

评级：看好

何立中

电子行业首席分析师

SAC 执证编号：S0110522110002

helizhong@sczq.com.cn

电话：010-81152682

韩杨

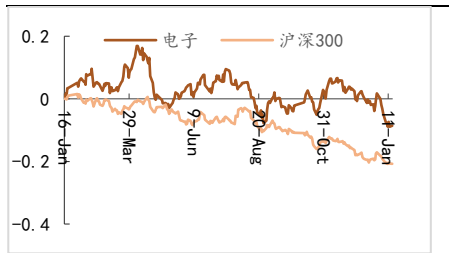
电子行业分析师

SAC 执证编号：S0110523070002

hanyang@sczq.com.cn

电话：010-81152681

市场指数走势（最近1年）



资料来源：聚源数据

相关研究

- 电子行业：消费电子新机热度高企，关注半导体设计弹性标的
- 基本面或已完成筑底，看好需求复苏兑现业绩增长
- 消费电子持续修复，华为表现亮眼

核心观点

- **电子年度涨幅排名已进入上升周期。**从 2007 年至今，电子板块涨幅相对排名呈现约 4 年周期：2007~2011(4 年)、2011~2014(3 年)、2014~2018(4 年)、2018~2022(4 年)。A 电子板块排名垫底年份的下一年有较好的相对排名。我们在 2023 年年度策略报告里面提出 2023 年电子板块涨幅排名将要进入上升周期。2023 年的市场走势已经证明了我们一年前的判断：电子涨幅排名从 2022 年的倒数第一上升至 2023 年的第七名，按照 4 年周期规律，2024 年的电子涨幅排名有望继续上升。
- **AI 带动下终端销量有望回暖。**手机：Canalys 预测 2024 年全球手机销量将达到 11.7 亿部，同比增长 4%，实现由负转正。Counterpoint 预计折叠手机 2024 年全年出货量将达 3060 万台，同比增长 83%，渗透率 2.6%。服务器：Trendforce 预计全球 AI 服务器出货量 2024 年增长 27.1% 达到 150.4 万台。AI PC：群智咨询预测随着 AI CPU 和 Windows 12 的发布，明年将成为 AI PC 的规模性出货元年，全球 AI PC 整机的出货量预计将达到约 1300 万台。
- **半导体周期或已筑底，开启新一轮增长周期。**国内：集成电路产量月度增速自 2023 年 4 月转正以来，连续 7 个月实现正增长，2023 年 10 月增速达 34.5%。全球：全球半导体月度销售额同比增速在 2023 年 4 月达到最低的-21.6% 以来，降幅在持续收窄至 10 月的-0.7%。六家国际机构对 2024 年全球半导体市场规模增速预测分布在 9%-20% 之间。
- **人工智能带动存储&计算表现较好。**Semiconductor Intelligence 认为 2024 年收入增长强劲的有三类公司：1、内存公司，如三星、SK 海力士和美光。2、PC 和计算的公司。如英伟达、英特尔和 AMD。3、智能手机的公司，如高通和联发科。WSTS 预测 2024 年全球存储器增长 44.8%、逻辑增长 9.6%、微处理器/控制器增长 7.0%，模拟销售额增长 3.7%。
- **数字芯片：**AI 驱动硬件创新，边缘算力重点布局。伴随生成式 AI 的飞速发展，算力从云端向边缘端拓展将是大势所趋。**模拟芯片：**消费电子新机发布拉动需求，下游消费电子占比较多的模拟芯片公司率先迎来复苏。**功率半导体：**国产替代持续推进，关注 SiC 渗透率拐点。
- **投资建议：**在终端出货量见底回升的基本面见底的大背景下，我们认为“乐观想象”将是市场的内在逻辑，看好“空间大、渗透率低、创新产品”相关的板块估值提升。1、在国产化率低的 CPU/GPU 领域，关注国产化纯度较高的公司。2、AI PC、AI 手机等新终端带来的 SOC、存储器、先进封装的机会。3、关注苹果 MR 产品 Vision Pro 从发布到正式量产出货的机会。
- **风险提示：**1. 基本面见底，但底部时间较长。短期手机、电脑等终端出货量降幅收窄或微增长，中期出货量或保持稳定。2. 下游终端销售不及预期，向上游电子元器件厂商压价，影响毛利率。3. 公司主动去库存，导致产业链各环节订单不足、产能闲置。

## 目录

1 复盘：电子涨幅排名已进入上升期.....	1
2 半导体相关基金增多 .....	2
2.1 最近 3 年名称含半导体相关的新发基金增多 .....	2
2.2 半导体相关占比超过 25% .....	5
2.3 机构持股占比超 30% .....	5
3 下游终端：AI 带动销量有望回暖.....	6
3.1 国内手机销量有筑底回暖迹象.....	6
3.2 全球手机 2024 或出现反弹，同比增长 4%.....	7
3.3 折叠手机 2024 年有望增长 83% .....	8
3.4 全球 PC 出货量有望继续保持增长.....	9
3.5 人工智能服务器值得期待.....	11
4 半导体周期：或已筑底，开启新一轮增长周期.....	12
4.1 国内集成电路产量连续 7 个月正增长.....	12
4.2 全球半导体销售额增速有望见底.....	12
4.3 全球半导体有望回升.....	13
4.4 人工智能带动存储&计算表现较好.....	14
4.5 美光季度收入增速有望在 2024 年实现转正.....	15
5 AI 需求爆发高端显卡供不应求，国产算力技术突破产品迭代 .....	15
5.1 智能算力需求爆发，英伟达高端算力芯片供不应求.....	15
5.2 英伟达下游客户持续投入 AI 芯片自研，全球 ASIC 市场百花齐放.....	21
5.3 禁令下国产替代重要性凸显，国内厂商技术持续突破.....	23
6 数字芯片：AI 驱动硬件创新，边缘算力重点布局.....	27
6.1 AI 算力正逐渐向边缘端拓展.....	27
6.2 生成式 AI 有望促进 XR 加速创新.....	29
6.4 国产数字 SoC 厂商积极布局端侧算力.....	29
7 模拟芯片：消费电子新机发布拉动需求，关注新品新领域扩展.....	31
7.1 平台厂商发布 AI 主控芯片，有望拉动换机需求.....	31
7.2 终端库存去化尾声，消费电子需求率先复苏.....	33
7.3 关注国产模拟公司新产品和新领域扩展.....	33
8 功率半导体：国产替代持续推进，关注 SiC 渗透率拐点 .....	34
8.1 国产替代持续推进，新能源占比大幅提升.....	34
8.2 碳化硅深入布局，关注渗透率拐点.....	35
9 投资建议 .....	36
9.1 国产化纯度较高的 CPU 公司.....	36
9.2 关注 AI 终端的机会 .....	36
9.3 关注苹果 MR 产品 Vision Pro 正式量产的机会.....	37
10 风险提示 .....	37

## 插图目录

图 1 A 股行业指数年度涨幅排名 .....	1
图 2 电子板块涨幅排名 .....	2
图 3 新发基金名称含“半导体”数量（是指基金名称里面有“半导体、芯片、集成电路、半导体材料设备”） .....	2
图 4 当年新发含“半导体”基金份额&半导体指数涨跌幅（是指基金名称里面有“半导体、芯片、集成电路、半导体材料设备”） .....	3
图 5 全市场基金名称里面含“半半导体、芯片、集成电路、半导体材料设备”的基金 .....	4
图 6 2023 行业&主题指数基金新增个数 .....	5
图 7 2023 行业&主题指数基金新增份额 .....	5
图 8 机构持股占比 .....	5
图 9 月度国内手机出货量及增速 .....	6
图 10 年度国内手机出货量（万部）及增速 .....	7
图 11 2023 年全球手机出货量预测（百万部） .....	7
图 12 印度智能手机出货量 .....	8
图 13 中东(不包括土耳其)地区智能手机出货量 .....	8
图 14 华为最新款折叠屏手机 .....	8
图 15 三星最新款折叠手机 .....	8
图 16 全球折叠手机出货量及渗透率（百万） .....	9
图 17 季度全球 PC 季度出货量（百万台） .....	9
图 18 年度全球 PC 年度出货量（百万台） .....	10
图 19 年度全球 AI PC 年度出货量（百万台） .....	10
图 20 1Q23-4Q24 全球服务器出货量及增速预测 .....	11
图 21 AI 服务器出货量 .....	11
图 22 国内集成电路产量（万块）及增速（%） .....	12
图 23 全球半导体销售额月度增速（%） .....	13
图 24 全球半导体市场规模季度变化 .....	13
图 25 2023~2024 全球半导体市场规模预测 .....	14
图 26 WSTS 对 2024 全球半导体市场规模预测 .....	14
图 27 美光收入和毛利率变化周期（%） .....	15
图 28 CPU 服务器处理器出货量规模 .....	16
图 29 全球数据中心 CPU 市场份额 .....	16
图 30 算力当量与交叉熵对应关系 .....	17
图 31 大模型发布时间及对应的算力当量 .....	17
图 32 全球 AI 服务器芯片种类规模占比 .....	17
图 33 2021 年中国数据中心加速计算芯片种类 .....	17
图 34 全球 GPU 市场规模预测 .....	18
图 35 AIB 显卡市场份额变化 .....	18
图 37 英伟达 GPU 架构演进历史 .....	20
图 38 英伟达 H100 芯片成本主要构成 .....	20
图 39 2023 年终端客户对于 H100 的预计订单量 .....	21
图 40 2022-2028 全球 ASIC 市场规模 .....	22
图 41 微软 Azure Maia 100 GPU .....	22
图 42 微软 Azure Cobalt 100 CPU .....	22

图 43 谷歌 9 月 MLPerf™ 3.1 推理基准测试结果 .....	23
图 44 谷歌 11 月 MLPerf™ 3.1 训练基准测试结果 .....	23
图 45 AWS Trainium2 .....	23
图 46 中国 AI 大模型市场规模 .....	24
图 47 中国 AI 芯片市场规模 .....	24
图 48 龙芯中科高速片间互联技术——龙链技术 .....	25
图 49 思元 370 .....	26
图 50 AI 处理的重心正向边缘转移 .....	27
图 51 在以终端为中心的混合 AI 架构中云端仅用于分流处理终端无法充分运行的 AI 任务 .....	28
图 52 数量可观的生成式 AI 模型可从云端分流到终端上运行 .....	28
图 53 生成式 AI 模型将面向 XR 赋能对话式 AI 和全新渲染工具 .....	29
图 54 阿里利用 15 天时间训练出“鸟鸟分鸟” .....	30
图 55 全志科技 AI 语音专用芯片 R329 .....	30
图 56 炬芯科技最新一代高端 AI 音频芯片将于 2024 年 Sample .....	31
图 57 乐鑫科技 ESP32-S3-BOX 功能示意图 .....	31
图 58 高通骁龙 8Gen3 芯片概况 .....	32
图 59 联发科天玑 9300 芯片概况 .....	33
图 60 2022Q1-2023Q3 下游消费类收入较多的模拟芯片公司营业收入 .....	33
图 61 斯达半导 2022 年各下游领域收入占比 .....	35
图 62 斯达半导 2023H1 各下游领域收入占比 .....	35
图 63 2021 年-2027E 碳化硅各下游市场规模 .....	35
图 64 国产化重要指标 .....	36

## 表格目录

表 1 算力分类 .....	15
表 2 英伟达高算力芯片：数据中心 GPU .....	19
表 3 Instinct MI300X VS H100 .....	21
表 4 海光深算一号性能对比 .....	26

## 1 复盘：电子涨幅排名已进入上升期

2023 电子涨幅排名大幅上升，新周期已开启。从 A 股行业指数年度涨幅排名看，电子板块排名垫底年份的下一年有较好的相对排名。除了 2006 年之外，其他年份只要当年电子板块涨幅排名进入倒数前三名，下一年的相对排名表现较好。截至 2023 年 11 月 30 日，电子指数排名第七。

图 1 A 股行业指数年度涨幅排名

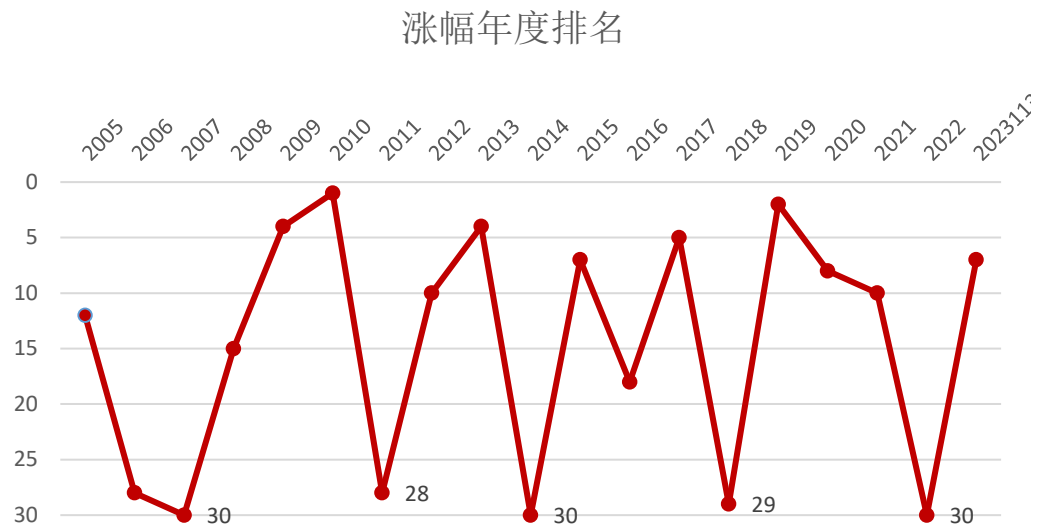
2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	20231130
军工	非银	煤炭	电新	汽车	电子	银行	房地产	传媒	非银	计算机	食品饮料	食品饮料	消费	食品饮料	电新	电新	煤炭	传媒
银行	银行	有色	医药	有色	医药	食品饮料	非银	计算机	建筑	消费	家电	家电	银行	电子	食品饮料	化工	交通	通信
食品饮料	食品饮料	计算机	建筑	煤炭	计算机	煤炭	建筑	通信	房地产	通信	银行	银行	煤炭	石化	家电	消费	有色	消费
房地产	有色	综合	农业	电子	机械	消费	银行	电子	银行	轻工	煤炭	非银	石化	建材	军工	煤炭	综合	煤炭
零售	房地产	化工	通信	家电	食品饮料	公用	综合	军工	交通	纺织	石化	电子	农业	农业	医药	钢铁	石化	计算机
石化	上证	建筑	家电	计算机	军工	房地产	家电	家电	钢铁	农业	建筑	银行	非银	非银	汽车	公用	房地产	家电
上证	机械	建材	公用	建材	有色	上证	有色	医药	公用	电子	钢铁	钢铁	上证	计算机	化工	建筑	建筑	电子
有色	零售	纺织	食品饮料	农业	石化	医药	医药	电新	军工	化工	有色	建材	计算机	医药	电子	汽车	农业	石化
消费	石化	医药	传媒	消费	电新	煤炭	公用	综合	计算机	综合	建材	有色	医药	通信	机械	综合	纺织	纺织
化工	钢铁	家电	化工	房地产	建材	家电	电子	轻工	上证	传媒	汽车	上证	军工	轻工	有色	电子	公用	非银
电新	军工	钢铁	建材	轻工	消费	纺织	汽车	消费	综合	电新	化工	交通	房地产	消费	家电	石化	零售	机械
电子	建材	交通	军工	综合	纺织	轻工	消费	零售	消费	机械	医药	非银	医药	建筑	银行	建材	机械	银行
传媒	传媒	电新	零售	零售	零售	钢铁	上证	农业	机械	医药	纺织	石化	公用	房地产	轻工	轻工	有色	建筑
交通	汽车	汽车	纺织	纺织	传媒	通信	建材	汽车	有色	房地产	上证	房地产	钢铁	机械	计算机	军工	机械	银行
医药	农业	机械	电子	机械	家电	医药	石化	公用	纺织	汽车	轻工	通信	煤炭	传媒	农业	建材	上证	上证
家电	电新	公用	机械	食品饮料	综合	建筑	化工	化工	汽车	家电	医药	汽车	交通	有色	上证	纺织	医药	钢铁
建材	消费	非银	综合	银行	汽车	交通	食品饮料	纺织	零售	零售	农业	消费	建材	化工	钢铁	传媒	通信	医药
通信	通信	零售	石化	非银	化工	轻工	石化	电子	电新	建材	电子	公用	家电	军工	煤炭	上证	汽车	轻工
非银	建筑	轻工	房地产	医药	通信	化工	交通	交通	建材	军工	零售	建筑	零售	电新	交通	计算机	非银	军工
农业	医药	房地产	计算机	化工	建筑	建材	钢铁	机械	通信	交通	机械	电新	汽车	上证	非银	交通	钢铁	有色
机械	化工	农业	轻工	钢铁	轻工	汽车	农业	银行	化工	食品饮料	电新	化工	纺织	汽车	石化	通信	计算机	零售
公用	纺织	消费	银行	农业	煤炭	农业	煤炭	非银	煤炭	公用	综合	综合	通信	交通	公用	零售	化工	食品饮料
汽车	轻工	食品饮料	上证	通信	石化	零售	机械	建材	轻工	有色	公用	机械	轻工	综合	零售	银行	军工	交通
建筑	综合	军工	消费	电新	公用	机械	传媒	上证	家电	建筑	通信	零售	电新	煤炭	银行	医药	电新	农业
纺织	交通	石化	煤炭	上证	交通	综合	军工	食品饮料	农业	上证	房地产	轻工	化工	纺织	传媒	食品饮料	家电	综合
钢铁	煤炭	通信	非银	交通	上证	军工	零售	建筑	石化	银行	军工	农业	机械	石化	纺织	农业	轻工	化工
综合	家电	上证	汽车	石化	房地产	计算机	纺织	房地产	食品饮料	石化	交通	军工	传媒	零售	综合	房地产	食品饮料	建材
计算机	电子	银行	交通	传媒	非银	电子	计算机	钢铁	传媒	钢铁	消费	计算机	有色	公用	建筑	家电	建材	房地产
轻工	公用	传媒	钢铁	公用	钢铁	有色	电新	有色	医药	煤炭	计算机	传媒	电子	钢铁	通信	非银	传媒	电新
煤炭	计算机	电子	有色	建筑	银行	电新	通信	煤炭	电子	非银	传媒	纺织	综合	建筑	房地产	消费	电子	消费

