光芯片



# ChatGPT掀起AI浪潮，打响全球AI军备竞赛

- 2022年11月30日OpenAI公司发布聊天机器人ChatGPT，仅用时5天用户人数已经突破100万人，截至2023年12月，ChatGPT拥有约1.8亿用户。ChatGPT平均每月产生17亿次网站浏览量。
- 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，随后阿里、商汤、360、华为等企业纷纷发布自己的大模型。

图：ChatGPT平均每月访问量17亿次



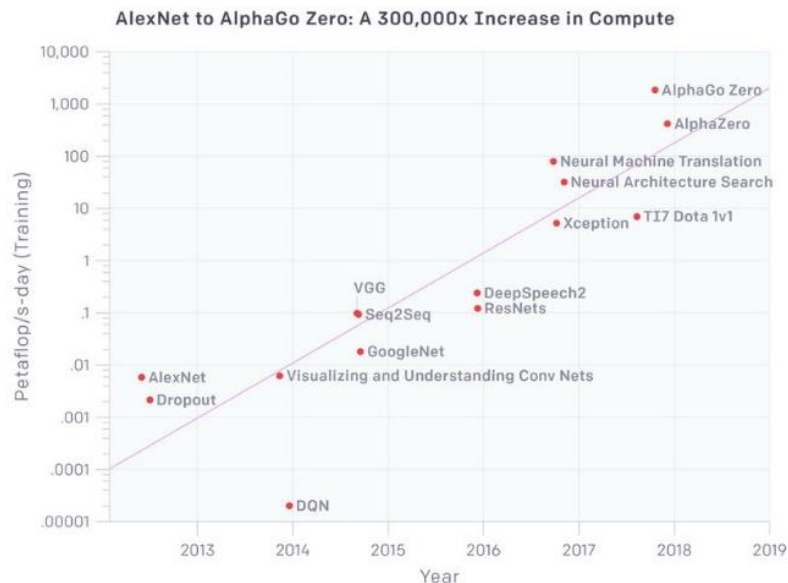
图：2023年海内外AI厂商相继发布大模型

| 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                      | Pika1.0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大且最有能力的模型；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。   |

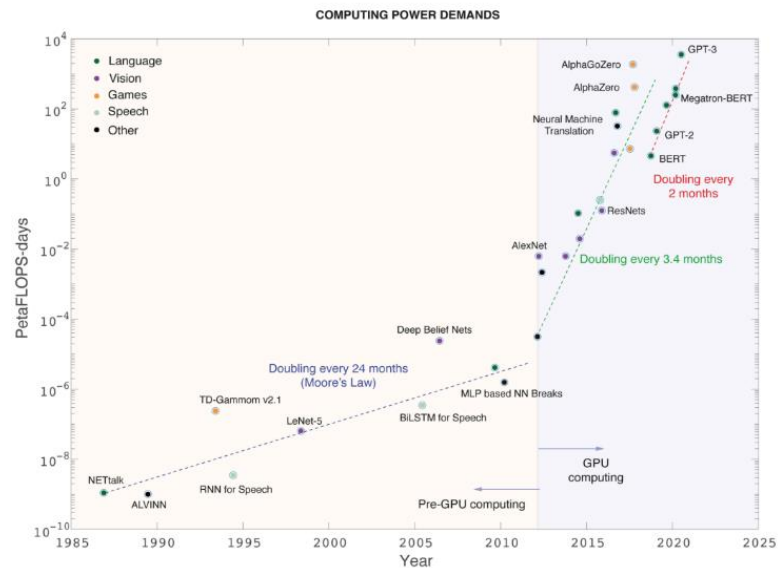
## 算力激增需要互联匹配，带动光模块需求增长

- GPT-1的模型参数量和预训练数据量分别为1.17亿个和5GB，到GPT-3已激增至1750亿个和45TB，且预计未来GTP-5的参数量将是GTP-3的100倍，需要的计算量则是GTP-3的200~400倍，随着模型不断升级迭代，所需要的数据量也在快速呈倍数增长。根据OpenAI发布的《AI and Compute》分析报告中指出，自2012年以来，AI训练应用的算力需求每3-4个月就会翻倍，且从2012年至今，AI算力增长超过了30万倍。

图：AlexNet至AlphaGo Zero6年时间内算力需求翻30万倍



图：2020年后算力需求约每两个月翻一番



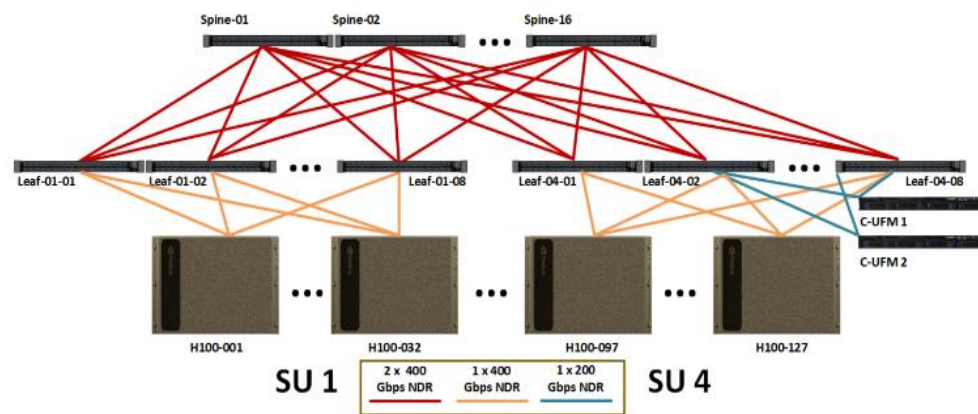
## 算力激增需要互联匹配，带动光模块需求增长

- 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。
- 2022年3月NVIDIA推出DGX H100，带动400G、800G光模块配套需求。由于组网方式不同，GPU与光模块用量配比关系较难精确测算，我们以NVIDIA DGX H100服务器集群举例测算：1) 在计算网络中，每单片H100对应1.5个800G光模块和0.9个400G光模块；2) 在存储网络中，每单片H100对应1.25个400G光模块。假设只考虑以上两种网络场景中对光模块的需求，则每单片H100对应1.5个800G光模块和2.15个400G光模块需求，每单片H100对应大约20个100G光芯片的需求。

图：不同组网方式下GPU与光模块用量配比

| 组网方式 | 芯片型号 | 网卡型号       | 交换机型号             | 网络架构   | 光模块速率     | 用量配比                              |
|------|------|------------|-------------------|--------|-----------|-----------------------------------|
| A    | A100 | ConnectX-6 | QM8700 Infiniband | 三层拓扑结构 | 200G      | GPU与200G比例为1:6                    |
| B    | A100 | ConnectX-6 | QM9700 Infiniband | 叶脊结构   | 200G、800G | GPU与200G比例为1:1, GPU与800G比例为1:0.75 |
| C    | H100 | ConnectX-7 | QM9700 Infiniband | 叶脊结构   | 400G、800G | GPU与400G比例为1:1, 和800G比例为1:1.5     |
| D    | H100 | ConnectX-8 | QM9700 Infiniband | 三层拓扑结构 | 800G      | GPU与800G比例为1:6                    |

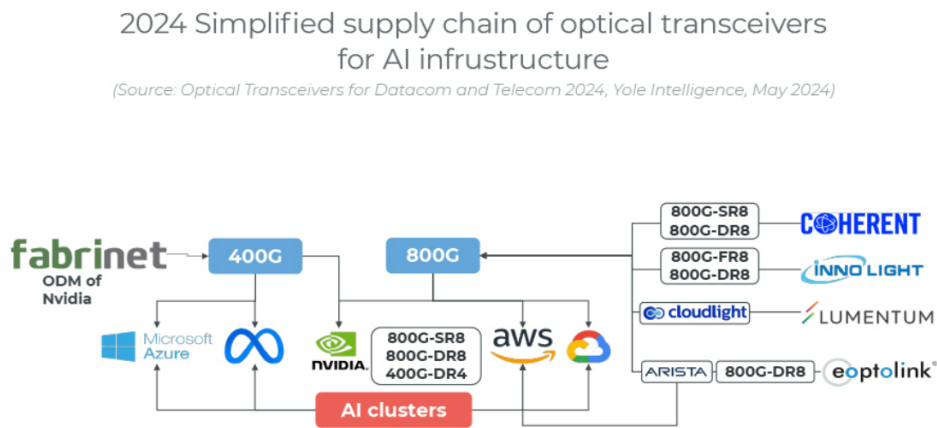
图：计算网络环境中NVIDIA DGX H100集群的网络架构



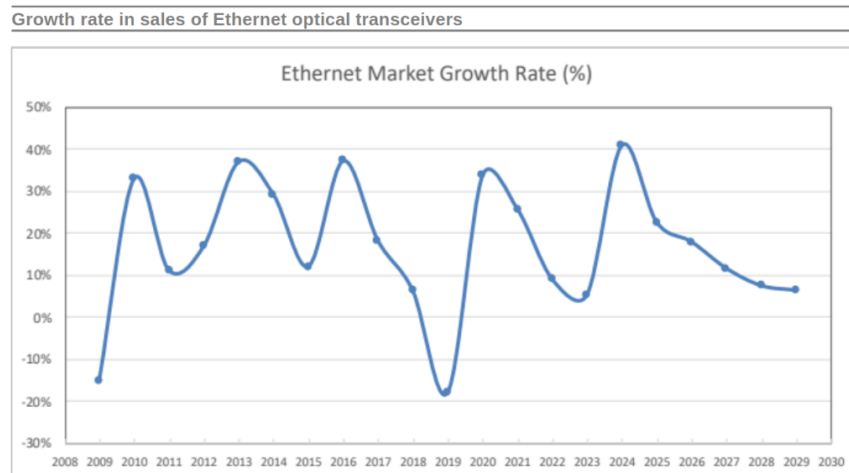
# 400G、800G海外需求23年起量，24年进一步释放；1.6T有望于24Q4落地，25年逐步上量

- 受AI驱动，2023年800G、400G光模块海外需求逐步起量。2023年3月开始，在谷歌、亚马逊和英伟达等海外大客户的推动下，市场对800G光模块的需求激增，导致订单和出货量大幅增长；2023年早些时候，微软和Meta也增加了对400G光模块的需求，从而反映市场受AI驱动正不断扩大。
- 2024年400G、800G需求将进一步释放。根据LightCounting预测，光模块在以太网中的使用量2024年增长40%，2025年增幅在20%以上，2026-2027年将实现两位数的增长。
- 1.6T有望于2024Q4落地，25年逐步上量。4月30日中际旭创在业绩说明会上表示1.6T光模块研发的市场导入正在加快，预计1.6T产品会有2024年末到2025年期间逐步上量。

图：2024年AI基建光模块供应链



图：2024年光模块在以太网中的使用量2024年增长40%



# 全球光模块行业格局呈“西退东进”趋势，中际旭创23年Top10排名第一

- ▶ 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，到2018年，大部分日本和美国厂商退出了这一市场，而以中际旭创为首的中国厂商的排名不断提升。根据LightCounting最新排名，Top10中中国厂商有7家入围：中际旭创（排名第1）、华为（排名第3）、光迅科技（排名第5）、海信宽带（排名第6）、新易盛（排名第7）、华工正源（排名第8）、索尔思光电（排名第9）。国外厂商共计3家入围：Coherent/Finisar（排名第2）、思科/Acacia（排名第4）、Marvell（排名第10）。
- ▶ 根据海关总署出口数据，5月光模块出口货值实现40.98亿元，同比+124.63%，环比+6.41%，3~5月光模块出口货值连续3月创下新高。