资料来源：wind，首创证券，注：因“综合金融”只有 11 只个股，故中信一级行业分类删除“综合金融”

可以从下图中看出，电子板块的涨幅排名变化具有明显的周期性规律。当年排名垫底，下一年排名较好。从 2007 年至今，电子板块涨幅相对排名呈现约 4 年的周期性 2007~2011（4 年）、2011~2014（3 年）、2014~2018（4 年）、2018~2022（4 年），2023 年电子板块涨幅排名大幅上升，从去年的 30 名上升到 2023 年的第七名，周期或已开启。



图 2 电子板块涨幅排名



资料来源: wind, 首创证券

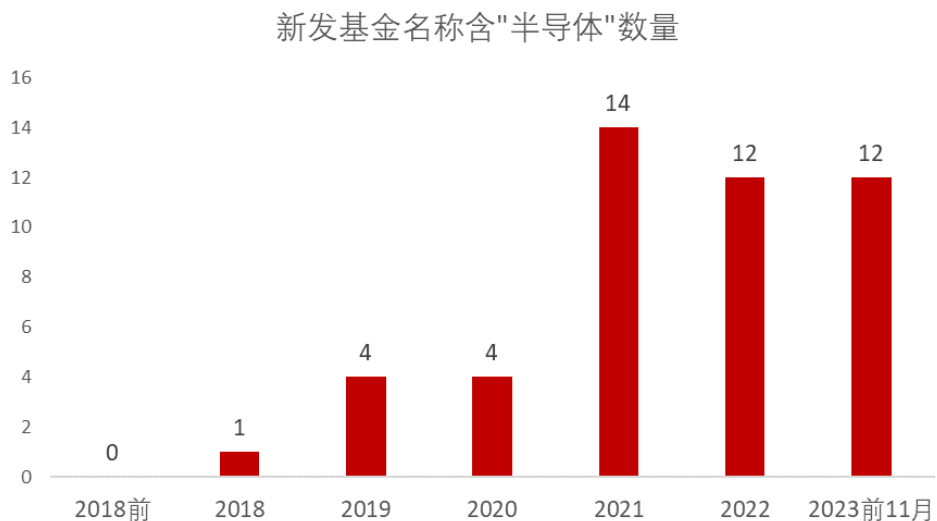
## 2 半导体相关基金增多

### 2.1 最近 3 年名称含半导体相关的新发基金增多

2018 年之前基金名称里面含“半导体”数量(是指基金名称里面有“半导体、芯片、集成电路、半导体材料设备”)的基金是零。

2018~2023 年前 11 个月, 新发基金名称里面含“半导体”的基金个数分别为 1、4、4、14、12、12 只。

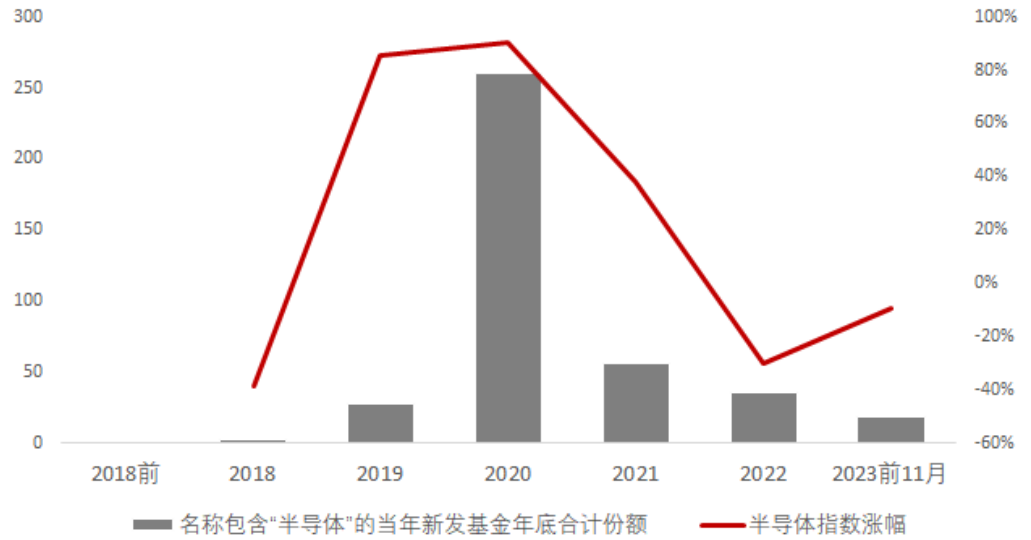
图 3 新发基金名称含“半导体”数量(是指基金名称里面有“半导体、芯片、集成电路、半导体材料设备”)



资料来源: wind, 首创证券

2018 年年底基金名称含“半导体”的基金份额 2 亿份, 2019~2023 前 11 个月, 名称包含“半导体”的当年新发基金年底合计份额分别为 28 亿份、260 亿份、56 亿份、35 亿份、16 亿份。

图 4 名称包含“半导体”的当年新发基金年底合计份额&半导体指数涨跌幅（“半导体”是指基金名称里面有“半导体、芯片、集成电路、半导体材料设备”）



资料来源: wind, 首创证券

截至 2023 年 11 月底, 全市场基金名称里面含“半导体、芯片、集成电路、半导体材料设备”的基金一共有 47 只, 合计规模 1302 亿元。

图 5 全市场基金名称里面含“半导体、芯片、集成电路、半导体材料设备”的基金

基金代码	基金名称	最新基金规模合计 (元)
006502.OF	财通集成电路产业A	104,874,905
512480.OF	国联安中证全指半导体ETF	26,812,204,054
512760.OF	国泰CES半导体芯片ETF	18,001,460,254
007300.OF	国联安中证全指半导体ETF联接A	4,213,806,227
008281.OF	国泰CES半导体芯片行业ETF联接A	6,157,506,151
159995.OF	华夏国证半导体芯片ETF	26,656,260,101
159801.OF	广发国证半导体芯片ETF	2,784,646,249
159813.OF	鹏华国证半导体芯片ETF	4,557,315,273
008887.OF	华夏国证半导体芯片ETF联接A	9,801,707,568
012650.OF	博时半导体主题A	1,706,469,105
012552.OF	天弘中证芯片A	922,215,944
516920.OF	汇添富中证芯片产业ETF	407,848,584
012837.OF	华安CES半导体芯片行业A	215,275,831
012629.OF	广发国证半导体芯片ETF联接A	932,761,963
516640.OF	富国中证芯片产业ETF	1,293,035,932
012969.OF	鹏华国证半导体芯片ETF联接A	907,264,581
013445.OF	西藏东财中证芯片产业A	333,265,743
013339.OF	创金合信芯片产业A	582,704,631
014193.OF	汇添富中证芯片产业指数增强A	337,628,114
013840.OF	银华集成电路A	3,539,778,291
516350.OF	易方达中证芯片产业ETF	312,220,614
014418.OF	西部利得CES半导体芯片行业指数增强A	704,314,511
014319.OF	德邦半导体产业A	2,866,231,189
014776.OF	富国中证芯片产业联接A	281,273,123
015336.OF	嘉实中证芯片产业A	231,256,696
014854.OF	嘉实中证半导体增强A	1,586,344,896
016500.OF	华夏半导体龙头A	765,633,335
015967.OF	永赢半导体产业智选A	53,776,876
588200.OF	嘉实上证科创板芯片ETF	7,440,495,253
588290.OF	华安上证科创板芯片ETF	1,315,883,682
513310.OF	华泰柏瑞中证韩交所中韩半导体ETF	121,781,144
017075.OF	宝盈半导体产业A	265,261,922
017469.OF	嘉实上证科创板芯片联接A	808,482,767
159665.OF	工银瑞信国证半导体芯片ETF	517,538,795
017559.OF	华安上证科创板芯片联接A	202,765,108
017653.OF	创金合信全球芯片产业A	19,572,499
501225.OF	景顺长城全球半导体芯片产业A	244,088,779
018000.OF	鹏华芯片产业A	16,116,320
018411.OF	易方达中证芯片产业联接A	43,148,702
159516.OF	国泰中证半导体材料设备主题ETF	797,062,836
561980.OF	招商中证半导体产业ETF	311,966,675
019454.OF	华泰柏瑞中韩半导体联接A	10,095,181
019632.OF	国泰中证半导体材料设备主题联接A	10,000,000
562590.OF	华夏中证半导体材料设备主题ETF	229,872,223
159546.OF	国泰中证全指集成电路ETF	406,373,772
159560.OF	景顺长城中证芯片产业ETF	441,779,065
019759.OF	中欧半导体产业A	10,410,868
	规模合计 (元)	130,281,776,333

资料来源: wind, 首創證券



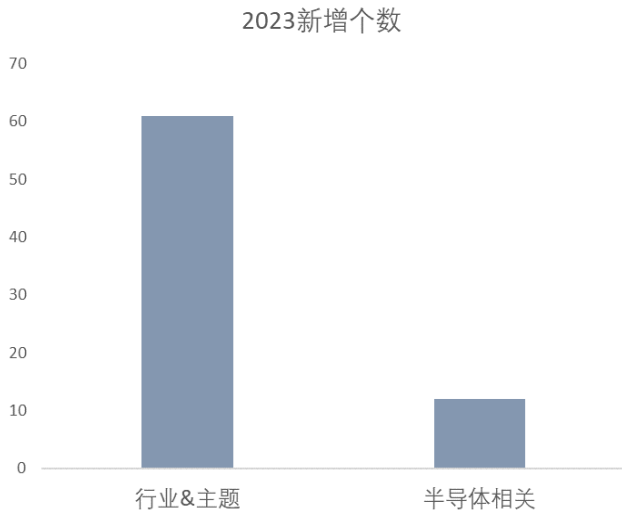
## 2.2 半导体相关占比超过 25%

按新增基金个数统计，截至 2023 年 11 月 30 日，全市场行业指数基金&主题指数基金新增个数为 61 只，其中半导体相关新增 12 只，占有类型的新增基金数量的比例为 20%。

按新增基金份额统计，截至 2023 年 11 月 30 日，全市场行业指数基金&主题指数基金份额新增 2066 亿份，其中半导体相关新增 538 亿份，占比 26%。

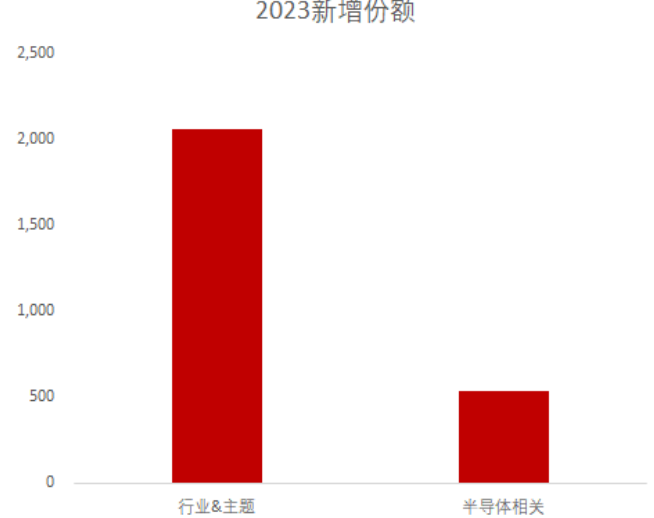
从半导体相关基金占比看，半导体较受投资人欢迎。若保持新发基金的占比，半导体板块的股价表现有望强于其他板块，或整体市场。

图 6 2023 行业&主题指数基金新增个数



资料来源: wind, 首创证券

图 7 2023 行业&主题指数基金新增份额

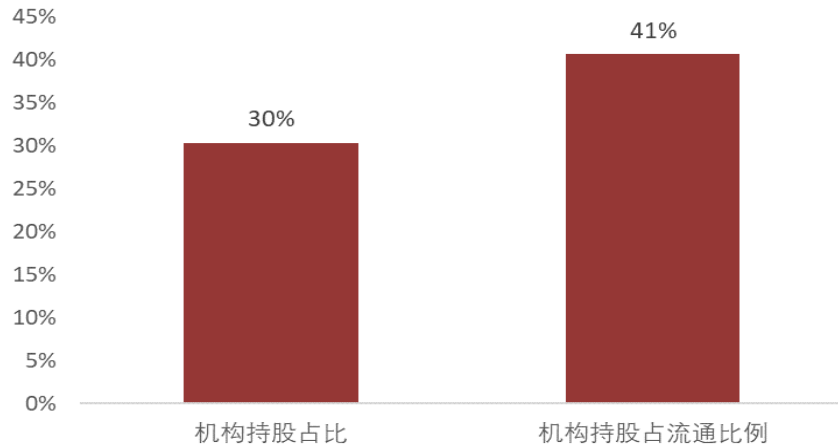


资料来源: wind, 首创证券

## 2.3 机构持股占比超 30%

截至 2023 年三季报，半导体板块（SW 分类）市值 30618.8 亿元，机构持股总市值 9285 亿元，占比 30.3%。按照流通股比例看，机构持股占半导体板块流通比例达到 41%。

图 8 机构持股占比



资料来源: wind, 首创证券

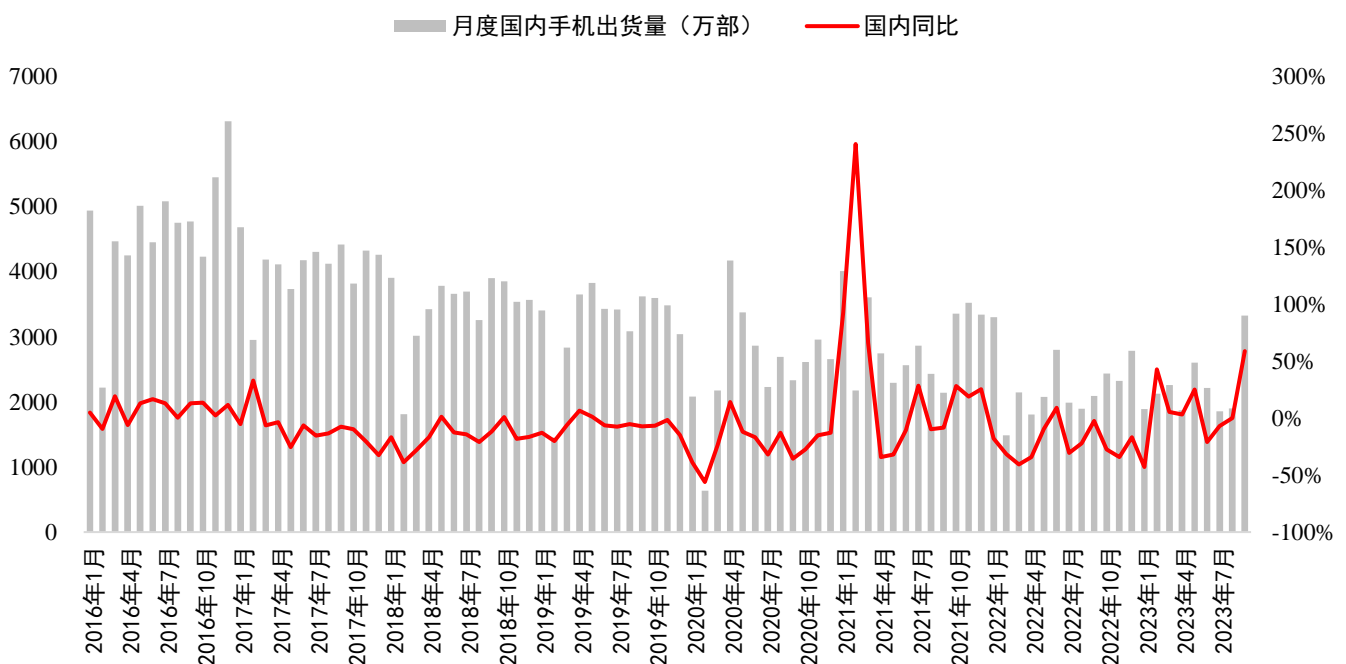
### 3 下游终端：AI 带动销量有望回暖

#### 3.1 国内手机销量有筑底回暖迹象

从月度视角看，国内手机出货量从 2023 年 1 月的 1890 万部波动上升至九月的 3328 万部，同比增速从年初的-42.8%提升至 59.1%，实现由负转正。

2016 年~2020 年，国内手机销量总体呈现减少趋势。从月度出货量及增速看，2020~2022 三年国内手机出货量横盘波动，从 2023 年三季度出货量看，国内手机市场或已完成筑底，有回暖迹象。

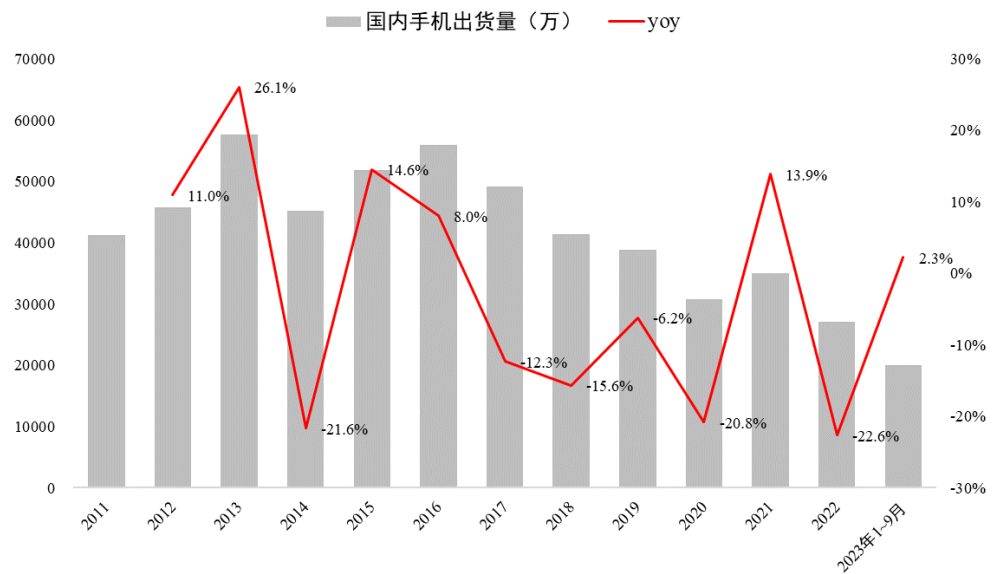
图 9 月度国内手机出货量及增速



资料来源：wind，首创证券

从年度视角看，国内手机出货量存在周期性波动，2022 年度国内手机出货量达 27151 万部，同比减少 22.6%，2023 年 1-9 月实现出货量 20047 万部，同比增长 2.3%，2023 年度有望实现同比再度增长。

图 10 年度国内手机出货量（万部）及增速

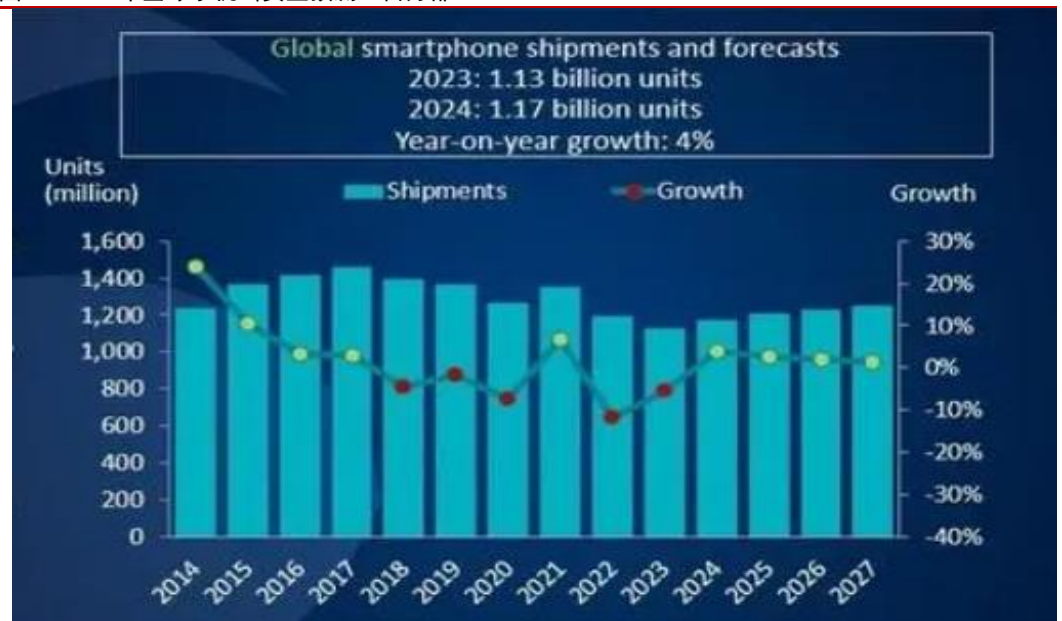


资料来源: wind, 首创证券

### 3.2 全球手机 2024 或出现反弹，同比增长 4%

根据 Canalys 数据，预计 2023 年度全球手机将实现销量 11.3 亿部，同比下降 5%，实现同比进一步修正。报告预测 2024 年将会成为转折点，预计手机销量将达到 11.7 亿部，同比增长 4%，实现由负转正。

图 11 2023 年全球手机出货量预测（百万部）



资料来源: Canalys, 首创证券

分地区看，根据 Canalys 数据显示，印度第三季度智能手机出货量为 4300 万部，同比降幅低达 3%，市场复苏迹象明显。2023 年第二季度中东地区（不含土耳其）智能手机出货量达 950 万部，同比增长 2%，相较于全球市场同期下降 10%，增长势头良好。

图 12 印度智能手机出货量



资料来源: Canlys, 首创证券

图 13 中东(不包括土耳其)地区智能手机出货量



资料来源: Canlys, 首创证券

### 3.3 折叠手机 2024 年有望增长 83%

折叠手机市场逐渐高端化, 根据京东官网, 华为新品 Mate X5 折叠手机 12+512G 售价 13999 元, 销量较好的 Mate X3 1TB 售价 12694 元。三星新品 Galaxy Z Fold5 售价 10799 元, W24 Flip 12+512G 售价 8699 元。

图 14 华为最新款折叠屏手机

**HUAWEI Mate X5**

四曲折叠机身 | 玄武钢化昆仑玻璃\* | 高分辨率微珠双摄

10-08/18-08 限时抢购

¥13999.00

华为 (HUAWEI) Mate X5 折叠屏手机 12GB+512GB 羽砂黑 1万+条评价

京东自营官方旗舰店

自营 新品

**HUAWEI Mate X3**

银盾时代 轻薄如直板机 超可靠昆仑玻璃\* 超强灵犀通信

¥12694.00

华为/HUAWEI Mate X3 折叠屏手机 超轻薄 超可靠昆仑玻璃 超强灵犀通信 1TB 青 1万+条评价

京东自营官方旗舰店

自营

资料来源: 京东 2023 年 11 月 26 日价格, 首创证券

图 15 三星最新款折叠手机

**Galaxy Z Fold5**

(12G+256G) (12G+512G) (12G+1TB)

品质保障 超薄合折 轻薄手感 PC级强大生产力 辅助性能 保障游戏体验 6+ 续航

¥10799.00

三星 SAMSUNG Galaxy Z Fold5 超闭合折叠 IPX8级防水 5G折叠手机 宇夜黑 100+条评价

寰悦数码电玩买手店

**SAMSUNG 三星 W24 Flip**

16GB+512GB 超薄合精工玻璃 贴合屏幕 新品上市 支持全国联保

¥8699.00

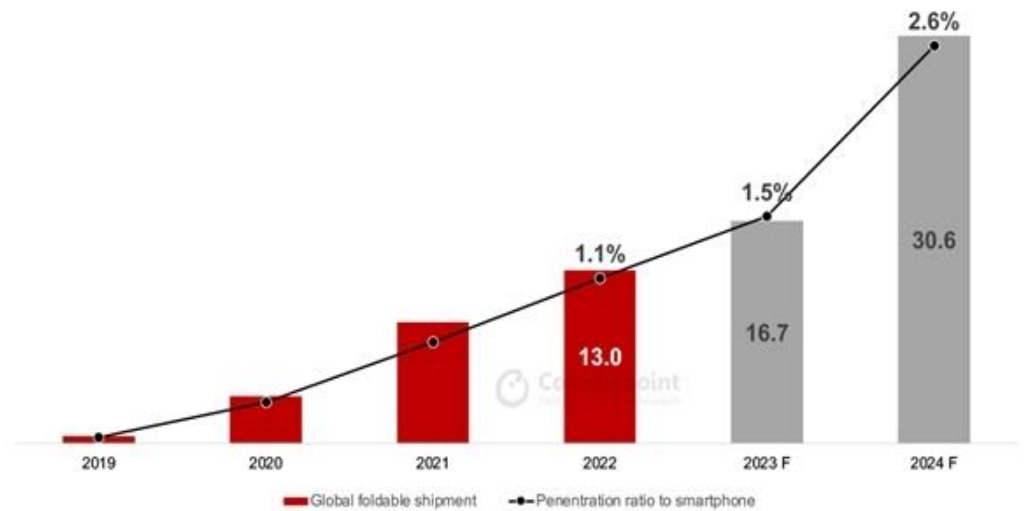
三星 SAMSUNG W24 Flip 12GB+512GB 冰瓷白 大视野智能外屏 掌心 折叠屏手机 0条评价

寰悦数码电玩买手店

资料来源: 京东 2023 年 11 月 26 日价格, 首创证券

根据 Counterpoint Research 数据, 2022 年度全球折叠手机出货量达 1300 万台, 对于智能手机整体市场渗透率为 1.1%。预计 2023-2024 年全年出货量将达到 1670 和 3060 万台, 2024 年出货量同比增长 83%, 渗透率分别达到 1.5%和 2.6%。

图 16 全球折叠手机出货量及渗透率（百万）

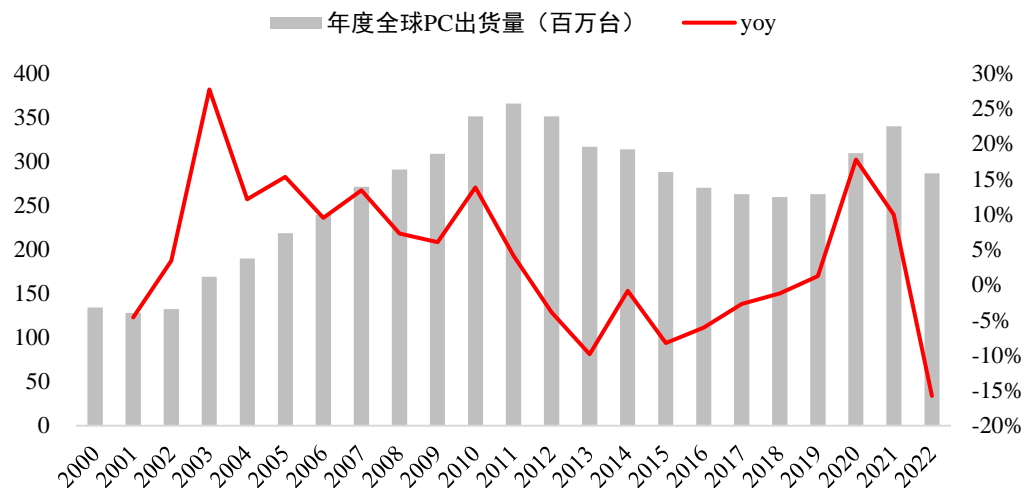


资料来源：Counterpoint Research，首创证券

### 3.4 全球 PC 出货量有望继续保持增长

从年度视角看，2022 年度全球 PC 出货量为 2.86 亿台，同比减少 15.77%，相较于 2021 年度 3.4 亿台出货量以及 9.93% 的同比增长，增速明显放缓。