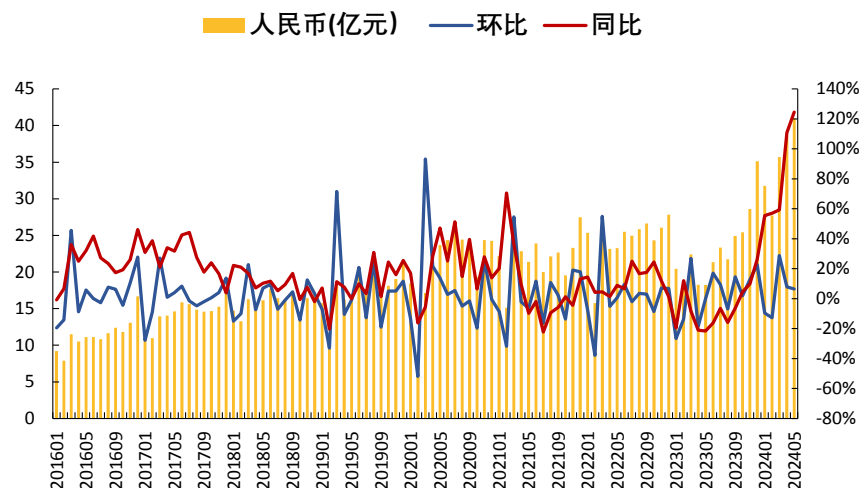
图：2023Top10光模块厂商中国企业入围7家

Figure 1: Top 10 Transceiver Vendors in 2010, 2016, 2018 and 2023

| Ranking of Top 10 Transceiver Suppliers |                  |    |                 |                    |
|---|------------------|----|-----------------|--------------------|
| 2010                                    | 2016             |    | 2018            | 2023               |
| Finisar                                 | Finisar          | 1  | Finisar         | Innolight          |
| Opnext                                  | Hisense          | 2  | Innolight       | Coherent           |
| Sumitomo                                | Accelink         | 3  | Hisense         | Huawei (HiSilicon) |
| Avago                                   | Acacia           | 4  | Accelink        | Cisco (Acacia)     |
| Source Photonics                        | FOIT (Avago)     | 5  | FOIT (Avago)    | Accelink           |
| Fujitsu                                 | Oclaro           | 6  | Lumentum/Oclaro | Hisense            |
| JDSU                                    | Innolight        | 7  | Acacia          | Eoptolink          |
| Emcore                                  | Sumitomo         | 8  | Intel           | HGGenuine          |
| WTD                                     | Lumentum         | 9  | AOI             | Source Photonics   |
| NeoPhotonics                            | Source Photonics | 10 | Sumitomo        | Marvell            |

Source: LightCounting

图：光模块出口海关数据



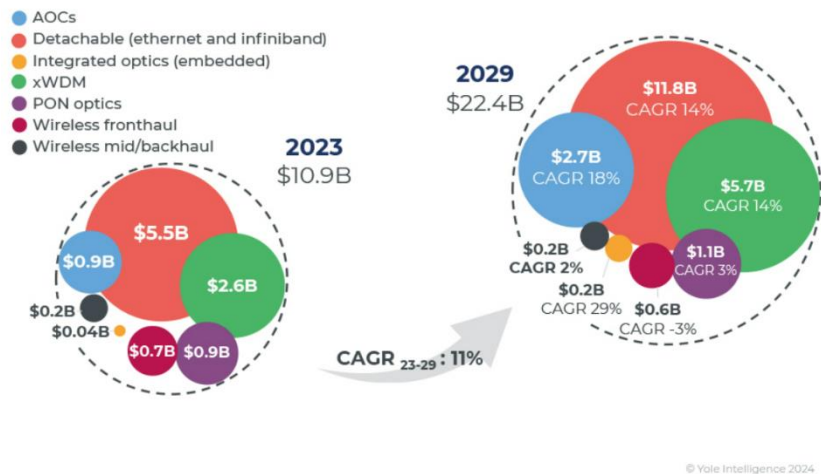
# Yole Intelligence: 23~29年全球光模块市场年化复合增速预计将达到11%

- 2023-2029年全球光模块市场规模CAGR将实现11%，2029年有望达到224亿美元。根据Yole Intelligence数据，2023年全球光模块市场规模为109亿美元，预计2029年将达到224亿美元，源于云服务运营商和国家电信运营商对400G以上高数据速率光模块的高需求。2024年的收入增长率有望实现27%，原因为英伟达的大量AI基础设施订单和数据中心网络升级到800G。
- 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：2023-2029年全球光模块市场CAGR预计将达到11%

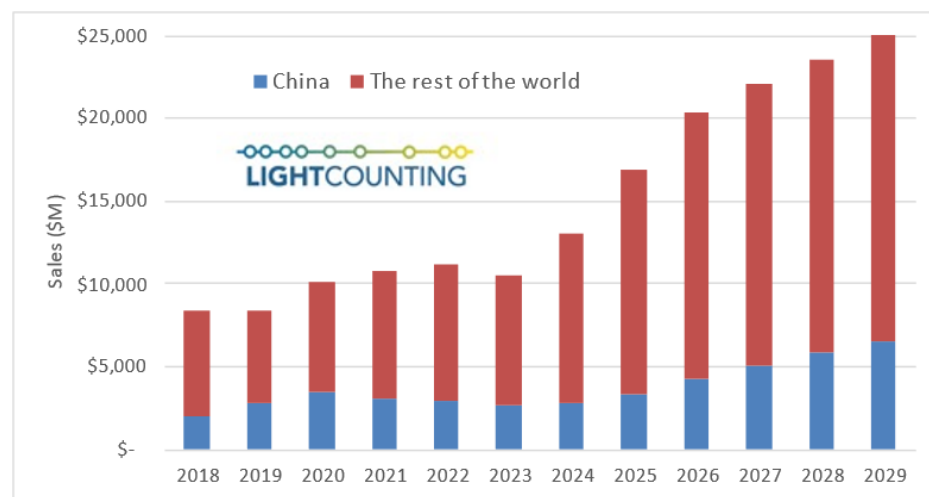
2023-2029 optical transceiver revenue growth forecast by segment

(Source: Optical Transceivers for Datacom and Telecom 2024, Yole Intelligence, May 2024)



图：2024-2029年中国光模块部署量占比全球20%-25%

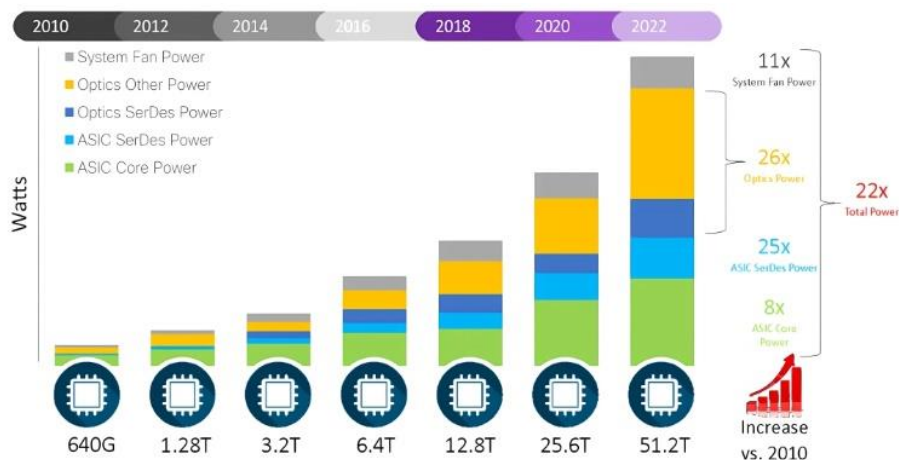
Figure: Global Market for Optical Transceivers



# 当下光模块技术向低功耗、高带宽方向演进

- **算力时代对光模块有着低功耗、高带宽的需求：**2021年我国数据中心耗电量为2166亿千瓦时，约为三峡电站同期年发电量1036.49亿千瓦时的2倍；2022年，我国数据中心耗电量达到2700亿千瓦时，占全社会用电量约3%；根据中国能源报统计，预计2025年该比重将接近5%。目前光模块行业技术演进方向包括：
  - **1) LPO方案。**LPO采用线性驱动技术代替传统DSP（数字信号处理）/CDR（时钟数据回复），可实现降功耗、压成本的作用，但代价在于拿掉DSP后会导导致系统误码率提升，通信距离缩短，因此LPO技术只适用于短距离的应用场景，例如数据中心机柜到交换机的连接等。
  - **2) CPO方案。**CPO是指将光引擎和开关芯片封装在一起的电光集成，而不是使用可插拔的光模块格式。这种方法通过缩短光学引擎和开关芯片之间的距离，可以在引擎和芯片之间更快地传输电信号。其优势在于尺寸小、功耗低、效率高。
  - **3) 硅光技术。**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技术。硅光子技术的核心理念是“以光代电”，即采用激光束代替电子信号传输数据，将光学器件与电子元件整合至一个独立的微芯片中，提升芯片之间的连接速度。

图：2010-2022年光引擎功耗提升26倍



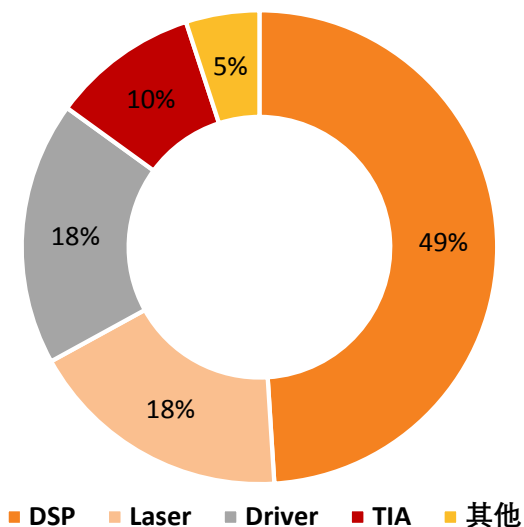
图：传统光模块、CPO与LPO优劣势对比

|        | 光模块 (DSP) | CPO | LPO |
|--------|-----------|-----|-----|
| 功耗     | 高         | 低   | 低   |
| 时延     | 高         | 低   | 低   |
| 链路性能   | 优         | 中   | 差   |
| 灵活性    | 高         | 低   | 中   |
| 可维护性   | 优         | 差   | 优   |
| 产品成熟度  | 高         | 低   | 低   |
| 异厂商互操作 | 高         | 中   | 低   |

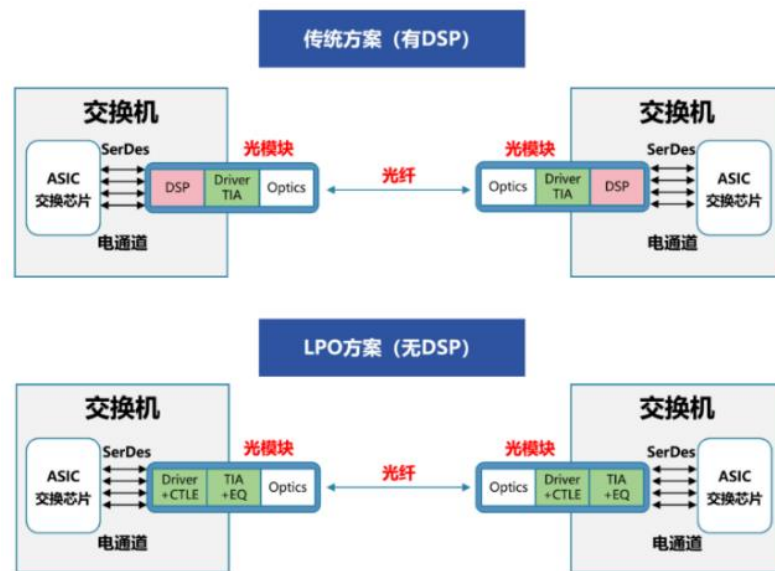
# LPO方案

- ▶ 传统DSP可对高速信号在光-电、电-光之间转换后出现的失真问题进行修复，从而降低失真对系统误码率的影响，但功耗大成本高：1) 400G光模块中，7nm DSP的功耗约为4W，占整个模块功耗的50%；2) 400G光模块中，DSP BOM成本约占20%-40%。相较于传统光模块，LPO技术去除了DSP，将其相关功能集成到设备侧的交换芯片中，只留下具有高线性度的Driver和TIA，用于对高速信号进行一定程度的补偿。
- ▶ LPO技术的优势包括：1) **低功耗**：OFC 2023 Macom展示出的单通道100G单模800G DR8、多模800G SR8 Linear-drive方案中多模功耗节省70%，单模功耗节省50%。根据Macom的数据，具有DSP功能的800G多模光模块的功耗可以超过13W，而采用Macom Pure drive技术的800G多模光模块的功耗不到4W。2) **低延迟**：没有DSP后处理步骤减少，数据传输延迟减少，Macom的Linear-drive方案中延时可降低75%。3) **低成本**：800G光模块中去除DSP后系统总成本可降低约8%。4) **可热拔插**：LPO封装沿用传统热拔插技术，便于后期维护。

图：DSP在400G光模块中的功耗占比约49%



图：LPO光模块架构

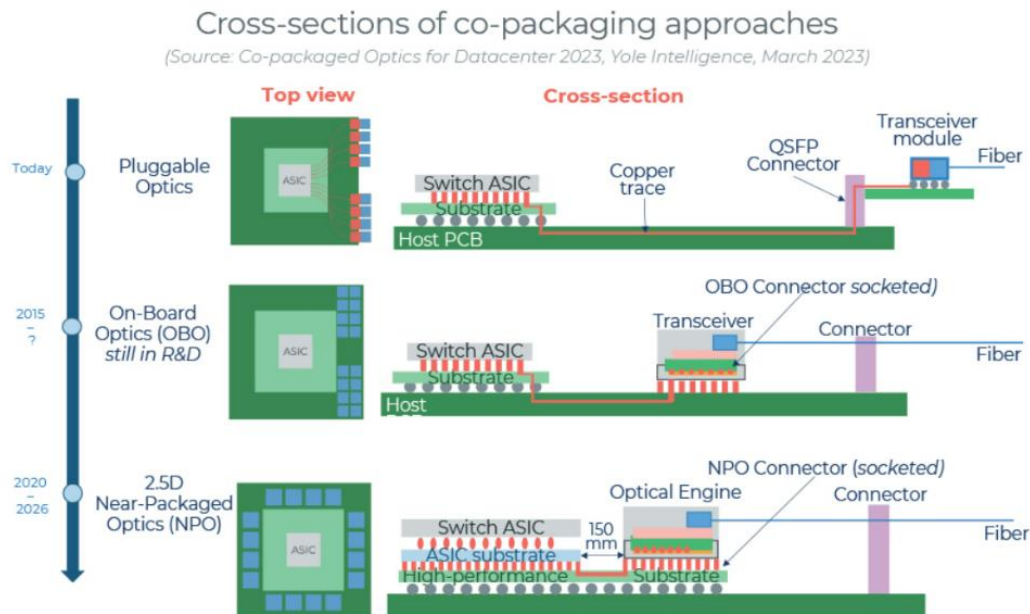




# CPO方案

- CPO是指将光引擎和开关芯片封装在一起的电光集成，这种方法通过缩短光学引擎和开关芯片之间的距离，可以在引擎和芯片之间更快地传输电信号。与传统可热拔插式技术相比，CPO技术的优势包括：
  - 1) **低延迟，低功耗**：由于光模块和交换芯片在同一个封装内，信号传输路径更短，可以实现更低的延迟。另外光电共封装技术可以减少信号传输的功耗，并提高整体系统的能效。
  - 2) **高带宽**：光电共封装技术支持高速光通信，可以提供更大的数据传输带宽。
  - 3) **小尺寸**：相比传统的光模块和电子芯片分离封装的方式，光电共封装技术可以实现更紧凑的尺寸，有利于在高密度集成电路中的应用。

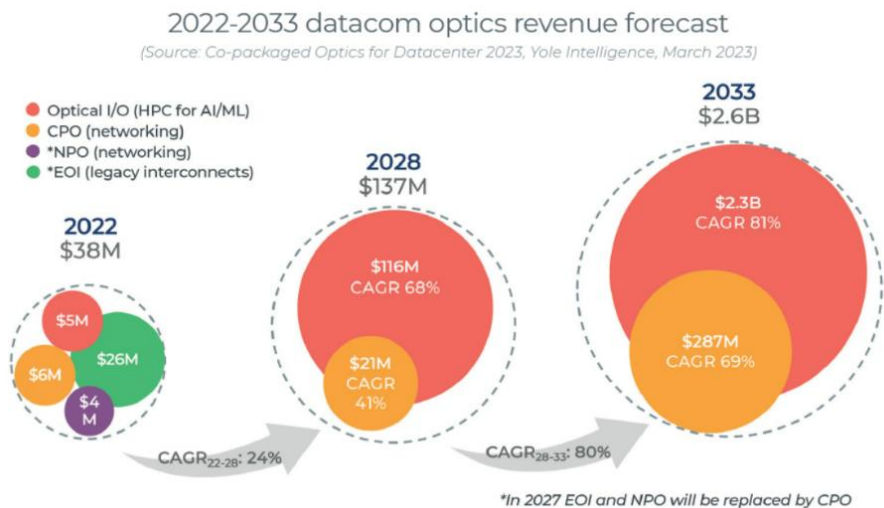
图：传统可热拔插、CPO形态区别



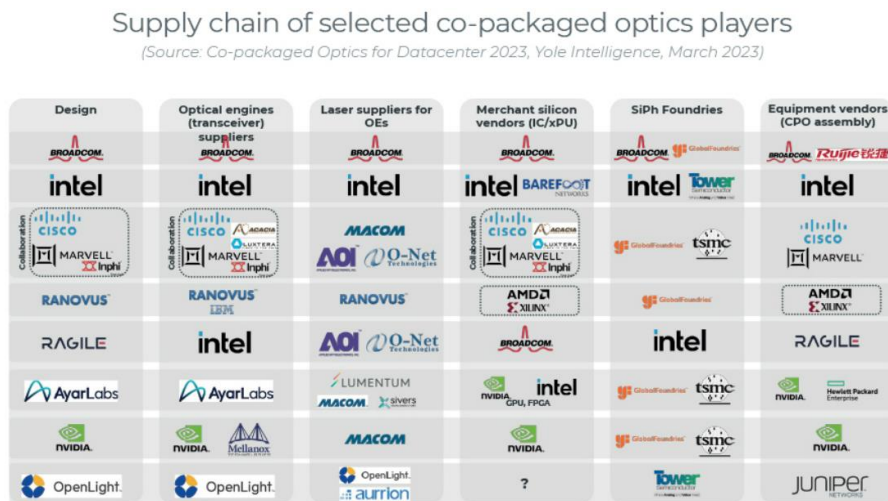
# CPO方案

- CPO目前处于产业化初期，在技术路径、采用紧迫性、商业模式等方面仍面临诸多挑战：
- 1) **技术路径和成熟度**：相干相控振荡器（CPO）技术仍处于早期发展阶段，技术路径需要进一步成熟。虽然 CPO 显示出前景，但它尚未达到迫切需求或广泛采用的阶段。传统的可插拔解决方案继续主导市场，CPO 成为主流选择需要时间。
- 2) **采用的紧迫性**：在3.2T光模块时代，对CPO外形尺寸的需求相对较高。然而在当前1.6T时代，可插拔光模块已经提供了成熟的主流解决方案，配置为8\*200G，有效满足了行业需求。缺乏对更高速 CPO 解决方案的紧迫性可能会减缓其采用速度。
- 3) **商业模式挑战**：实施CPO技术需要大量的硅光子学技术储备。CPO制造工艺的紧密集成使其很可能由开关芯片制造商主导。然而，这种对特定制造商的依赖可能会影响 CPO 技术的进步和广泛应用。
- **算力时代背景下AI对网络速率需求提升，市场空间未来有望突破**。LightCounting在2022年12月报告中称，AI对网络速率的需求是目前的10倍以上，在这一背景下，CPO有望将现有可插拔光模块架构的功耗降低50%，将有效解决高速高密度互联传输场景。Yole报告数据显示，2022年，CPO市场产生的收入达到约3800万美元，预计2033年将达到26亿美元，2022-2033年复合年增长率为46%。

图：2022-2033CPO市场收入CAGR有望实现46%



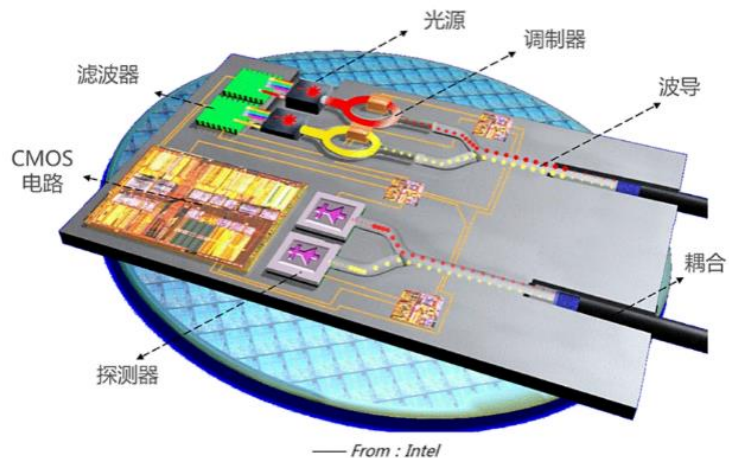
图：2023年CPO产业链企业图谱



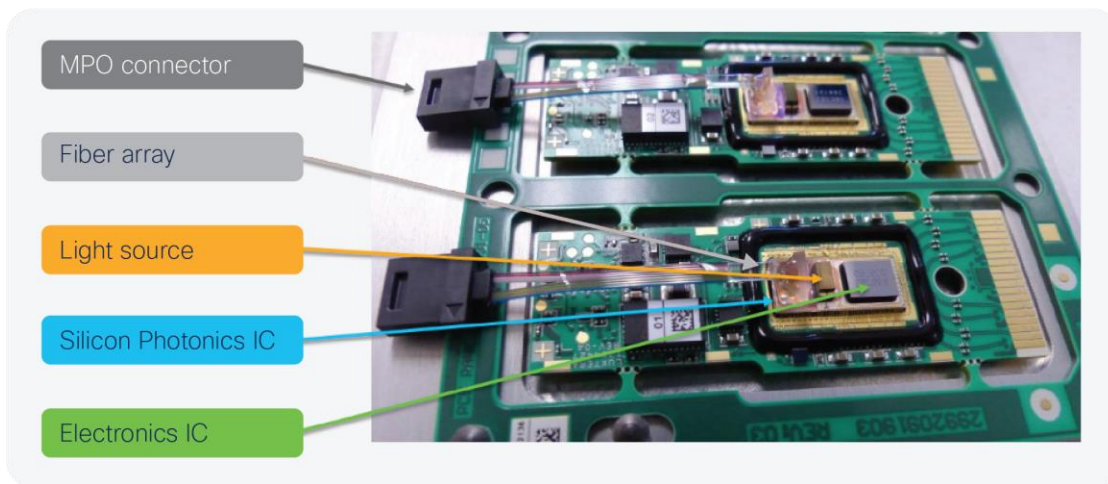
# 硅光方案

- 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技术。硅光子技术的核心理念是“以光代电”，即采用激光束代替电子信号传输数据，将光学器件与电子元件整合至一个独立的微芯片中，提升芯片之间的连接速度。
- 相较传统分立器件方案，硅光子技术的优势包括：**1) 高集成度**：其采用半导体制造工艺将硅光材料和器件集成在同一硅基衬底上，形成由光调制器、探测器、无源波导器件等组成的集成光子器件。相较磷化铟（InP）等有源材料制作的传统分立器件，硅光光模块无需ROSA（光接收组件）、TOSA（光发射组件）封装，因而硅光器件体积与数量更小、集成度更高。**2) 低成本**：相较于传统的分立式器件，硅光模块的集成度更高，封装与人工成本降低；此外硅基材料成本较低且可以大尺寸制造，意味着硅基芯片成本得以大幅降低。**3) 兼容成熟CMOS工艺**：硅光子技术能利用半导体在超大规模、微小制造和集成化上的成熟工艺积累优势。