图 17 季度全球 PC 季度出货量（百万台）

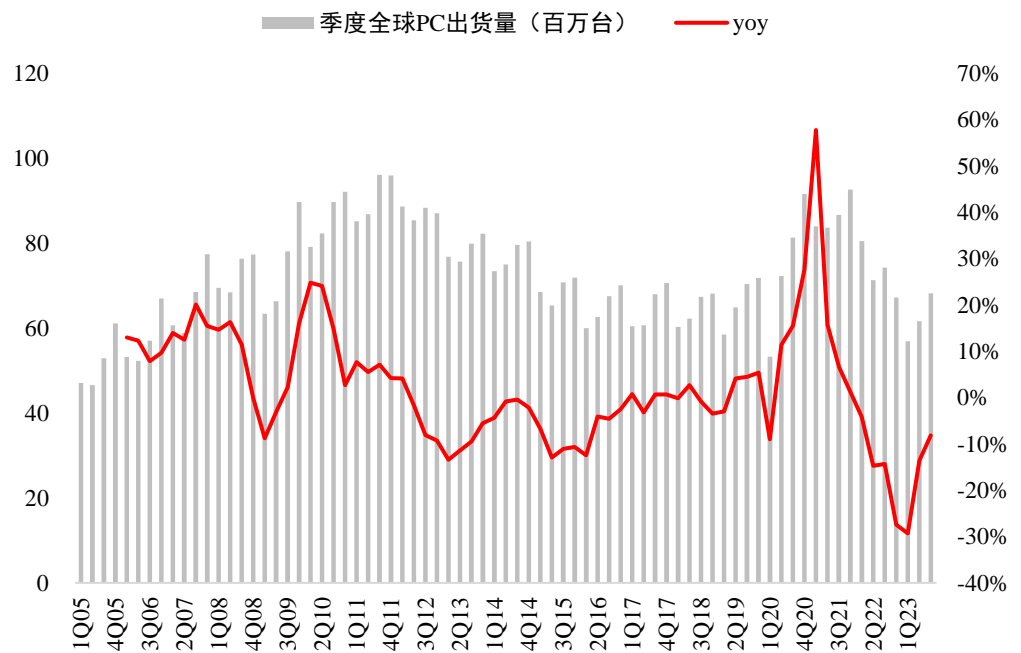


资料来源：Wind，首创证券

进入 2023 年，2023 年度第一到三季度全球 PC 出货量分别为 5690、6160 和 6820 万台，呈现增长态势，同比增速由第一季度的 -29.3% 逐步修正至第三季度的 -8.2%，降幅逐渐收窄。



图 18 年度全球 PC 年度出货量（百万台）

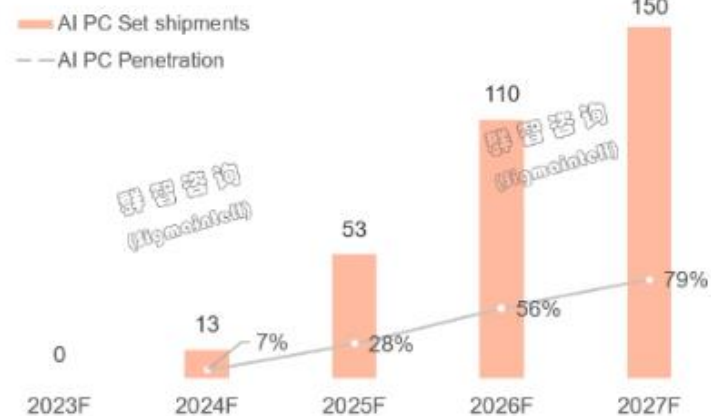


资料来源：IDC，首创证券

根据群智咨询的预测，随着 AI CPU 和 Windows 12 的发布，明年将成为 AI PC 的规模性出货元年，全球 AI PC 整机的出货量预计将达到约 1300 万台。

图 19 年度全球 AI PC 年度出货量预测（百万台）

Y2023~Y2028 Global AI PC Shipments Trend (Units: M sets; %)



Remark: The PC that can provide hybrid (integrated software and hardware) intelligent learning and reasoning capabilities, is defined as AI PCs according to Sigmaintell.

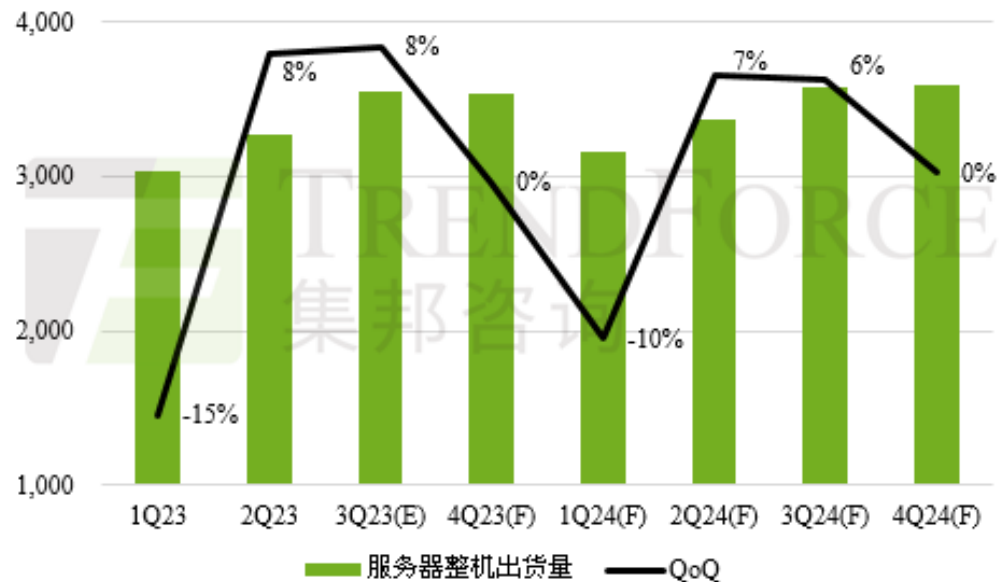
资料来源：群智咨询，首创证券

### 3.5 人工智能服务器值得期待

2023 年度第一季度全球服务器整机出货量超过 300 万台，且在第二季度和第三季度都保持了 8% 的环比增速。根据 TrendForce 预测，这一数据将在 2024 年度保持较为稳定的趋势，特别是 2024 年 Q2~Q3 环比增速有望超过 5%。

图 20 1Q23-4Q24 全球服务器出货量及增速预测

图、1Q23~4Q24全球服务器整机出货量预估（单位：千台）

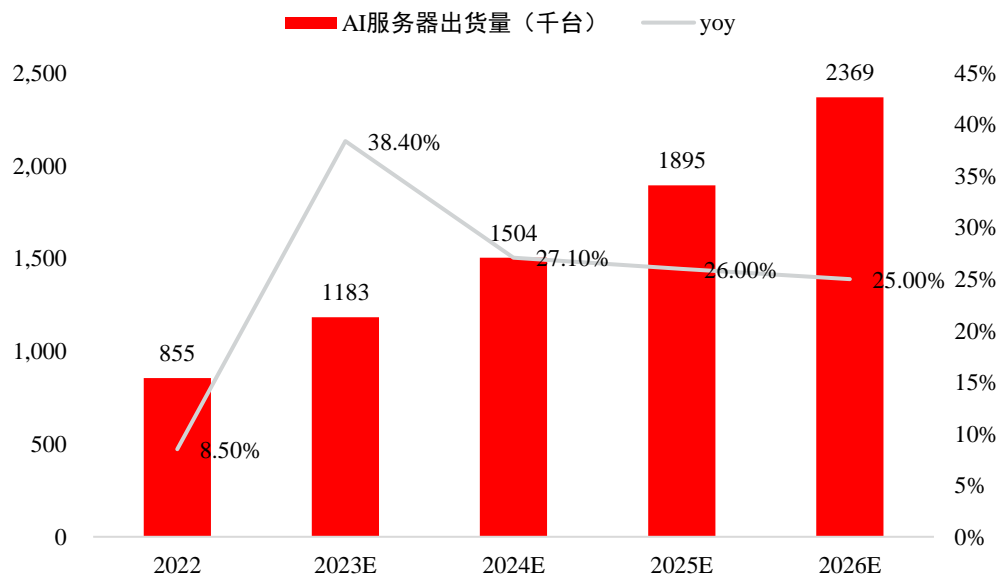


Source: TrendForce, Aug., 2023

资料来源: TrendForce, 首创证券

伴随生成式 AI 在各应用领域加速渗透, AI 服务器需求持续上扬。根据 Trendforce, 2022 年全球 AI 服务器出货量为 85.5 万台, 预计 2023 年将达到 118.3 万台, 预计 2024 年增长 27.1% 达到 150.4 万台, 2026 年在 23 年基础上翻倍, 达到 236.9 万台。

图 21 AI 服务器出货量



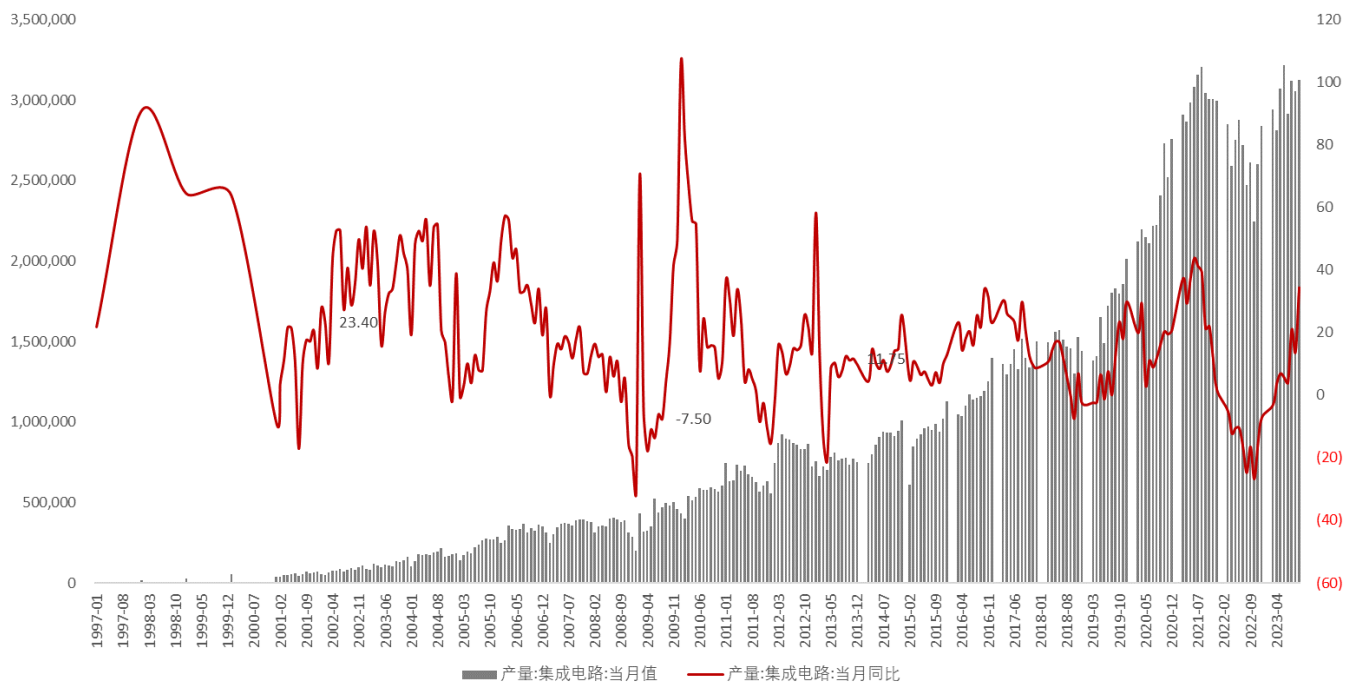
资料来源: Trendforce, 首创证券

## 4 半导体周期：或已筑底，开启下一轮增长周期

### 4.1 国内集成电路产量连续 7 个月正增长

从国内集成电路增速看，半导体周期或已经见底。国内集成电路产量月度增速自 2023 年 4 月转正以来，连续 7 个月份实现正增长，并且增速在提升。2023 年 10 月国内集成电路产量 312.8 亿块，同比增长 34.5%。

图 22 国内集成电路产量（万块）及增速（%）



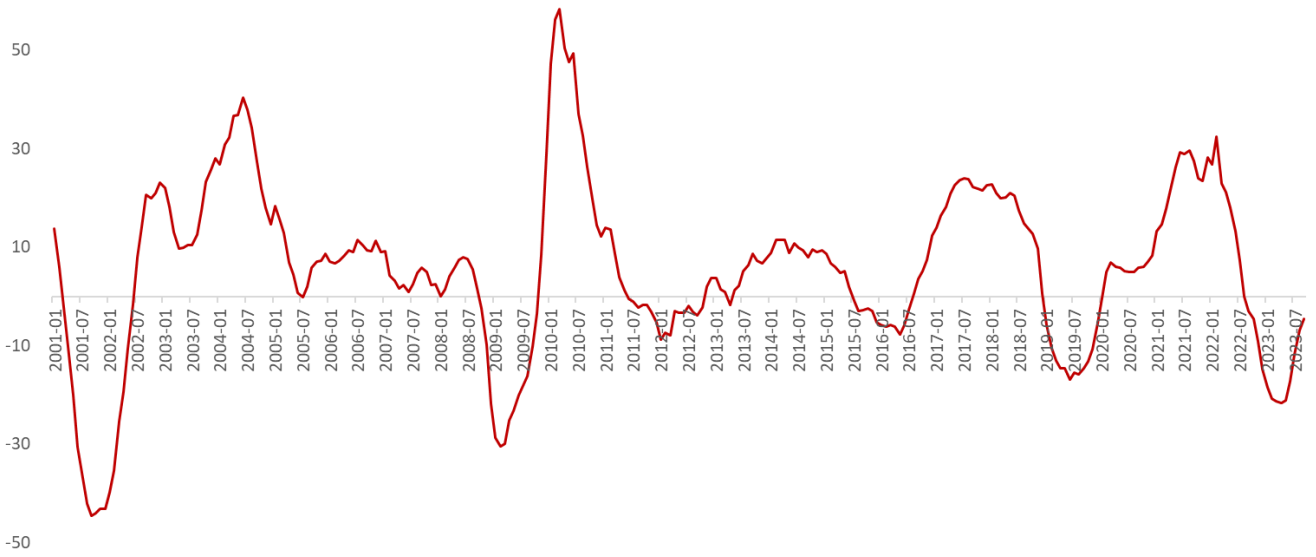
资料来源: wind, 国家统计局, 首创证券

### 4.2 全球半导体销售额增速有望见底

从全球范围看，随着时间推移半导体销售额增速有筑底的可能性在增加。根据半导体产业协会(SIA)公布的 2023 年 10 月全球半导体产业销售总额为 466 亿美元，较 2023 年 9 月的 449 亿美元增长 3.9%，同比降幅继续收窄至-0.7%。