图：硅基光电子集成芯片概念图



图：基于硅光子的 100 Gbps 光模块示例





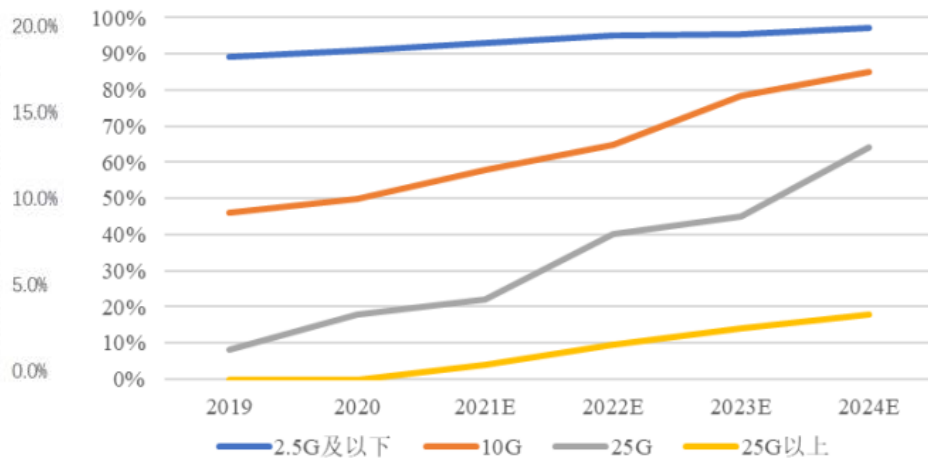
## 光芯片将进入高速增长期，国产化率空间广阔

- C&C数据显示，2022年全球光芯片市场规模达到175亿元人民币（约合27.2亿美元），同比2021年增长了13.8%，同时预测未来受益于互联网应用、人工智能、机器学习、电信城域网升级、家庭宽带网络升级以及专网专线业务等因素的驱动，2022年~2026年间，CAGR仍将保持在10.87%，至2026年全球光芯片市场规模有望突破300亿元人民币（折合46.9亿美元）。预计2025年前后，机器学习应用的成熟将带来更多的算力网络、数据交换的需求，因此增长比例将进入高速增长期，光芯片市场规模增速可达18.4%。
- 我国光芯片企业已基本掌握2.5G及以下速率光芯片的核心技术，10G光芯片方面，2021年国产光芯片占全球比重约60%，但不同光芯片的国产化情况存在一定差异，性能要求较高、难度较大，如10G VCSEL/EML激光器芯片等，国产化率不到40%；25G及以上光芯片方面，2021年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%，仍以海外光芯片厂商为主。

图：2022-2026年全球光芯片市场规模



图：2019-2024年中国光芯片占全球光芯片市场比例预测

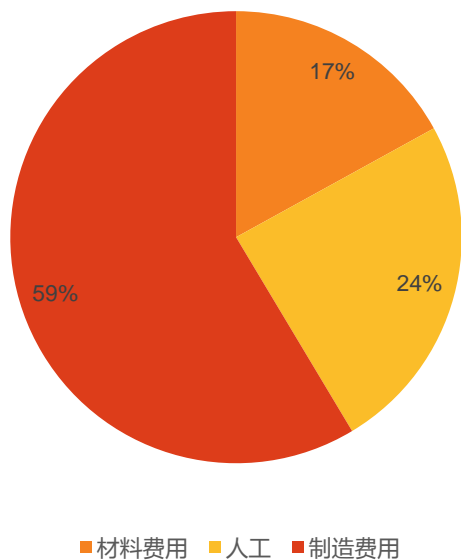


# 光芯片的制造成本构成

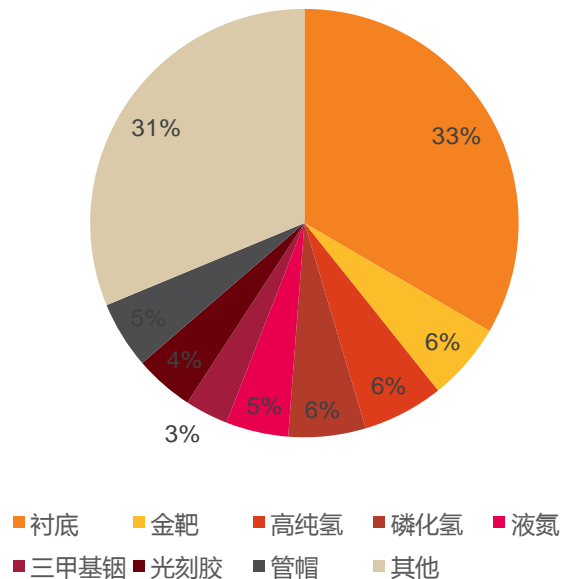
成本中，制造成本占比达59%、人工成本占24%、材料成本占17%（源杰2021年各速率产品成本合并统计）

- （1）制造费用主要包括折旧费、装修费摊销、水电费、光栅加工费等其他费用。
- （2）光芯片的原材料包括衬底、金靶、特殊气体（主要包括高纯氢、磷化氢、液氮等）、三甲基镱、光刻胶、封装材料（包括管帽等）和其他材料等，其他原材料包括显影液、光刻掩模板、异丙醇、砷化氢等材料，其他原材料品种较多且占比较低。
- （3）衬底供应商为通美、住友、云南锗业等。

图：光芯片成本构成（源杰2021年各速率产品成本合并统计）



图：光芯片成本中材料费用的构成（源杰21年数据）





## 2. AI算力：铜缆

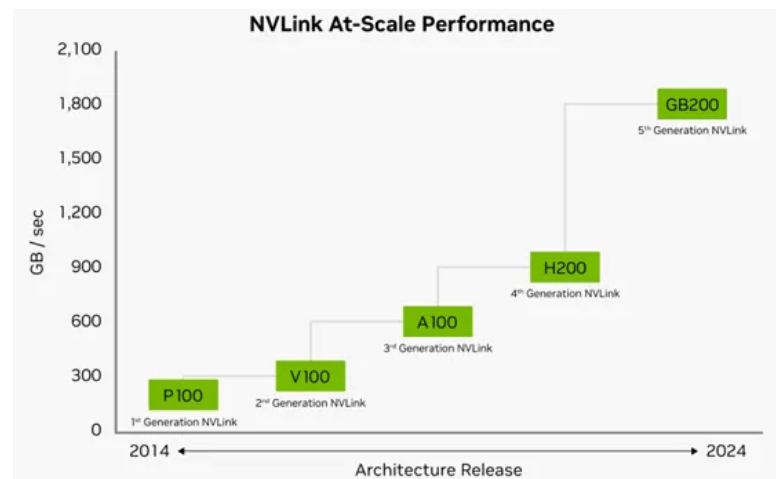
# NVIDIA GTC2024发布新一代 Blackwell 架构平台

- 2024 年 3 月 19 日，NVIDIA 在 GTC 上推出了新一代 Blackwell 架构平台，该平台包括新的 AI 芯片 GB200、第二代 Transformer 引擎（利用专有的 Blackwell Tensor 核心技术将 AI 推理能力和模型大小翻倍）、第五代 NVLink 解决方案（旨在加速数万亿个参数和复杂的 AI 模型，能够互连多达 576 个 GPU，将 GPU 吞吐量提高 9 倍）、RAS（可靠性、可用性和可维护性）引擎（及早识别潜在故障，减少停机时间，增强智能恢复和维护能力），智能安全服务（在不影响整体性能的情况下保护 AI 模型和客户数据，支持为医疗保健和金融等数据隐私要求高的行业量身定制的下一代原生接口加密协议）。
- NVIDIA 的第五代 NVLink 显著增强了大型多 GPU 集群的可扩展性，每个单核 Blackwell GPU 的总带宽为 1.8TB/s。每个 Blackwell Tensor Core GPU 最多可支持 18 个 NVLink 100GB/s 连接，提供高达 1.8TB/s 的总带宽。该带宽性能是 NVIDIA 上一代 H200 产品的两倍，是 PCIe 5.0 技术的 14 倍。

图：NVIDIA 推出下一代 Blackwell 平台



图：NVIDIA 第 5 代 NVLink 网络架构的性能





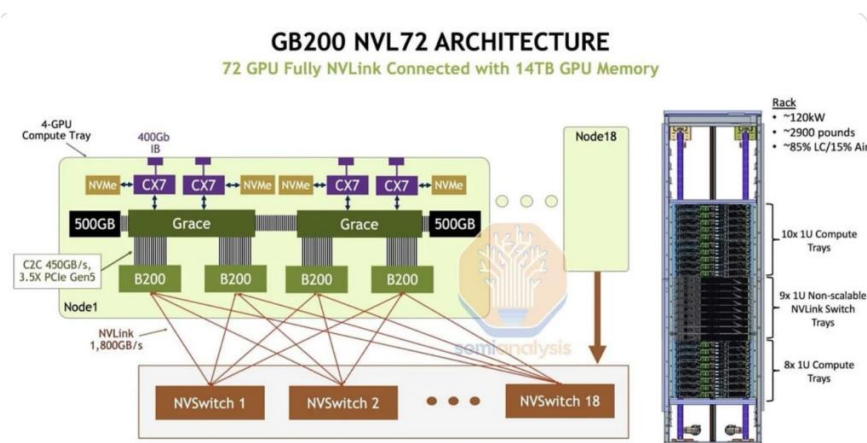
# NVIDIA GB200 NVL72 解决方案

- NVIDIA GB200 NVL72 解决方案将万亿参数AI 模型的推理速度提高了30 倍，训练速度提高了4倍，能效提高了25倍，数据处理性能提高了18倍。
- 1) 在推理性能方面，GB200 NVL72 解决方案采用支持 FP4 AI 的最新一代 Transformer 引擎，并采用第五代 NVLink，为万亿参数大型模型提供 30 倍的大型语言模型（LLM）推理性能提升。它通过新的 Tensor Core 微格式实现了更高的精度和吞吐量，并通过液体冷却在单个机柜中实现了由 72 个 GPU 组成的大型 GPU 集群。
- 2) 在训练性能方面，FP8精度的第二代Transformer引擎将大型语言模型和大规模训练速度提高了4倍。它使用 InfiniBand 交换机网络和 NVIDIA Magnum IO 软件，通过第五代 NVLink 提供 1.8Tb/s 的GPU到GPU互连速度。
- 3) 在电源效率方面，液冷GB200 NVL72显著降低了数据中心的能耗。液体冷却技术提高了计算密度，同时减少了服务器机架占用空间，从而在大型 NVLink 域架构中实现了高带宽、低延迟的GPU通信。与上一代 NVIDIA H100 风冷机柜相比，GB200 液冷机柜在相同功耗下性能提升25倍，同时有效减少用水量。
- 4) 在数据处理性能方面，GB200 利用 NVIDIA Blackwell 架构中的高带宽内存性能、NVLink-C2C 技术和专用解压缩引擎，与CPU相比，将关键数据库查询速度提高了18倍，同时将TCO成本降低了5倍。

图： GB200 NVL72 解决方案



图： GB200 NVL72 内部架构



## 光铜互补：内部铜缆互联，外部光缆互联

- 在交换机和服务器高密度集群化的趋势中，铜缆连接具有性价比和功耗优势，铜缆连接有望成为224Gb/s交换机时代阶段的最佳解决方案。NVIDIA GB200 解决方案的一个重要变化在于交换机和计算节点在单个机柜中的互连，以及通过铜缆连接而不是以前的 PCB-光模块-电缆连接来连接交换机的内部连接。**GB200互连分为三大类：**
- 1) **GB200 NVL72机柜间连接（外接电缆）**：大型数据中心往往需要大量的机柜进行并行计算，如果机柜需要对外联网，则通过TOR交换机用DAC/AOC电缆连接。对于大量的机柜，需要将外部互连安装在机柜上方的布线设备上，以便有序连接，电缆长度往往较长，铜缆超过2至4米后无法满足连接要求，因此长距离互连主要采用光纤进行连接，此链路中的铜缆不能完全取代光纤电缆。

图： NVIDIA GB200 NVL72 机柜互联图



## 光铜互补：内部铜缆互联，外部光缆互联

- 2) GB200 NVL72 单机柜连接（机柜内部电缆）——全部更换为铜缆：8个计算节点和9台交换机通过铜缆内部连接，铜缆背板通过 PCB 背板-光模块-电缆连接取代了之前使用的 TOR 交换机和计算节点。对于单通道 224GB/s 的新一代交换机，800G/1.6T 光模块的功耗通常在 16W 以上，如果 GB200 NVL72 的连接方案是基于之前光模块的连接，则会导致高功耗问题。与光学模块相比，铜缆连接消耗的功率更少，并且更具成本效益。

图：NVIDIA GB200 NVL72 单机箱内交换机和计算节点连接示意图



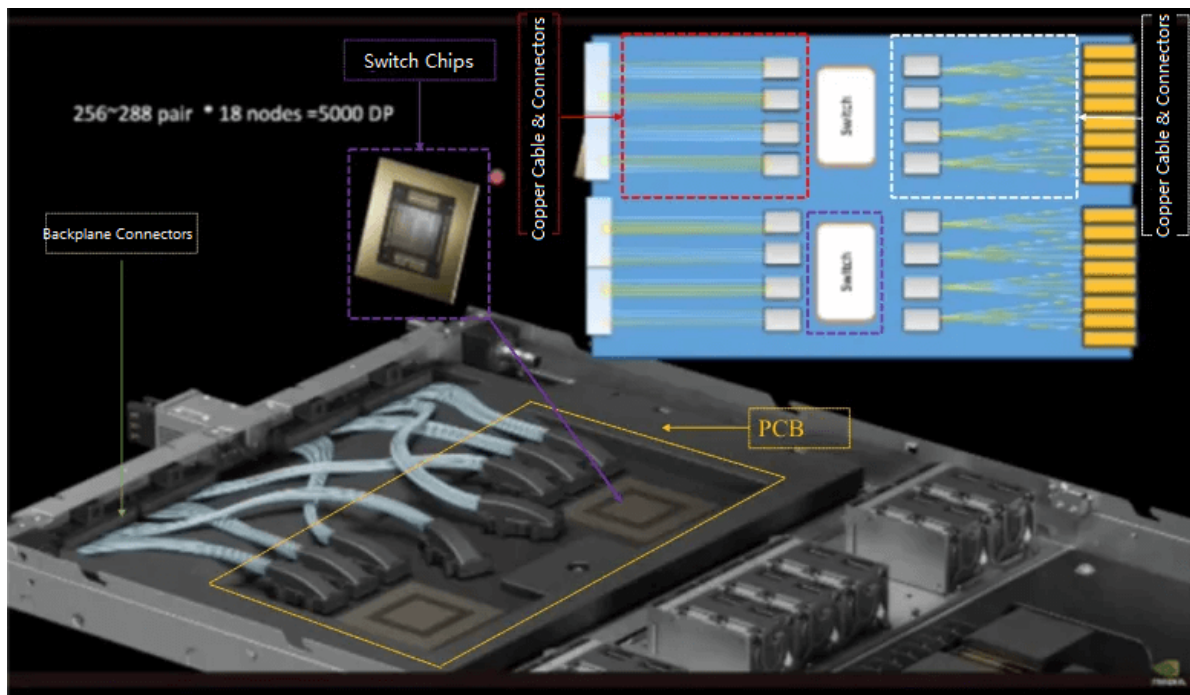
图：NVIDIA GB200 NVL72 铜缆背板和背板连接器原理图



## 光铜互补：内部铜缆互联，外部光缆互联

- 3) NV开关内部——使用铜线实现从背板连接器到开关芯片的连接：对于单通道224Gb/s开关，PCB板面积也有限，不足以覆盖整个区域，因此无法实现更远距离的链路连接，而铜跳线方式可以实现从背板到开关芯片的连接。
- 整体来看，短距用铜，长距用光。在需要超大带宽和远距离传输时（如工业以太网布线），选择光缆更为合适；而在短距离传输和低传输速率的情况下（如大楼内或企业办公室内），考虑到成本问题，铜缆更为合适。

图：NVIDIA GB200 NVL72 交换机内部铜缆连接方案示意图



## 光铜互补：Blackwell平台推动800G、1.6T光模块需求提升

- 英伟达的新一代Blackwell平台将带动对更高传输速率的1.6T光模块的需求：根据NVIDIA发布的Quantum-X800 Q3400-RA 4U InfiniBand交换机的性能，全球首款采用单通道200Gb/s技术的交换机，提供144个端口，连接速度为800GB/s，由72个1.6T OSFP光模块实现。因此，预计新一代交换机在GB200解决方案中的逐步应用将带动对1.6T光模块的需求。
- Blackwell平台在大规模GPU集群应用场景中仍需光模块实现机柜间互联，对800G光模块的需求将维持：
  - 1) GB200单机柜（对应72个GPU）：新一代GB200单机柜方案将不再需要光模块实现互联。
  - 2) 在1-8个GB200 NVL72集群（对应72-576个GPU）之间，还需要一些800G光模块来实现机柜之间的互联互通。如果20%的数据需要跨机柜传输，NVLink单向总传输带宽为7200Gb，对应单个GPU和800G光模块的需求比为1：2。
  - 3) 8个以上的大型GB200 NVL72集群（对应576个以上的GPU），预计配置InfiniBand三层网络，按GPU和800G光模块的需求比例1：2.5，第二层1：2，预计GB200整体需求比例为1：4.5。

图：NVIDIA 新一代 IB 交换机平台 Quantum-X800





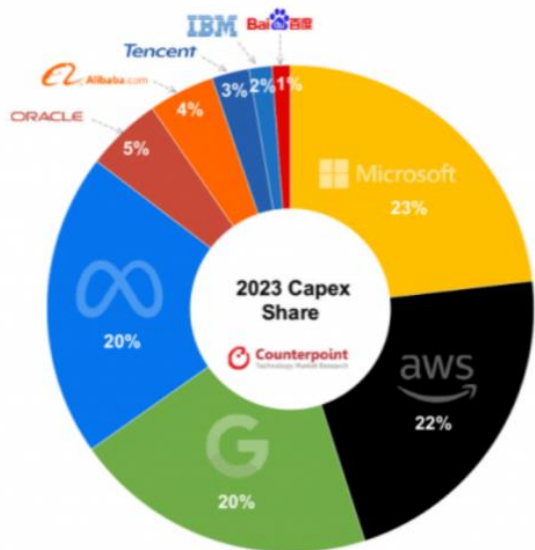
### 3. AI算力：IDC、交换机、服务器

## IDC：全球数据中心数量、容量持续攀升

- 根据Synergy Research Group数据显示，截止2023年底全球Hyperscale数据中心数量已增加到992个，并在2024年突破1000个。同时随着设施数量的增长和平均容量的攀升，Hyperscale数据中心总容量仅用了4年时间就翻了一倍。按照关键IT负载的兆瓦计算，美国占比全球51%，欧洲和中国各占17%和16%。Synergy预测未来4年Hyperscale数据中心的总容量有望再次翻倍，且每年将新增120-130个，而生成式AI技术是规模扩大的主要原因。此外根据Business Research Insights预测，2022年全球数据中心服务市场规模为756.2亿美元，2028年该市场将达到1484.6亿美元，年化复合增速11.9%。

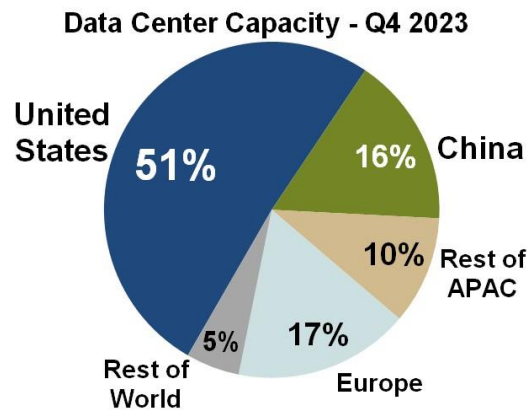
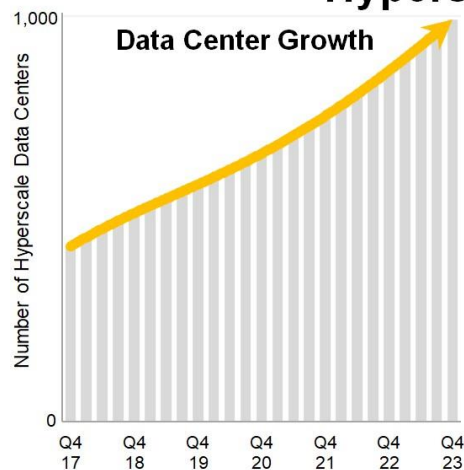
图：2023年全球云服务提供商资本开支占比

Global Cloud Service Providers' Capex Share



图：2023年全球超大数据中心图谱

### Hyperscale Data Centers



Source: Synergy Research Group

# IDC：受新基建、数字化及数字中国远景目标，行业整体稳健发展：

2021年，我国数据中心行业市场收入达到1500亿元左右，近3年年均复合增长率达到30.69%。

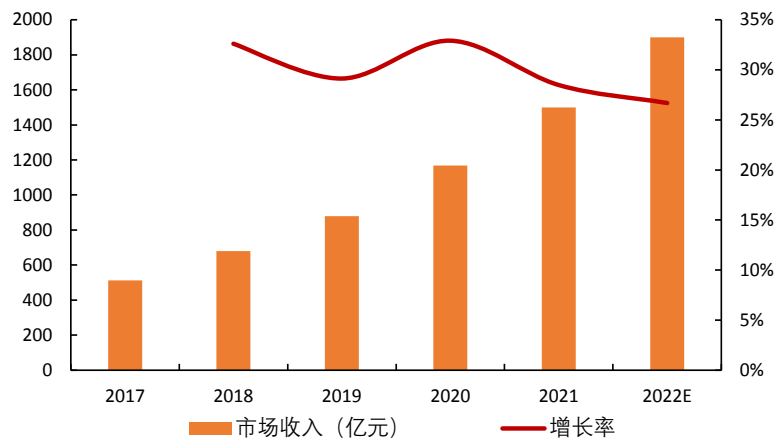
- 我国数据中心机架规模持续稳步增长，大型以上数据中心规模增长迅速。按照标准机架2.5kW统计，截止到21年年底，我国在用数据中心机架规模达到520万架，近五年年均复合增速超过30%。其中，大型以上数据中心机架规模增长更为迅速，按照标准机架2.5kW统计，机架规模420万架，占比达到80%。

中西部地区自然环境优越，土地、电力等资源充足，但本地数据中心市场需求相对较低；东部地区市场需求旺盛，但土地、电力、人员等生产要素成本较高，东西部协同发展逐渐成为趋势。

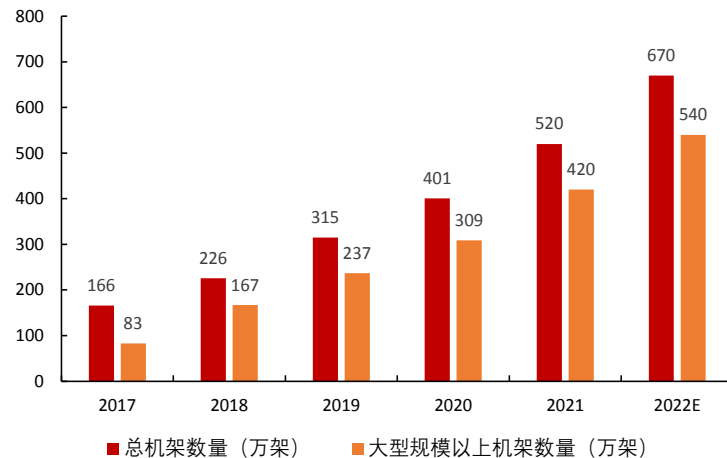
- 随着后续网络质量的优化，中西部将不再仅是进行冷存储的灾备数据中心聚集区，也将承载更多的应用。