从全球半导体月度销售额增速变化趋势看，同比增速在 2023 年 4 月达到最低的-21.6%以来，降幅在持续收窄。2023 年 4 月-10 月增速分别为-21.6%、-21.1%、-17.3%、-11.8%、-6.8%、-4.5%、-0.7%。

图 23 全球半导体销售额月度增速 (%)

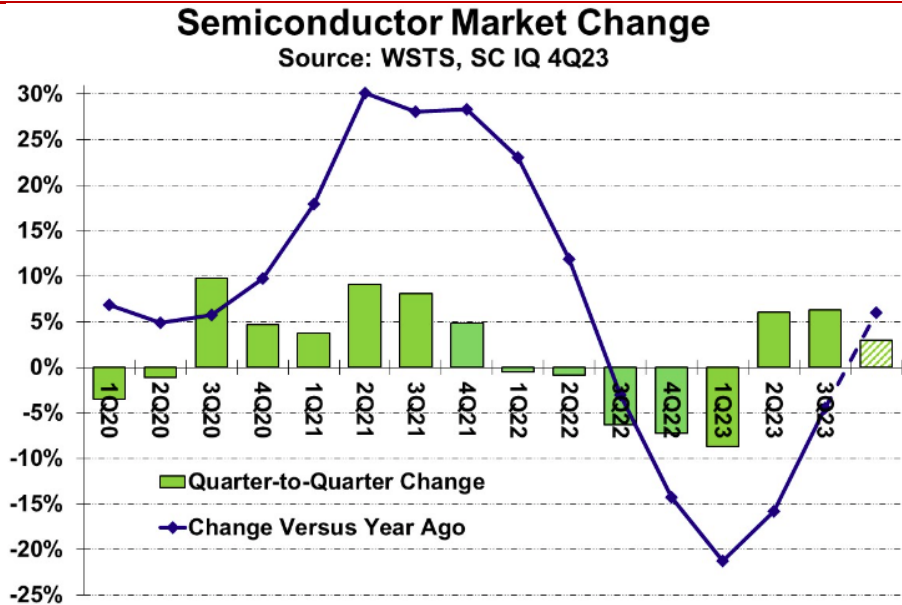


资料来源: wind, 首创证券

### 4.3 全球半导体有望回升

全球半导体市场目前正稳步好转。WSTS 将 2023 年第 2 季度的数据从之前的 4.2% 修订为 6.0%。2023 年第 3 季度较 2023 年 2 季度增长 6.3%。根据 Semiconductor Intelligence 2023 年第四季度增长 3% 的预测, 2024 年第一季度的同比增长率将为正 6%。这将为 2024 年的同比增长奠定基础。

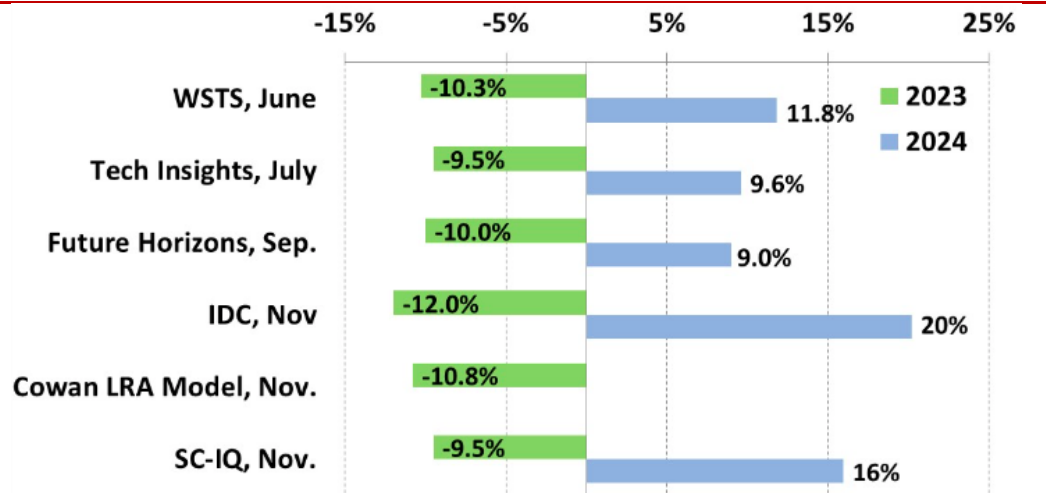
图 24 全球半导体市场规模季度变化



资料来源: Semiconductor Intelligence, 首创证券

进入 2023 年年末, 国际第三方机构对 2024 年全球半导体市场规模做出预测, 六家机构对 2024 年的增速预测分布在 9%-20% 之间。相比较 2023 年增速有明显回升。

图 25 2023~2024 全球半导体市场规模预测



资料来源: Semiconductor Intelligence, 首创证券

#### 4.4 人工智能带动存储&计算表现较好

Semiconductor Intelligence 认为 2024 年收入增长最强劲的有三类半导体公司:

1、内存公司,如三星、SK 海力士和美光。2、PC 和计算的公司。如英伟达、英特尔和 AMD。3、智能手机的公司,如高通和联发科。而主要专注于汽车、工业和物联网的公司,如 TI、Infineon、ST、NXP 和 Analog Devices 的收入增长相对较少。

WSTS 预测 2024 全球半导体市场 5883.64 亿美元,比上年增长 13.1%。WSTS 同样看好生成型人工智能相关产品需求持续增长。

根据 WSTS 预测,2024 年按产品类别划分的市场规模如下:分立半导体将增长 4.2% 至 375 亿美元,光电子将增长 1.7%至 433 亿美元,传感器将增长 3.7%至 201 亿美元,整体 IC 将同比增长 15.5%至 4875 亿美元。其中:存储器增长 44.8%、逻辑增长 9.6%、微处理器/控制器增长 7.0%,模拟销售额增长 3.7%。

图 26 WSTS 对 2024 全球半导体市场规模预测

Fall 2023	Amounts in US\$M			Year on Year Growth in %		
	2022	2023	2024	2022	2023	2024
Americas	141,136	132,536	162,154	16.2	-6.1	22.3
Europe	53,853	57,048	59,480	12.8	5.9	4.3
Japan	48,158	47,209	49,275	10.2	-2.0	4.4
Asia Pacific	330,937	283,333	317,455	-3.5	-14.4	12.0
<b>Total World - \$M</b>	<b>574,084</b>	<b>520,126</b>	<b>588,364</b>	<b>3.3</b>	<b>-9.4</b>	<b>13.1</b>
Discrete Semiconductors	33,993	35,951	37,459	12.0	5.8	4.2
Optoelectronics	43,908	42,583	43,324	1.2	-3.0	1.7
Sensors	21,782	19,417	20,127	13.7	-10.9	3.7
Integrated Circuits	474,402	422,174	487,454	2.5	-11.0	15.5
Analog	88,983	81,051	84,056	20.1	-8.9	3.7
Micro	79,073	76,579	81,937	-1.4	-3.2	7.0
Logic	176,578	174,944	191,693	14.0	-0.9	9.6
Memory	129,767	89,601	129,768	-15.6	-31.0	44.8
<b>Total Products - \$M</b>	<b>574,084</b>	<b>520,126</b>	<b>588,364</b>	<b>3.3</b>	<b>-9.4</b>	<b>13.1</b>

资料来源: WSTS, 首创证券



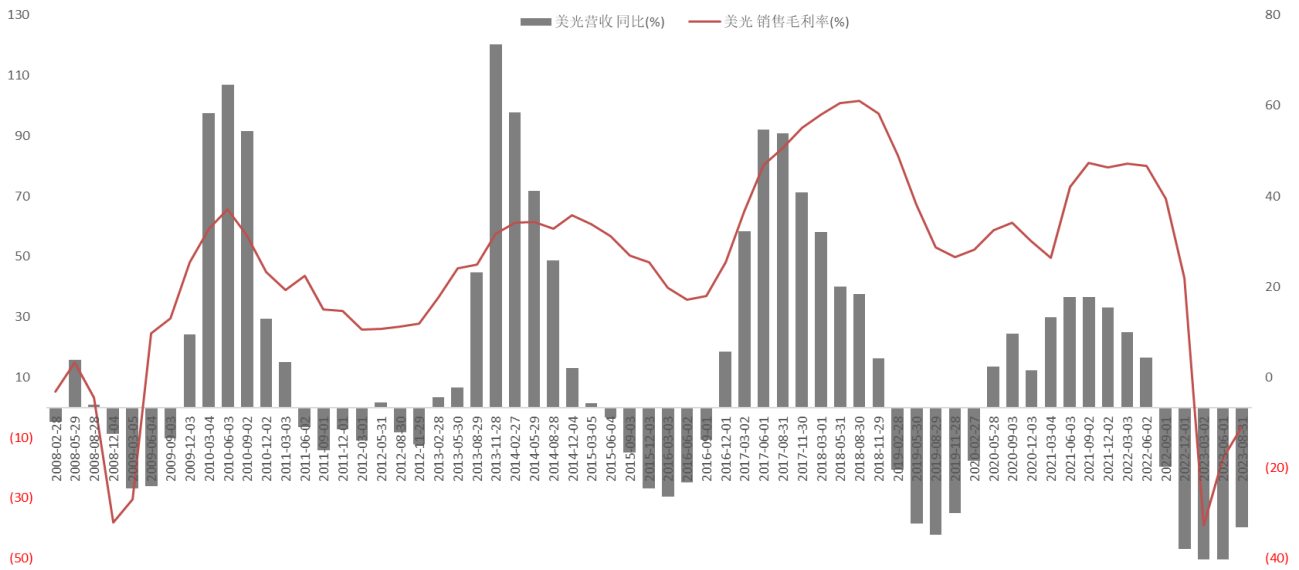
#### 4.5 美光季度收入增速有望在 2024 年实现转正

美光科技从事 DRAM、NAND 闪存业务，美光自 2008 年以来，季度收入增速和毛利率变化趋势有明显的周期性，并且收入增速和毛利率保持同步变化趋势。

2008 年至今有 4 个明显负增长的区间，分别是 2008Q4~2009Q3、2011Q2~2012Q4、2015Q2~2016Q3、2019Q1~2020Q1，每次负增长的时间长度分别为 4 个季度、7 个季度、6 个季度、5 个季度。

从 2022Q3 开始的负增长已经持续 5 个季度，从收入增速规律看，按照过去的规律，美光的收入增速有望在 2024 年实现季度正增长。

图 27 美光收入和毛利率变化周期 (%)



资料来源: wind, 首创证券

### 5 AI 需求爆发高端显卡供不应求，国产算力技术突破产品迭代

#### 5.1 智能算力需求爆发，英伟达高端算力芯片供不应求

算力是对信息数据进行处理并输出目标结果的计算能力，主要通过 CPU、GPU、FPGA、ASIC 等各类计算芯片实现。根据中国信通院的分类，算力可以分为通用算力、智能算力和超算算力。其中，通用算力基于搭载 CPU 芯片的服务器，2022 年我国通用服务器出货量为 384.6 万台，同比增长 3%；智能算力基于搭载 GPU、FPGA、ASIC 等芯片的加速计算平台，2022 年 AI 服务器出货量 28 万台，同比增长 23%；超算算力基于超级计算机等高性能计算集群，用于处理极端复杂或数据密集型问题。

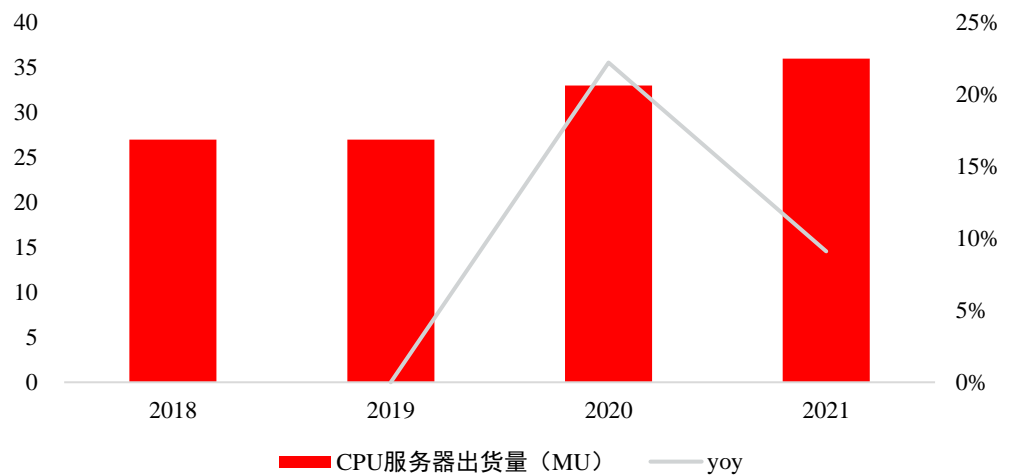
表 1 算力分类

	定义	适用场景
通用算力	基于 CPU 芯片的服务器所提供的计算能力。	执行大部分通用计算任务和操作系统管理。
智能算力	基于 GPU、FPGA、ASIC 等芯片的加速计算平台所提供的计算能力。	提供人工智能训练和推理的计算能力。
超算算力	是基于超级计算机等高性能计算集群所提供的计算能力。	主要运用于尖端科研、国防军工等大科学、大工程、大系统中。

资料来源: 中国信通院, 首创证券

在通用算力方面，全球 CPU 服务器处理器出货量稳定增长。根据中国信通院数据，2021 年全球 CPU 服务器的处理器出货量为 3600 万 (unit)，同比增长 10%。

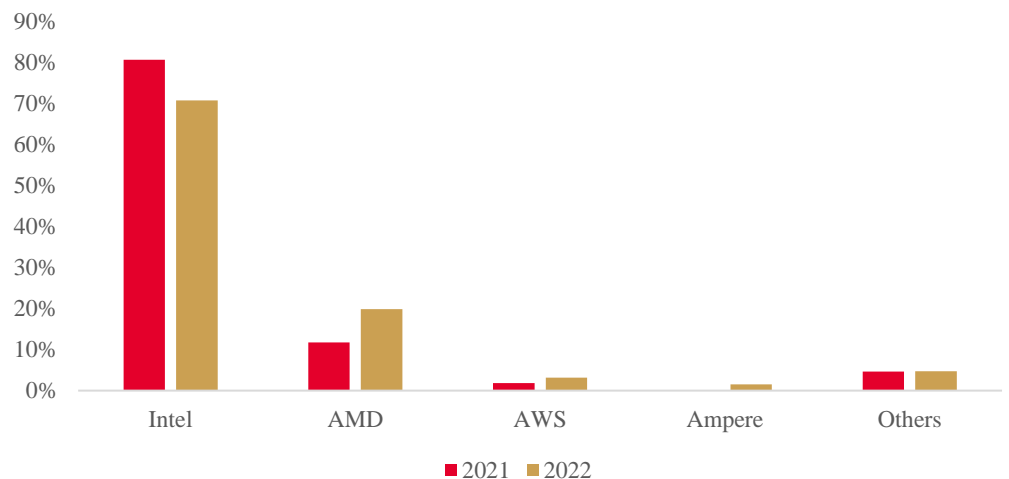
图 28 CPU 服务器处理器出货量规模



资料来源：中国算力白皮书（2022 年），首创证券

在全球数据中心 CPU 市场，基于 X86 架构的 Intel 和 AMD 占据市场主导地位。根据 Counterpoint，2022 年 Intel 以 70.77% 的市场份额绝对领先其他对手，但是相较 2021 年市场份额下降了近 10 个百分点；AMD 以 19.84% 的市场份额位列第二，较 2021 年提升 8.1pp。同时，基于 ARM 架构的处理器市场份额不断提升，2022 年收入首次超过 10 亿美元。