作为算力服务中枢，数据中心既是数字经济底座，也是数字技术创新的高地。随着新一代信息技术的不断发展，数据中心正逐渐突破传统机房运营模式，产业发展逐渐由资本驱动迈向创新驱动。

图：我国数据中心市场规模



图：我国数据中心机架规模





# IDC：算力多元化发展，IDC中长期趋势向好

算力呈现多元化发展趋势：多样化的智能场景需要多元化的算力，以AIGC为代表的人工智能应用、大模型训练等新应用、新需求快速崛起都对算力提出更高要求。

- 基础算力方面，在全球数据中心快速发展的推动下，基础算力规模持续增长，其中云计算、边缘计算等场景依然是基础算力的主要应用场景。根据IDC数据，2022年全球云计算IaaS市场规模增长至1155亿美元，同比上涨26.2%。
- 智能算力方面，近年新推出的大语言模型所使用的数据量和参数规模呈现“指数级”增长，带来智能算力需求爆炸式增加。
- 超算算力方面，十年千倍定律依然在持续，橡树岭国家实验室（ORNL）的Frontier仍然是全球TOP500上唯一的百亿亿次级机器。

在算力形态方面，我国数据中心形态多样化发展趋势逐渐显现，智算中心、边缘数据中心将保持高速增长。

- 当前，通用算力的数据中心仍是市场主力，按机架规模统计，占比超过90%；
- 超算中心主要应用于国家重大科研领域，商业应用场景较少；智算中心从早期实验探索逐步走向商业试点，随着我国各类人工智能应用场景的丰富，智算需求将快速增长，预期规模增速将达到70%；
- 边缘数据中心能够为智能终端、物联网设备提供实时算力，边缘数据中心的规模增速有望达到30%。

图：不同场景下的异构算力需求



图：我国算力规模及增速（EFlops）



# IDC：PUE逐步趋紧，但高算力IT设备成为主流，单机柜功率、密度有望持续提升，关注液冷发展机遇

- 22年6月工业和信息化部、国家发展改革委、财政部、生态环境部、国务院国资委、市场监管总局联合印发《工业能效提升行动计划》，其中提出“到2025年，新建大型、超大型数据中心电能利用效率（PUE，指数据中心总耗电量与信息设备耗电量的比值）优于1.3。”
- 同时，向数据中心提出：加快液冷、自然冷源等制冷节能技术应用，鼓励采用分布式供电、模块化机房及虚拟化、云化IT资源、高温型IT设备等高效系统和设备，推广高压直流供电、集成式电力模块等技术，发展智能化能源管控系统。鼓励数据中心在保证安全运行的前提下，优化减配冗余基础设施，自建余热回收设施。
- 高算力IT设备已经成为主流，因算力、密度持续提升，多样性算力协同要达到15-30kW，混合密度将成为主流。芯片更新周期为2年，且未来5年持续向高算力、高功率密度演进，不同类型算力功耗差异拉大。

图：能源利用效率（PUE）逐年收紧

| 颁布时间（年） | 发布单位： | 涉及 PUE 的内容  |
|---------|-------|---|
| 2016 年  | 国务院   | 印发“十三五”节能减排综合工作方案的通知：进一步推广云计算技术应用，新建大型云计算数据中心能源利用效率（PUE）值优于 <b>1.5</b> 。  |
| 2019 年  | 工信部   | 关于加强绿色数据中心建设的指导意见：到 2022 年，数据中心平均能耗基本达到国际先进水平，新建大型、超大型数据中心的电能使用效率值达到 <b>1.4</b> 以下，高能耗老旧设备基本淘汰，水资源利用效率和清洁能源应用比例大幅提升，废旧电器电子产品得到有效回收利用。   |
| 2021 年  | 工信部   | 印发《新型数据中心发展三年行动计划（2021-2023 年）》的通知：到 2023 年底，全国数据中心机架规模年均增速保持在 20%左右，平均利用率力争提升到 60%以上，总算力超过 200 EFLOPS，高性能算力占比达到 10%，国家枢纽节点算力规模占比超过 70%。新建大型及以上数据中心 PUE 降低到 <b>1.3</b> 以下，严寒和寒冷地区力争降低到 1.25 以下。 |
| 2022 年  | 上海市   | 关于推进本市数据中心健康有序发展的实施意见：集聚区新建大型数据中心综合 PUE 降至 1.25 左右。   |
| 2023 年  | 北京市   | 关于征求《数据中心能源效率限额》北京市地方标准意见的通知：已建成并稳定运行一个自然年以上以及改扩建的数据中心，其 PUE 值应符合 1.3 的要求；新建的数据中心，其 PUE 值应符合 <b>1.2</b> 的要求；数据中心管理者应通过节能技术改造和加强节能管理，使数据中心 PUE 值达到 1.15 的要求。                                     |

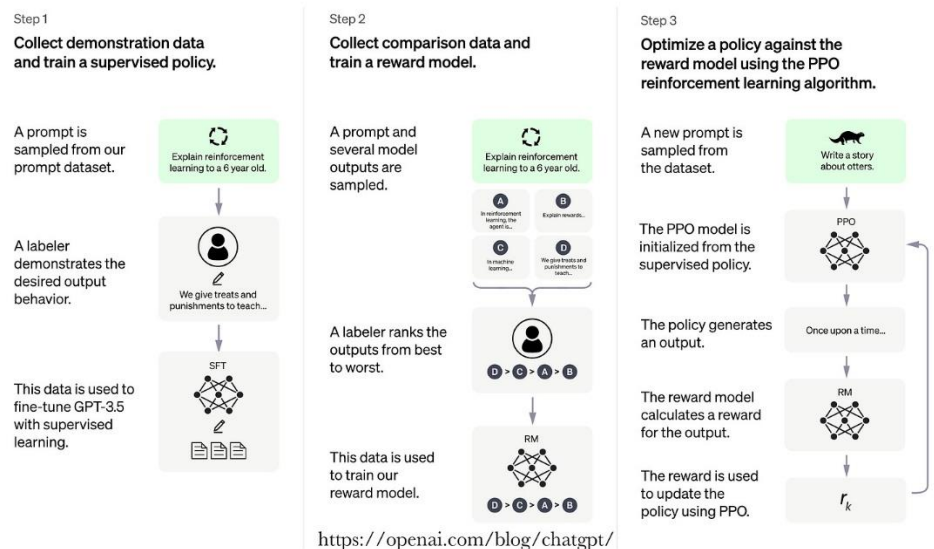
图：高算力设备已成为主流



# IDC：算力租赁乘风起，短期供需格局推动单P租金持续提升

- GPT模型通过海量参数，进一步提升了模型的精确度。初代的GPT模型参数是1.17亿，而GPT2的模型参数达15亿。GPT3模型参数达到了1750亿，是GPT2参数的100倍；GPT-4的模型参数在1.8万亿左右、13万亿训练数据。如果用ChatGPT部署到谷歌进行的每次搜索中，需要512820.51台英伟达顶级超算服务器和总共4102568个顶级超算GPU，这些服务器和网络的总成本仅资本开支就超过1000亿美元。
- 由此，算力租赁乘风起，**算力租赁是对算力进行出租，是一种通过云计算服务提供商租用计算资源的模式**。用户可以根据自己的需求租赁服务器或虚拟机实现大规模的计算任务，而无需拥有自己的计算资源。算力租赁是一种灵活、高效、成本低廉的计算服务，适用于各种大规模计算需求的场景。
- 由于GPU的紧缺，算力租赁价格持续上行，以中贝通信为例，公司23年11月16日发布《关于签订算力服务框架合同》的公告中显示，单价为18万元/P/年。相较于23年9月7日发布的《关于签订算力服务框架合同》的公告中的单价12万元/P/年，涨幅达50%。

图：ChatGPT训练步骤



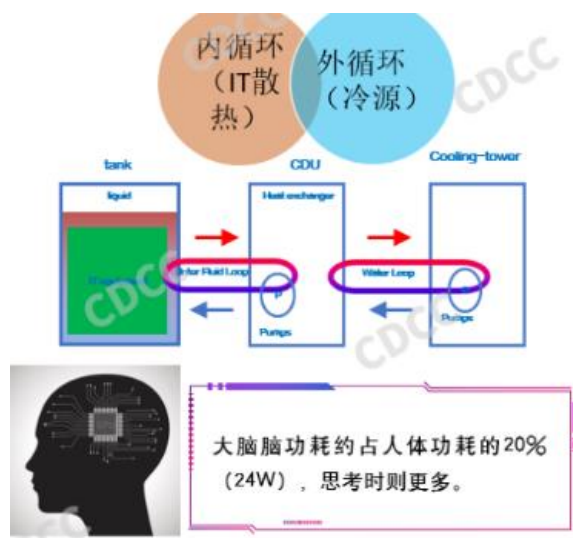
## 液冷：功率提升带来的散热需求或促使液冷发展

- A100/H100服务器系统功耗明显提升。以NVIDIA的DGX A100 640GB为例，其配置了8片A100 GPU，系统功耗达到最大6.5千瓦。H100更是最高功耗达到10.2kW。同时这将使得单机柜功率或将大幅提升，液冷方案或将更为适用。
- 在数据中心的环境中，散热对于整个数据中心和基础设施的稳定性至关重要。同时，数据中心也追求以最少的电力来确保整个IT系统的可靠运行。
- 技术的发展、节能的需求以及国家政策要求等诸多因素，再加上人工智能的催化，使得液冷技术的发展正当其时。随着能源紧张和环境保护的重要性日益凸显，液冷技术的高效节能特点使其成为可持续发展的重要解决方案。
- 我们看到，A100到H100，功耗有持续提升，到了性能进一步增强的B100，功耗或将进一步增加，带来液冷的需求，液冷行业或迎来催化。

图：PUE成为数据中心重要指标



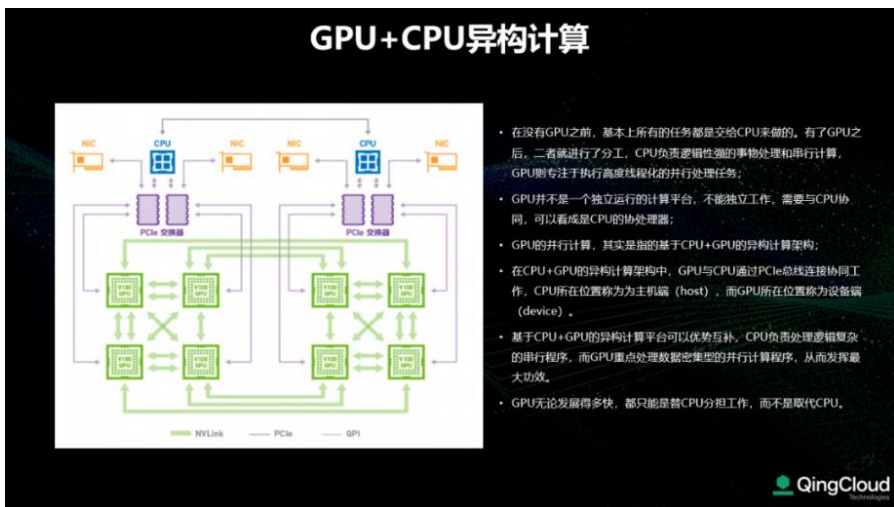
图：数据中心冷却技术液冷的两个循环



# 服务器：AI带来异构计算演进，服务器价值量提升

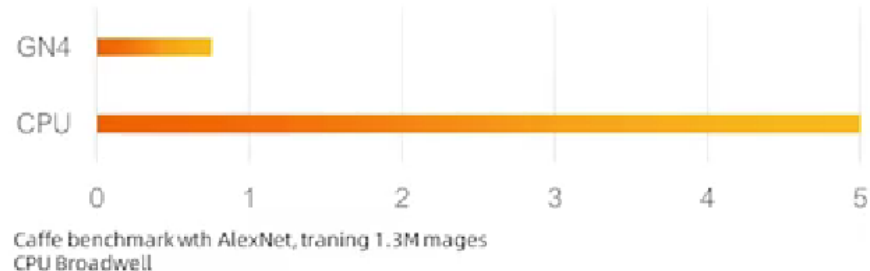
- ▶ 未来异构计算或成为主流
- ▶ 异构计算（Heterogeneous Computing）是指使用不同类型指令集和体系架构的计算单元组成系统的计算方式，目前主要包括GPU云服务器、FPGA云服务器和弹性加速计算实例EAS等。让最适合的专用硬件去服务最适合的业务场景。
- ▶ 在CPU+GPU的异构计算架构中，GPU与CPU通过PCIe总线连接协同工作，CPU所在位置称为主机端(host)，而GPU所在位置称为设备端(device)。基于CPU+GPU的异构计算平台可以优势互补，CPU负责处理逻辑复杂的串程序序，而GPU重点处理数据密集型的并行计算程序，从而发挥最大功效。
- ▶ 越来越多的AI计算都采用异构计算来实现性能加速。
- ▶ 阿里第一代计算型GPU实例，2017年对外发布GN4，搭载Nvidia M40加速器.，在万兆网络下面向人工智能深度学习场景，相比同时代的CPU服务器性能有近7倍的提升。