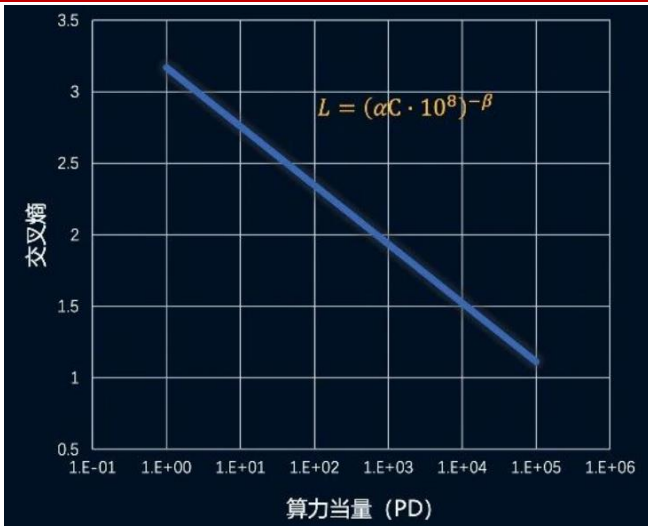
图 29 全球数据中心 CPU 市场份额



资料来源：Counterpoint，首创证券

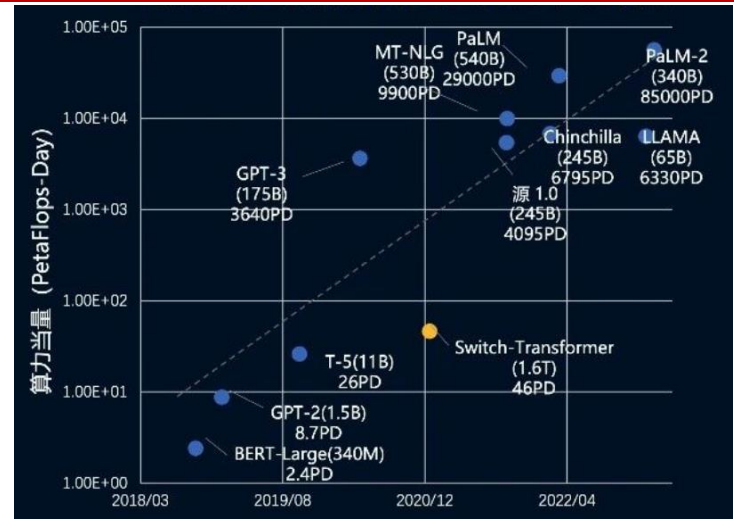
大模型参数量指数级增长，智能算力需求水涨船高。模型的参数量越大，训练数据量越大，模型泛化能力越强，但与此同时对算力的要求越高。根据浪潮信息 Owen ZHU 在 NPCon 大会的演讲，2019 年 2 月，GPT-2 的模型参数量为 15 亿，对应算力当量为 8.7PD (PetaFlops-Day)，2020 年 3 月，GPT-3 的模型参数提高至 1750 亿，对应算力当量为 3640PD，是上一代 GPT-2 的 418 倍，谷歌 2023 年 5 月推出的 PaLM-2 参数量为 3400 亿，算力当量已经达到 85000PD。

图 30 算力当量与交叉熵对应关系



资料来源：浪潮信息，首创证券

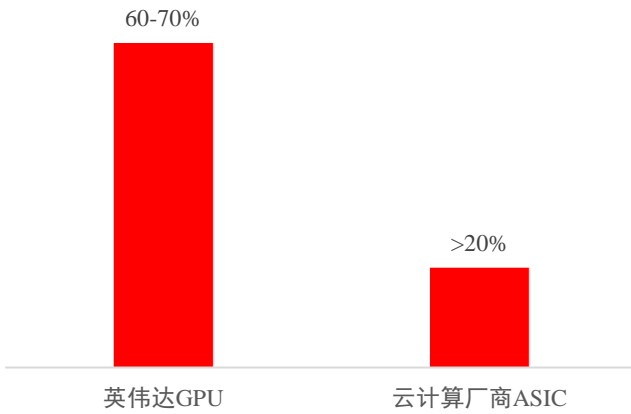
图 31 大模型发布时间及对应的算力当量



资料来源：浪潮信息，首创证券

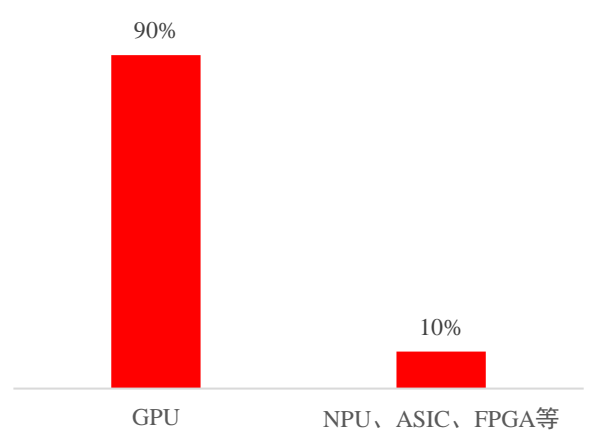
在众多适用密集型计算的芯片种类中，GPU 目前仍是最主流的选择。根据 TrendForce 2023 年 5 月发布的报告，全球 AI 服务器市场中，英伟达 GPU 市场份额约占 60-70%，其他厂商自主研发的 ASIC 芯片市场份额超过 20%。而根据 IDC，2021 年中国数据中心加速计算中 GPU 芯片市占率近 90%，NPU、ASIC、FPGA 等约 10%。

图 32 全球 AI 服务器芯片种类规模占比



资料来源：TrendForce，首创证券

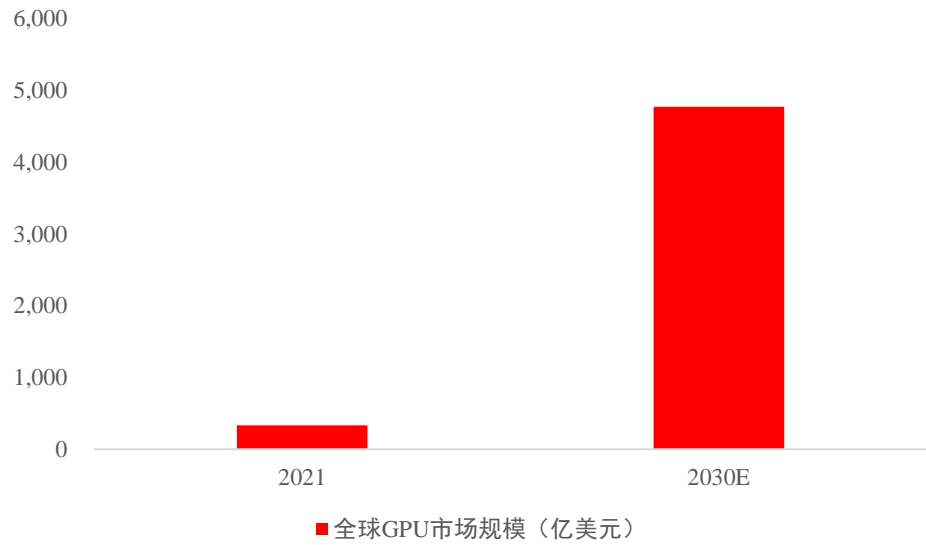
图 33 2021 年中国数据中心加速计算芯片种类



资料来源：IDC，首创证券

大规模算力需求下，GPU 规模高速扩张。根据中商情报网的预测，2030 年全球 GPU 市场规模将会达到 4773.7 亿美元，2021-2030 CAGR 34.4%。

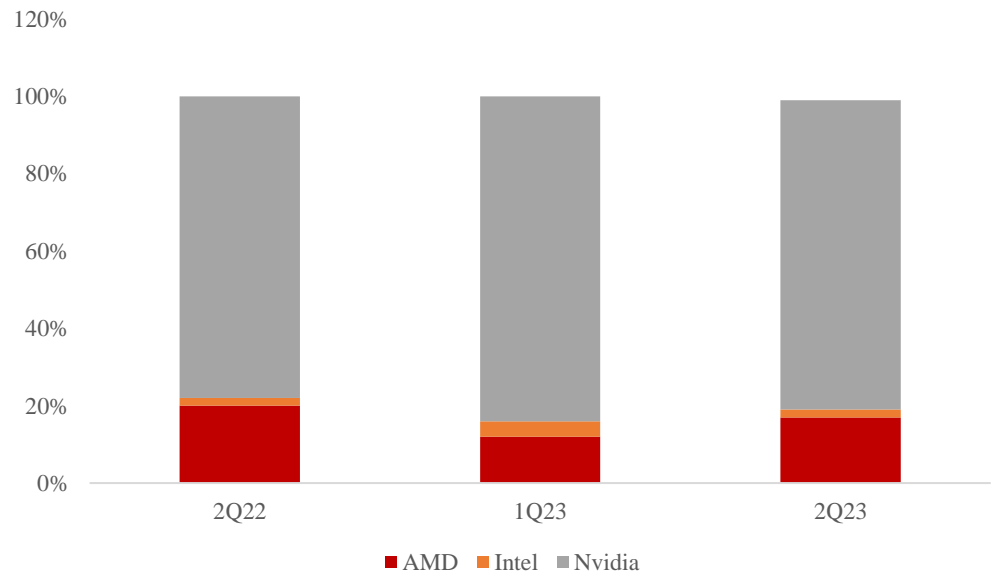
图 34 全球 GPU 市场规模预测



资料来源：中商情报网，首创证券

英伟达在 GPU 市场保持着绝对的龙头地位。除了不断推出的性能强大的 GPU，英伟达还凭借其 CUDA 生态不断拓宽自身护城河。根据 JPR 统计数据，2023 英伟达在独立显卡市场中份额为 80%，在 GPU 市场了保持着领先的市场份额。

图 35 AIB 显卡市场份额变化



资料来源：JPR，首创证券

英伟达 GPU 产品主要分为游戏、设计、数据中心三条线，Hopper 架构的 H200 是目前数据中心条线最强 GPU。英伟达 GPU 产品线主要分为面向游戏娱乐领域的 GeForce GTX/RTX 系列显卡、面向专业设计和虚拟化领域的 NVIDIA RTX/Quadro 系列显卡，以及面向数据中心的高算力 GPU。根据英伟达官网，数据中心产品线目前在售产品主要包括 Ampere 系列，Hopper 系列，Ada Lovelace 系列和 Turing 系列。其中，最新推出的 H200 是首款采用 HBM3e 的 GPU，显存配置较上一代显著提升。H200 显存 141GB，是 H100 的

1.8 倍，显存带宽 4.8TB/s，是 H100 的 2.4 倍，在处理 700 亿参数的大语言模型 Llama2 时，H200 的推理速度是 H100 的 1.9 倍。

表 2 英伟达高算力芯片：数据中心 GPU

系列	型号	性能	发布时间	供应状态	2023.12.7 京东售价
Ampere	A2	CUDA 核心数 1280，单精度 4.5T，16GB 显存，200GB/s 带宽	2021 年 11 月	正常	7999
	A10	CUDA 核心数 9261，单精度 31.2T，24GB 显存，600GB/s 带宽	2021 年 4 月	正常	29999
	A16	CUDA 核心数 2560，单精度 8.68T，64GB 显存，800GB/s 带宽	2021 年 4 月	正常	27899
	A30	CUDA 核心数 3584，单精度 10.3T，24GB 显存，933GB/s 带宽	2021 年 4 月	正常	45999
	A40	CUDA 核心数 10752，单精度 37.42T，48GB 显存，696GB/s 带宽	2020 年 10 月	正常	47999
	A100	CUDA 核心数 6912，单精度 19.5T，80GB 显存，1935GB/s 带宽	2020 年 5 月	禁止	144230
	A800	CUDA 核心数 6912，单精度 19.5T，80GB 显存，1935GB/s 带宽	2022 年 11 月	禁止	159899
Hopper	H100	CUDA 核心数 14592，单精度 51T，80GB 显存，2TB/s 带宽	2022 年 3 月	禁止	257229
	H200	CUDA 核心数 16896，单精度 67T，141GB 显存，4.8TB/s 带宽	2023 年 11 月	禁止	未开始供货
	H800	CUDA 核心数 14592，单精度 51.22T，80GB 显存，2039GB/s 带宽	2023 年 3 月	禁止	228229
Ada Lovelace	L4	CUDA 核心数 7680 个，单精度 30.3T，24GB 显存，300GB/s 带宽	2023 年 3 月	正常	19999
	L40	CUDA 核心数 18176 个，单精度 90.5T，48GB 显存，864GB/s 带宽	2022 年 10 月	禁止	89899
	L40S	CUDA 核心数 18176 个，单精度 91.6T，48GB 显存，864GB/s 带宽	2023 年 8 月	禁止	124899
Turing	T4	CUDA 核心数 2560 个，单精度 8.14T，16GB 显存，320GB/s 带宽	2018 年 9 月	正常	6699

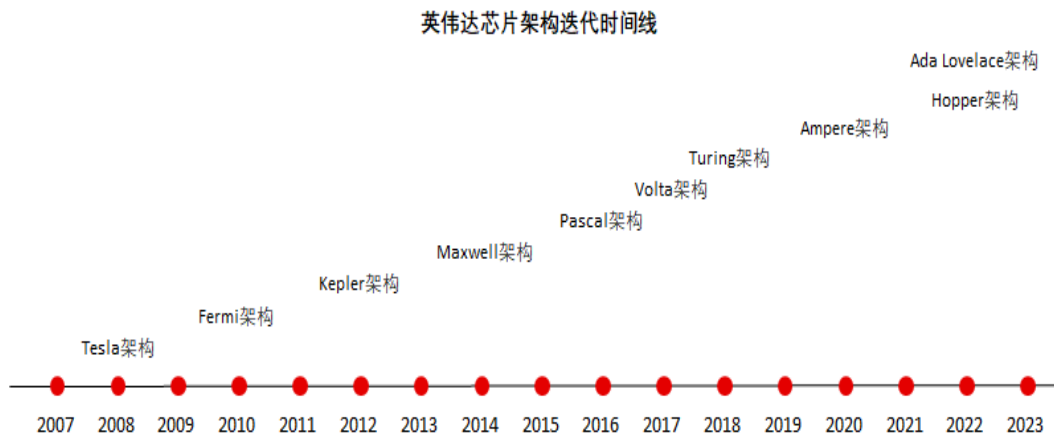
资料来源：英伟达，京东，TopCPU，芯参数，首创证券

注：命名规则：首字母为芯片微架构首字母，数字为芯片性能的编码，越大越强，如 A 为 Ampere 架构，H 为 Hopper 架构

**英伟达护城河之一：算力，持续架构迭代提升 GPU 性能。**2010 年，英伟达推出世界上第一个完整的 GPU 架构 Fermi，此后，英伟达不断通过扩展 Cuda 核心种类，增加 CUDA Core 数量，引入并升级 Tensor Core & RT core 等等途径，增强 GPU 在深度学习、AI 运算方面的性能。2022 年英伟达推出全新的 Ada Lovelace 架构和 Hopper 架构，其中专为数据中心打造的 Hopper 架构采用了台积电 4nm 制造工艺，与上一代相比，Hopper 将 TF32、FP64、FP16 和 INT8 精度的 FLOPS 提高了 3 倍。



图 36 英伟达 GPU 架构演进历史

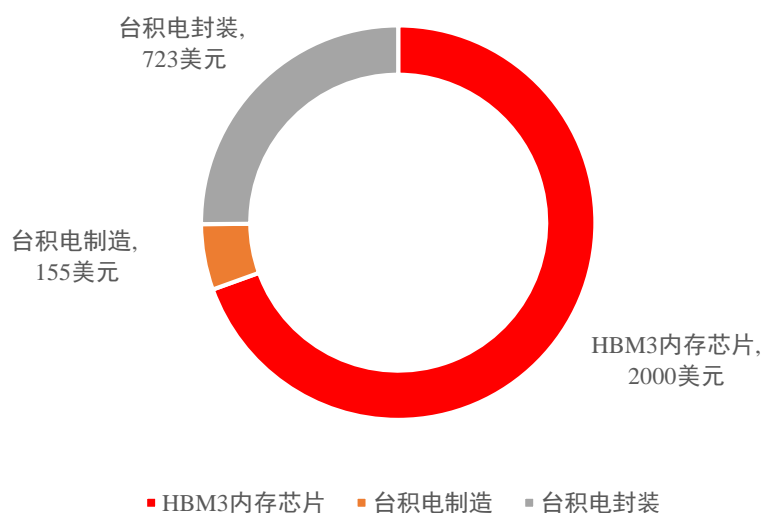


资料来源：CSDN，首创证券

**英伟达护城河之二：生态，CUDA 工具降低开发门槛。** CUDA 是英伟达 2006 年推出的并行计算框架，本质是一系列用于优化计算的编程函数，通过提供包括数据索引、内核函数、线程分配等在内的完整的工具套件，方便开发者针对不同任务对处理器进行编程，从而让 GPU 的功能由图形处理拓展至通用计算，具有了解决复杂计算问题的能力。但 CUDA 并非开源生态，英伟达拥有大量专利壁垒。随着不断迭代，CUDA 在针对 AI 或神经网络深度学习领域推出了非常多的加速库，构成了 CUDA 的软硬件生态站。完善的功能吸引更多开发者使用，大量的开发者亦不断完善 CUDA 生态，从而形成正向循环。截至 2023 年 5 月，CUDA 注册开发者增加至 400 万人。

**在算力和生态双重支撑下，英伟达 H100 毛利或高达 90%。** 英伟达作为芯片设计公司，主要制造成本为购买 BOM（物料清单）的支出和委托第三方的制造封装成本。根据晚点 LatePost 的测算，收入方面，一颗 H100 售价超 3 万美元，制造成本方面，BOM 支出主要为从 SK 海力士采购的 HBM3 内存芯片约 2000 美元，委托台积电制造封装预计花费 878 美元，粗略测算英伟达 H100 毛利或高达 90%。

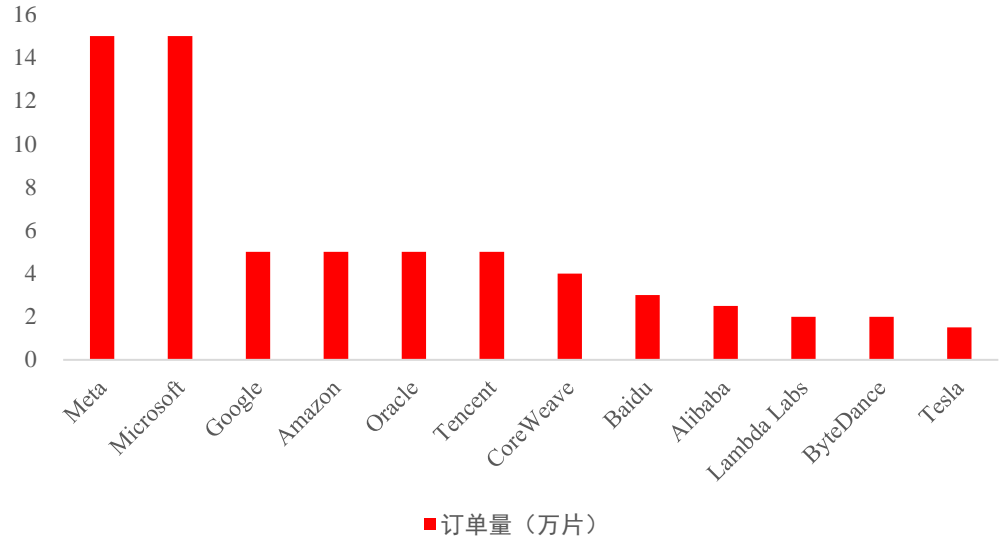
图 37 英伟达 H100 芯片成本主要构成



资料来源：晚点 LatePost，首创证券

英伟达高端芯片供不应求，货期最长可达 52 周。在庞大的需求与紧张的供货下，H100 最长需要 52 周才能交付。同时，大多数服务器 GPU 都供应给了超大规模云服务提供商，而服务器原始设备制造商（如戴尔、联想、HPE）目前还无法获得足够的 AI 和 HPC GPU。根据 Omdia 数据，Meta 与微软是 H100 最大的客户，预计 2023 年订购量为 15 万张，谷歌、亚马逊等预计订购 5 万张。

图 38 2023 年终端客户对于 H100 的预计订单量



资料来源: Omdia, 首创证券

## 5.2 英伟达下游客户持续投入 AI 芯片自研，全球 ASIC 市场百花齐放

一方面，英伟达直接竞争对手持续进行升级。2023 年 12 月 7 日，AMD 推出最新数据中心芯片 Instinct MI300X GPU，在算力和显存方面均实现了对 H100 的超越。算力方面，TF32(Sparsity)等指标是 H100 的 1.3 倍，FP64 (Vector)等指标是 H100 的 2.4 倍；显存方面，MI300X 有 192GB 的 HBM3 显存容量，容量是 H100 的 2.4 倍，显存带宽 5.3TB/s，是 H100 的 1.6 倍。

表 3 Instinct MI300X VS H100

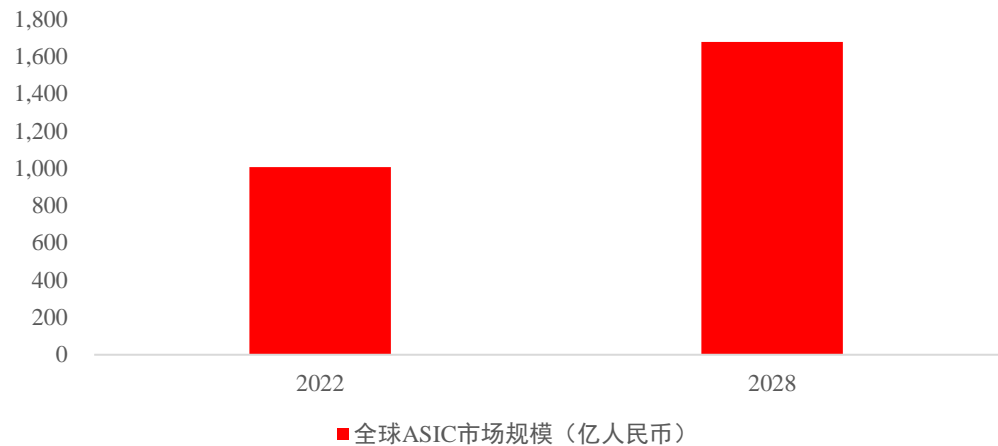
	指标	MI300X	H100	MI300X 优势
AI 性能 (Peak TFLOPs)	TF32(Sparsity)	1307.4	989.6	1.3x
	FP16/BF16(Sparsity)	2614.9	1978.9	1.3x
	FP8(Sparsity)	5229.8	3957.8	1.3x
HPC 性能 (Peak TFLOPs)	FP64 (Vector)	81.7	33.5	2.4x
	FP64 (Tensor / Matrix)	163.4	66.9	2.4x
	FP32 (Vector)	163.4	66.9	2.4x
显存	显存容量 (GB)	192	80	2.4x
	显存带宽 (TB/S)	5.3	3.4	1.6x

资料来源: AMD 官网, 首创证券

另一方面，众多英伟达客户投入 ASIC 芯片研发中，ASIC 市场高速发展。ASIC 不同于 CPU、GPU、FPGA，目前全球 ASIC 市场并未形成明显的头部厂

商。在庞大需求与有限供应的不对称影响下，一方面出于自身业务需求，另一方面寻求新的高利润增长点，许多云服务提供商以及新兴创业公司投入 ASIC 芯片研发中。根据贝哲斯咨询统计，2022 年全球 ASIC 市场规模达到了 1005.51 亿人民币，预计 2028 年市场规模将达到 1677.49 亿人民币，2022-2028 CAGR 9%。

图 39 2022-2028 全球 ASIC 市场规模



资料来源：贝哲斯咨询，首创证券

**微软：进军 AI 芯片领域，推出两款自研 AI 芯片。**2023 年 11 月，微软推出两款自研 AI 芯片 Azure Maia 100 GPU 和 Azure Cobalt 100 CPU。Maia100 芯片基于台积电的 5 纳米工艺打造，总共包含 1050 亿个晶体管，据 semianalysis，Maia 100 在 MXInt8 下的性能为 1600 TFLOPS。Cobalt 100 基于 Arm 架构，据 semianalysis，Cobalt 100 包含 128 个 Neoverse N2 内核和 12 个 DDR5 通道。

图 40 微软 Azure Maia 100 GPU



资料来源：微软官网，首创证券

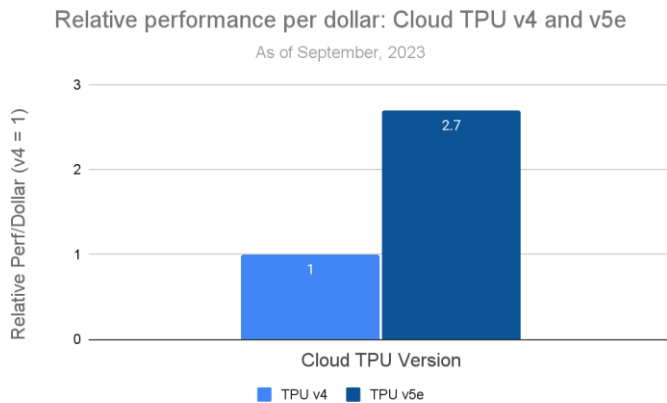
图 41 微软 Azure Cobalt 100 CPU



资料来源：微软官网，首创证券

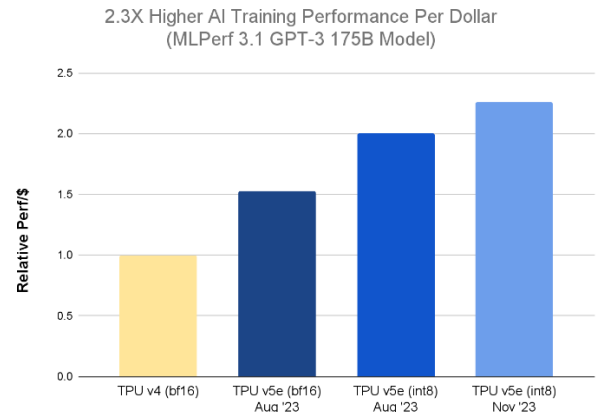
**谷歌：推出新一代高性价比 TPU 芯片 TPUv5e。**TPU 是谷歌专为机器学习和深度学习设计的专用 AI 加速芯片。2023 年 8 月，Google 推出新一代 AI 芯片 TPUv5e。根据谷歌 9 月 MLPerf™ 3.1 推理基准测试结果，TPUv5e 每美元推理性能是 TPU v4 的 2.7 倍；根据谷歌 11 月 MLPerf™ 3.1 训练基准测试结果，TPUv5e 每美元训练性能是 TPU v4 的 2.3 倍。

图 42 谷歌 9 月 MLPerf™ 3.1 推理基准测试结果



资料来源：谷歌官网，首创证券

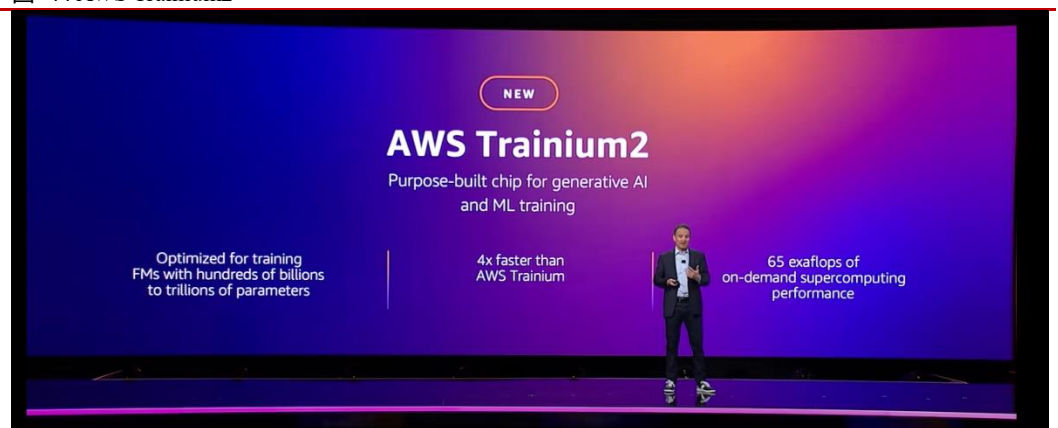
图 43 谷歌 11 月 MLPerf™ 3.1 训练基准测试结果



资料来源：谷歌官网，首创证券

**亚马逊：推出 Trainium2，训练速度较上一代提升四倍。**2023 年 11 月，亚马逊发布最新 AI 训练芯片 Trainium2。Trainium2 专为数万亿参数的基础模型（FM, foundation model）和大语言模型（LLM, large language models）的高性能训练设计，相较第一代芯片 Trainium，Trainium2 的训练速度快 4 倍，存储容量大 3 倍，能源效率提升 2 倍。Amazon EC2 Trn2 实例将采用 Trainium2 芯片，从而支持客户在下一代 EC2 UltraCluster 中大规模扩展至多 10 万张 Trainium2 芯片，通过 AWS EFA 的 PB 级网络连接，能够提供高达 65 exaflops ( $10^{18}$ flops) 的算力。