图：GPU+CPU异构计算



图：GPU面向深度学习场景性能比CPU提升近7倍

## 面向深度学习场景性能比CPU提升近7倍



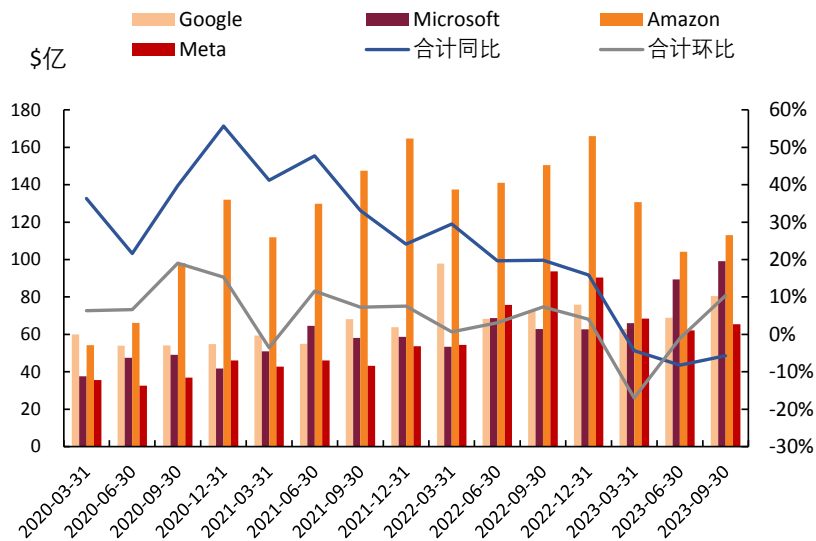
# 服务器：AI带来异构计算演进，服务器价值量提升

- 从全球AI服务器采购情况来看，据TrendForce统计，北美四大云端供应商微软、谷歌、Meta、AWS合计占比约66%。国内企业中，字节跳动采购量最高，占比6%；紧接在后的的是腾讯、阿里巴巴、百度，采购量分别占比2.3%、1.5%和1.5%。
- 北美云厂商资本开支2024年有望回升。亚马逊23Q4资本开支133.53亿美元，同比减少13.52%，公司CFO Brian Olsavsky表示2024年公司资本支出预计将有所增加，主要源于云业务的区域扩张及容量提升，特别是与AIGC项目相关的支出。微软23Q4资本开支97.35亿美元，同比增长55.16%，公司CFO Amy Hood预计公司资本开支将在24年Q1显著环比扩大。谷歌23Q4资本开支110.19亿美元，同比增长45.08%，公司CFO Ruth Porat在分析师电话会议上宣称24年资本支出将显著增加，因公司将继续投资服务器、数据中心等基础设施，为其人工智能产品铺路。

图：AI服务器采购量占比



图：北美云厂商资本开支（\$亿）

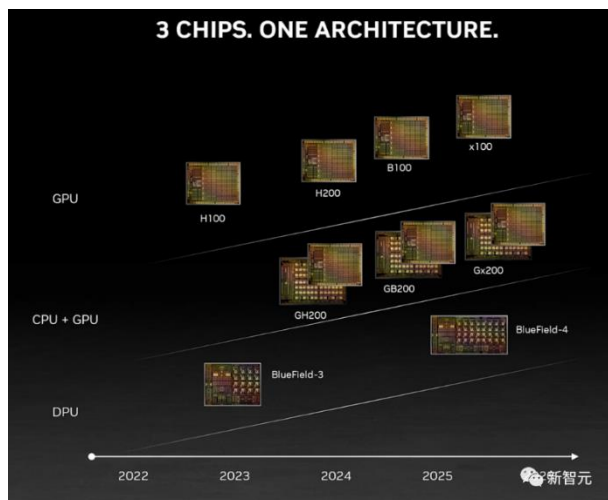


# 服务器：AI服务器持续升级迭代，关注国产相关机会

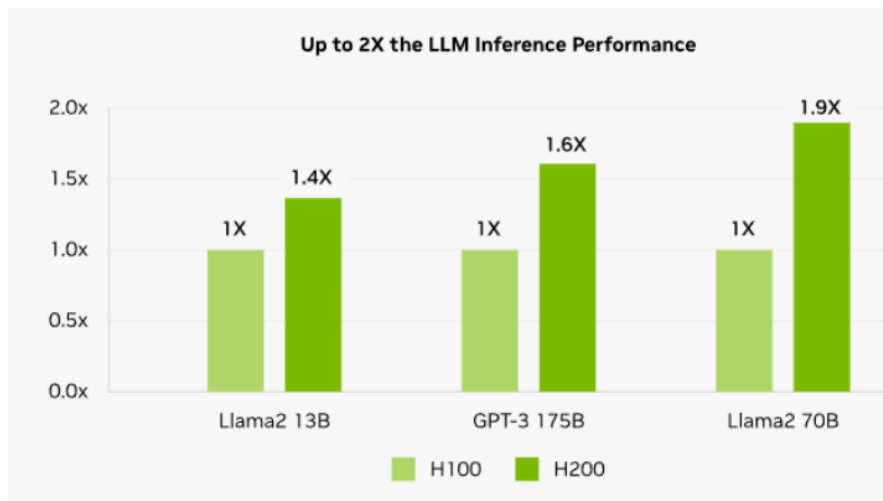
## 服务器关注几大核心变化：

- 1、**针对中国市场的AI芯片推出**：在美国商务部新的半导体出口限制，针对A800和H800芯片进行管控后，英伟达或将推出三款针对中国市场的AI芯片，以应对美国的最新芯片限售令。关注三款芯片的详细规格、价格等数据以及推出时间或对于产业在国内发展产生影响。
- 2、**服务器持续迭代升级，2024年B100或将面世**：英伟达近期发布AI芯片H200，与A100相比，H200的容量几乎翻了一番，带宽也增加了2.4倍。与H100相比，H200的带宽则从3.35TB/s增加到了4.8TB/s。此外2024年，英伟达还会发布基于Blackwell架构的B100，以推理1750亿参数的GPT-3为例，2023年刚发布的H100是前代A100性能的11倍，2024年即将上市的H200相对于H100则有超过60%的提升，而再之后的B100性能预计继续大幅提升。
- 3、**国产算力的机会**。2023年10月15日，中国电信AI算力服务器(2023-2024年)集中采购项目公布候选人，总集采金额超84亿元，其中国产算力的商业化已经有相当规模。据《IT时报》，基于昇腾910B的大规模智算集群基础设施，已经在全国一些智算中心和电信运营商内部开始测试。同时国内厂商包括寒武纪、海光等，均有望迎来良好发展机遇。

图：英伟达产品线路图



图：芯片性能持续迭代升级

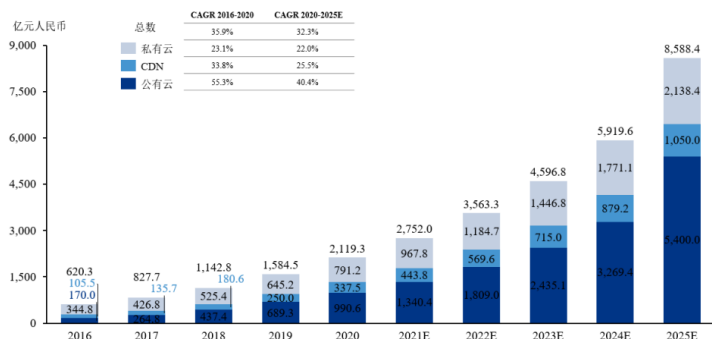


# 交换机核心驱动力：技术应用带来流量增长

## 核心驱动力：

- 本质是新技术应用驱动流量增长：5G普及，基站数量持续增长、移动流量的快速增长、固定宽带传输速率提升、云计算提升云流量的增长、物联网挖掘物体间流量价值等等。
- 而流量的增长将驱动数据中心、交换机等网络设备需求的提升。

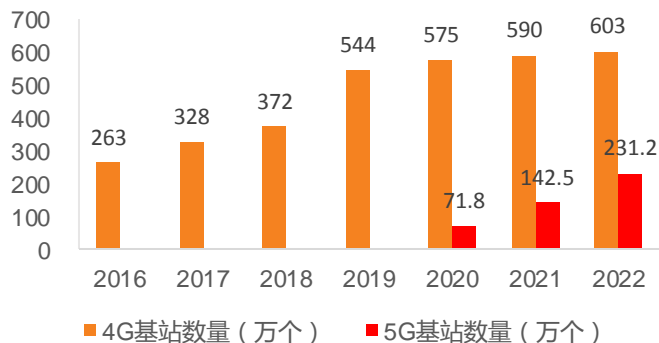
图：中国云计算规模



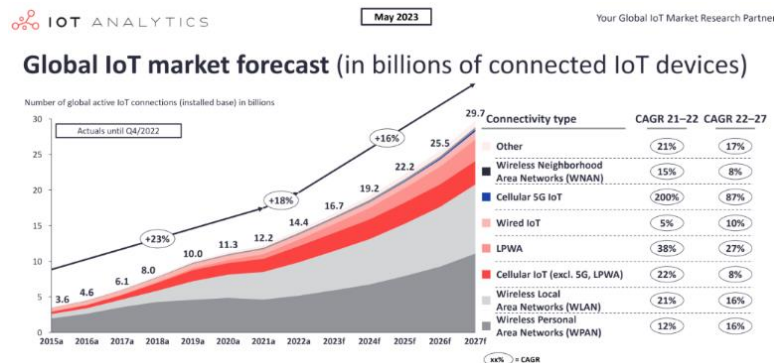
图：移动互联网流量增长



图：中国基站数量



图：IOT连接数

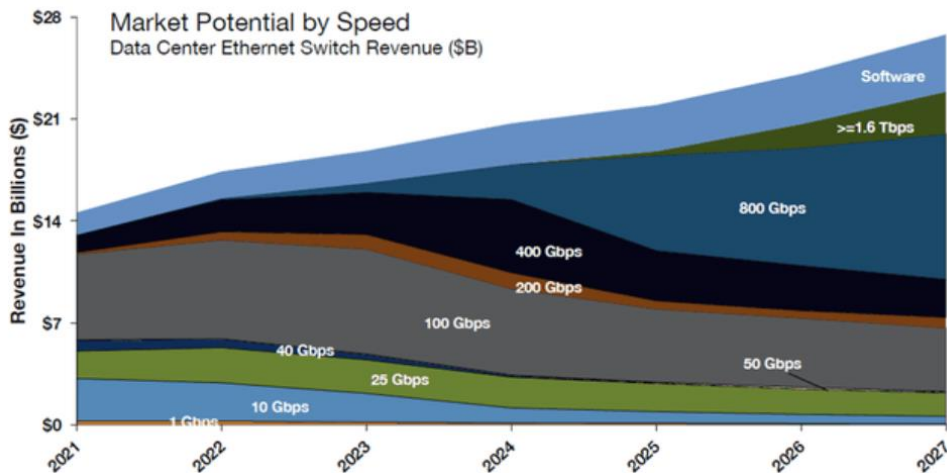




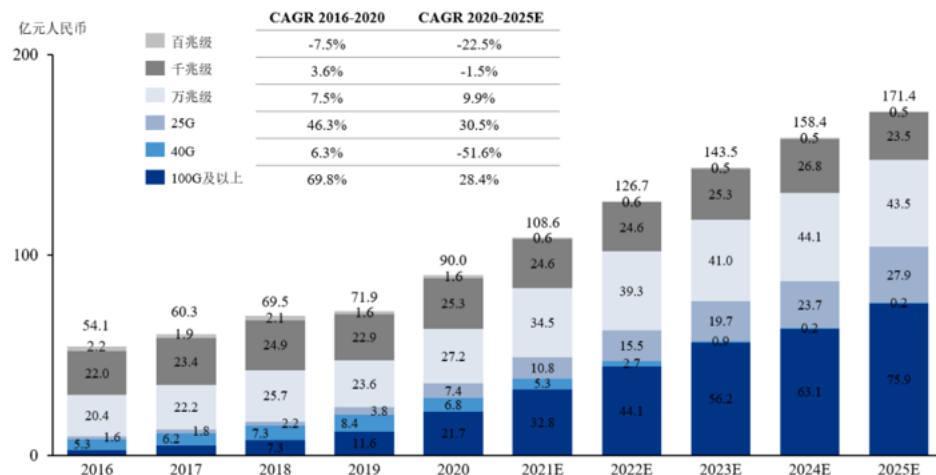
# 交换机：传输速率有望持续升级

- 从端口速率看，以太网交换芯片可分为百兆、千兆、万兆、25G、40G、100G 及以上不等。近年数字经济的快速发展，推动了云计算、大数据、物联网、人工智能等技术产业的快速发展和传统产业数字化的转型，均对网络带宽提出新的要求。
- 随着数据中心交换机向更高速度的过渡，高速率数据中心交换机的价值量也相应提升。Dell’Oro 预计全球以太网交换机数据中心市场将在 2021 年-2026 年达到近两位数的 CAGR 增长，到预测期末，400Gbps、800Gbps 和 1600Gbps 将占市场 50% 以上收入。国内方面，灼识咨询预计至 2025 年，100G 及以上和 25G 的中国商用以太网交换芯片市场规模将大幅增长，占比将分别达到 44.2% 和 16.3%，2020-2025 年年均复合增长率将分别达到 28.4% 和 30.5%。
- Chatgpt 可能进一步加速升级。

图：以太网交换机各端口速率市场空间预测



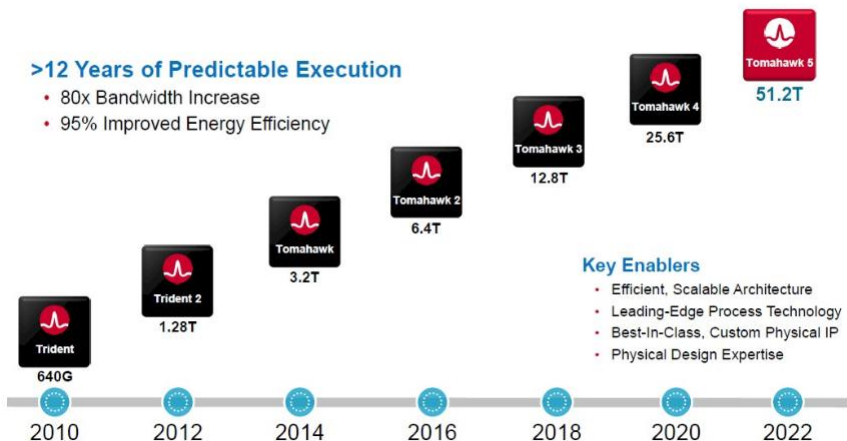
图：中国商用以太网交换芯片各端口速率市场规模情况（以销售额计）



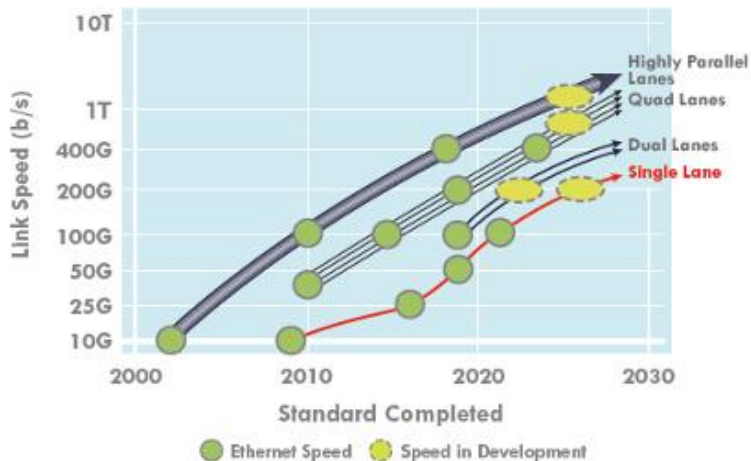
# 交换机：2025年800G交换机有望超过400G

- 交换机芯片容量两年翻一倍：博通始终代表着交换芯片的最先进解决方案，Tomahawk系统芯片是高带宽交换机平台芯片，Tomahawk 5是Broadcom推出的5纳米工艺的51.2T交换容量的芯片，主要针对超大规模企业和云服务商用交换机和路由器芯片市场。
- Tomahawk 5可以驱动51.2TB/秒的总带宽，64个以800Gb/秒运行的端口、128个以400Gb/秒运行的端口和256个以200Gb/秒运行的端口，超大规模企业和云服务商喜欢具有128端口基数、每端口运行速度为400Gb/秒的交换机，其中64个端口向下连接到机架中的服务器，64个端口向上连接到网络结构的主干层。
- 800G交换机预计在2024年开始快速渗透上量：根据Dell ‘Oro预测，到2027年，近一半的数据中心交换机端口将由400Gbps及更高的速度驱动，同时预计到2025年，800Gbps将超过400Gbps。

图：博通芯片方案持续升级迭代



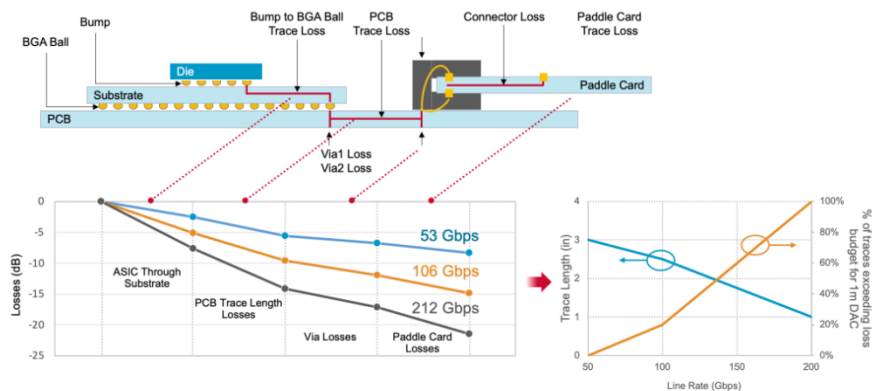
图：Ethernet链路速率展望图



# PCB：AI服务器及800G交换机升级带来增量

- 800G交换机对于PCB要求提升。随着高速信号的传输速率从50G增加到100G再到200G，传输系统的损耗从低于10dB增加到超过20dB。在信号传输速率超过200G时，PCB板的互联复杂性增加，几乎所有信号线的走线长度都将超出1m DAC线路传输的损耗预算。
- AI服务器PCB价值量明显提升：AI服务器对芯片性能及传输速率要求更高，通常需使用高端的GPU显卡以及更高的PCIe标准。同时，通用服务器PCB采用8-10层M6板为主，训练服务器PCB 18-20层M8板，推理型服务器PCB为14-16层M6板，AI服务器PCB价值量为1万-1.5万元，价值量明显提升。
- 对于PCB要求、层数的提升，使得高端产品竞争力重要性日渐凸显。其中，沪电股份拥有800G交换机PCB产品经验，公司在高阶数据中心交换机领域，应用于Pre800G（基于56Gbps速率，25.6T芯片）的产品已批量生产，应用于800G（基于112Gbps速率，51.2T芯片）的产品已实现小批量的交付。

图：速率增加对于PCB互联复杂性提升



图：800G交换机拆截图



## 风险提示

1. AI应用发展不及预期风险：AI应用持续迭代，存在技术发展不及预期的可能性，导致整体行业发展进度放缓。
2. 海外大厂投资不及预期风险：海外厂商资本开支不及预期，影响下游供应链需求。
3. 中美摩擦升级风险：中美科技、金融、贸易等领域摩擦加剧。
4. 行业竞争加剧，价格和盈利能力下降风险：进入行业的厂商增加，价格竞争使得盈利能力下降。
5. 公司业绩不及预期风险：相关产品研发、推广以及市场拓展不及预期。

## 分析师声明

本报告署名分析师在此声明：我们具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，本报告所表述的所有观点均准确地反映了我们对标的证券和发行人的个人看法。我们所得报酬的任何部分不曾与，不与，也将不会与本报告中的具体投资建议或观点有直接或间接联系。

## 一般声明

除非另有规定，本报告中的所有材料版权均属天风证券股份有限公司（已获中国证监会许可的证券投资咨询业务资格）及其附属机构（以下统称“天风证券”）。未经天风证券事先书面授权，不得以任何方式修改、发送或者复制本报告及其所包含的材料、内容。所有本报告中使用的商标、服务标识及标记均为天风证券的商标、服务标识及标记。

本报告是机密的，仅供我们的客户使用，天风证券不因收件人收到本报告而视其为天风证券的客户。本报告中的信息均来源于我们认为可靠的已公开资料，但天风证券对这些信息的准确性及完整性不作任何保证。本报告中的信息、意见等均仅供客户参考，不构成所述证券买卖的出价或征价邀请或要约。该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。客户应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专家的意见。对依据或者使用本报告所造成的一切后果，天风证券及/或其关联人员均不承担任何法律责任。

本报告所载的意见、评估及预测仅为本报告出具日的观点和判断。该等意见、评估及预测无需通知即可随时更改。过往的表现亦不应作为日后表现的预示和担保。在不同时期，天风证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。

天风证券的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。天风证券没有将此意见及建议向报告所有接收者进行更新的义务。天风证券的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

## 特别声明

在法律许可的情况下，天风证券可能会持有本报告中提及公司所发行的证券并进行交易，也可能为这些公司提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。因此，投资者应当考虑到天风证券及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突，投资者请勿将本报告视为投资或其他决定的唯一参考依据。

## 投资评级声明

| 类别     | 说明                         | 评级   | 体系               |
|--------|----------------------------|------|------------------|
| 股票投资评级 | 自报告日后的6个月内，相对同期沪深300指数的涨跌幅 | 买入   | 预期股价相对收益20%以上    |
|        |                            | 增持   | 预期股价相对收益10%-20%  |
|        |                            | 持有   | 预期股价相对收益-10%-10% |
|        |                            | 卖出   | 预期股价相对收益-10%以下   |
| 行业投资评级 | 自报告日后的6个月内，相对同期沪深300指数的涨跌幅 | 强于大市 | 预期行业指数涨幅5%以上     |
|        |                            | 中性   | 预期行业指数涨幅-5%-5%   |
|        |                            | 弱于大市 | 预期行业指数涨幅-5%以下    |

THANKS