图 44 AWS Trainium2



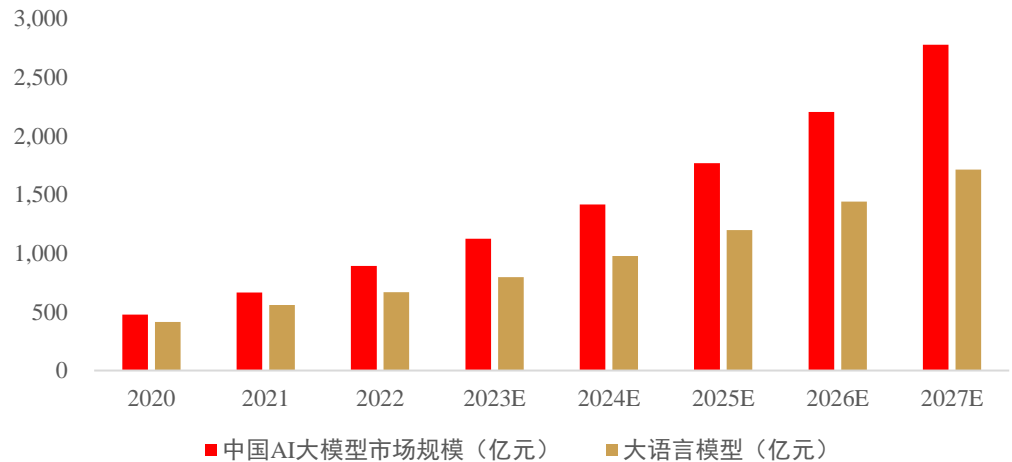
资料来源：2023 re:Invent，首创证券

### 5.3 禁令下国产替代重要性凸显，国内厂商技术持续突破

**GPT3.5 横空出世带领 LLM 实现革命性突破，我国 AI 大模型行业顺应趋势快速扩张。**2022 年 6 月 GPT-3.5 的发布，为语言大模型（LLM）行业带来革命性影响，2022 年中国 LLM 行业规模达到 668 亿元，根据头豹研究所和弗若斯特沙利文预测，2027 年市场规模将会达到 1713 亿元，2022-2027 CAGR 20.7%。除了 LLM，AI 大模型还包括视觉大模型和多模态大模型，2022 年中国 AI 大模型规模达到 891 亿元，根据头豹研究所和弗若斯特沙利文预测，2027 年市场规

模将增至 2778 亿元，2022-2027 CAGR 25.5%。

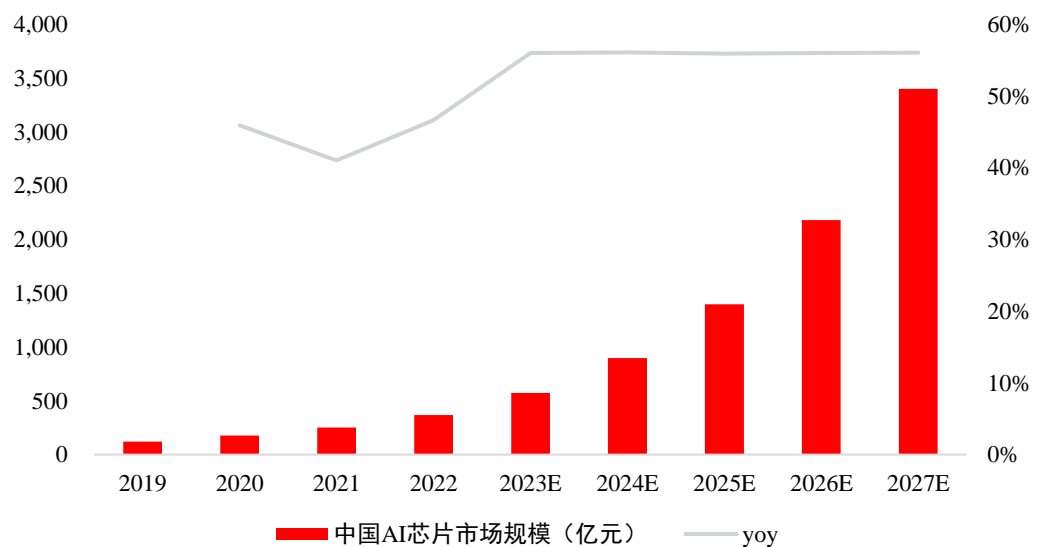
图 45 中国 AI 大模型市场规模



资料来源：头豹研究所，沙利文，首创证券

大模型装备竞赛下国内对高端算力芯片需求旺盛，高端算力芯片国产替代正当时。大模型能力的提升需要算力硬件进行支撑，算力充足与否，直接决定了 AI 产品供应商能否长期稳定的提供服务，从而保持行业竞争力。根据头豹研究所和沙利文联合发表的白皮书，2022 年中国 AI 芯片市场规模为 368 亿元，预计 2027 年将达到 3400 亿元，2022-2027 CAGR 56%。

图 46 中国 AI 芯片市场规模



资料来源：头豹研究所，沙利文，首创证券

2023 年 10 月 BIS 加强对华高端半导体管控，英伟达多款高端算力芯片限制对华出售。2023 年 10 月 17 日，美国商务部工业与安全局 (BIS) 发表政策文件，针对高端半导体出口政策进行限制。芯片方面，BIS 新引入总算力性能 TPP 和性能密度 TP 参数来明确芯片性能，对于符合 ECCN 3A090.a 和 3A090.b 规定的算力芯片进行出口限制。根据英伟达在 17 日提交给 SEC 的文件，其多款高性能计算芯片受到出口限制，包含 A100、



A800、H200、H100、H800、L40、L40S 等。

表 4 BIS 限制的芯片性能范围

算力芯片性能限制

<p>3A090.a:</p> <p>(1) 总计算性能 <math>\geq 4800</math>;</p> <p>(2) 总计算性能 <math>\geq 1600</math>, 且性能密度 <math>\geq 5.92</math>。</p>
<p>3A090.b:</p> <p>(1) <math>4800 &gt;</math> 总计算性能 <math>\geq 2400</math>, 且 <math>5.92 \geq</math> 性能密度 <math>&gt; 1.6</math>;</p> <p>(2) 总计算性能 <math>\geq 1600</math>, 且 <math>5.92 \geq</math> 性能密度 <math>&gt; 3.2</math>。</p>
<p>注释:</p> <p>总算力性能 TPP (Total Processing Performance) = <math>2 * \text{MacTOPS} * \text{操作比特长度}</math></p> <p>1) MacTOPS 衡量算力, 指的是每秒进行相乘累加计算的理论峰值次数;</p> <p>2) 操作比特长度衡量位宽, 指的是进行相乘累加计算输入值的最大比特长度。</p> <p>性能密度 PD (Performance Density) = TPP/对应芯片面积。</p>

资料来源: BIS, 首创证券

**龙芯中科: 龙链技术加速片间互联, 9A1000 显卡明年流片。**龙芯中科推出最新的龙链技术, 对标英伟达 NVLink、CXL, 可以实现片间高速互联, 是破解 Chiplet 的关键技术。龙芯中科表示, 龙链技术与公司目前使用的片间一致性互联接口相比, 能够有效降低互联延迟, 成倍提升实测带宽。目前公司正在进行国产显卡的研发, 9A1000 计划于 2024 年 Q3 流片, 对标 AMD RX550, INT8 大概 32TOPS。此外, 应用龙链技术可以实现 9A1000 多片互联, 形成更高性能。

图 47 龙芯中科高速片间互联技术——龙链技术



资料来源: 龙芯中科, 首创证券

**海光信息: 深算二号 DCU 性能翻番, 三号研发顺利。**深算系列为海光 DCU 产品, 以 GPGPU 架构为基础, 可用于大数据处理、人工智能、商业计算等应用领域。公司在招股书中表示, 在典型应用场景下, 深算一号指标达到国际上同类型高端产品的水平。目前, 公司已经推出深算二号, 性能相较于深算一号提升 100% 以上, 深算三号正处于研发中。



表 4 海光深算一号性能对比

公司	海光	NVIDIA	AMD
品牌	深算一号	Ampere 100	MI100
生产工艺	7nm FinFET	7nm FinFET	7nm FinFET
核心数量	4096 (64 CUs)	2560 CUDA processors 640 Tensor processors	120CUs
显存容量	32GB HBM2	80GB HBM2e	32GB HBM2
显存位宽	4096 bit	5120 bit	4096bit
显存频率	2.0 GHz	3.2 GHz	2.4 GHz
显存带宽	1024 GB/s	2039 GB/s	1228 GB/s
TDP	350 W	400 W	300W
CPU to GPU 互联	PCIe Gen4 x 16	PCIe Gen4 x 16	PCIe GEN4 x 16
GPU to GPU 互联	xGMI x 2, Up to 184 GB/s	NVLink, up to 600 GB/s	Infinity Fabric x 3, up to 276 GB/s

资料来源：海光信息招股书，首创证券

**寒武纪：思元 370，国内首款支持 LPDDR5 的云端 AI 芯片。**寒武纪云端智能芯片产品线包含思元 100，思元 270，思元 290，思元 370。其中，思元 370 基于 7nm 制程，是寒武纪首款采用 chiplet 技术的 AI 芯片，INT8 下的峰值算力达到 256TOPS，是思元 270 算力的 2 倍。思元 370 在业内率先支持 LPDDR5 内存，内存带宽是上一代产品的 3 倍。

图 48 思元 370



资料来源：寒武纪官网，首创证券

除了硬件升级，国内持续在前沿技术方面突破。在芯片制造材料方面，2023 年 11 月华为公布了和哈工大联合申请的全新芯片专利，是“一种基于硅和金刚石的三维集成芯片的混合键合方法”，突破了现有三维集成以硅为衬底的基础，利用金刚石作为连接材料，实现硅基半导体与金刚石之间的无缝融合，从而提高芯片集成度，降低制造成本，提升产品性能。在芯片制造设备方面，2022 年 12 月，复旦大学研究团队发明的晶圆级硅基二维互补叠层晶体管在国际顶尖期刊《自然-电子学》上发表，该技术可以在不依赖 EUV 光刻工艺下，在相同工艺节点下实现器件集成密度的翻倍，从而显著提高性能。

国内算力芯片厂商前景广阔。需求方面，生成式 AI 逐步渗透，同步带动相关算力需求集中释放，高端芯片蓝海市场；供给方面，BIS 禁令切断我国人工智能产业从国外进口高端芯片的途径，算力缺口提供了国产化芯片即时补位的窗口机遇，国内芯片厂商加速升级满足市场需求。在广阔市场需求和不断技术突破背景下，我看好国产算力芯片

板块发展。

## 6 数字芯片：AI 驱动硬件创新，边缘算力重点布局

### 6.1 AI 算力正逐渐向边缘端拓展

ChatGPT 引爆生成式 AI 应用热潮，人工智能应用有望迎来高速增长。2023 年上半年迎来了以 GPT 为代表的人工智能大模型热潮。人工智能大模型技术的创新突破为社会、经济、产业发展等带来变革，更多应用场景的落地将进一步拓展产业生态。在大模型技术不断成熟的基础上，大模型的小型化、专业化也会加速推进，生成式 AI 的潜在应用场景不断扩大，将逐渐渗透到教育、金融、医疗、文娱、智能家居、消费电子等行业，将引发未来人工智能应用的高速增长。

伴随生成式 AI 的飞速发展，算力从云端逐渐向边缘端拓展将是大势所趋。生成式 AI 的快速发展带来了大量的算力需求，将 AI 处理逐渐从云端向边缘端扩展能够保证 AI 实现规模化扩展并实现最大潜能。与仅在云端进行处理不同，混合 AI 架构在云端和边缘终端之间分配并协调 AI 工作负载。云端和边缘终端如智能手机、汽车、个人电脑和物联网终端协同工作，能够实现更强大、更高效且高度优化的 AI。

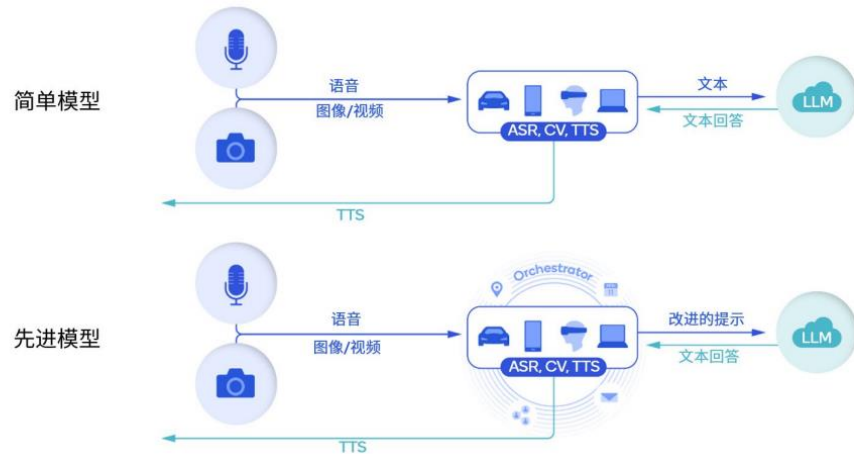
图 49 AI 处理的重心正向边缘转移



资料来源：高通官网，首创证券

云端和终端协同工作，可形成成本、能耗、性能、隐私、安全和个性化等多方面优势。将 AI 处理从云端转移到边缘端可以减轻云基础设施的压力并减少开支，同时边缘终端能够以更低的能耗运行生成式 AI 模型，尤其是将处理和数据传输相结合时。并且端侧 AI 还可以在无网络连接时正常运行生成式 AI 应用。由于查询和个人信息保留在终端上，端侧 AI 还可以从本质上保护用户的隐私，且提供个性化的体验。

图 50 在以终端为中心的混合 AI 架构中云端仅用于分流处理终端无法充分运行的 AI 任务

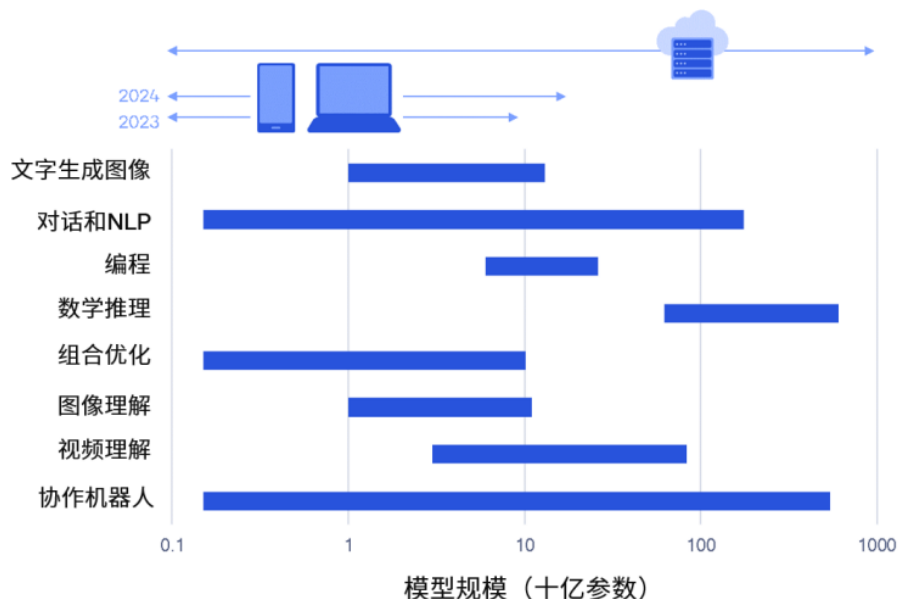


资料来源：高通官网，首创证券

将自动语音识别等模型工作负载转移至端侧可以节省计算和连接带宽。在混合 AI 架构中，当用户对着智能手机说话时，自动语音识别模型可以将语音转换为文字，然后将其作为请求提示发送至云端，并在云端运行大语言模型，随后将生成的文本回复至终端，终端侧运行文本生成语音模型，给予用户语音回复。随着大语言模型变为多模态并支持图像输入，计算机视觉处理也可以在终端上运行，以进一步分流计算任务并减少连接带宽，从而节省成本。

伴随生 AI 模型逐渐变小且端侧处理能力的提升，更多的 AI 模型可以在端侧运行。端侧嵌入式 AI 算力载体从 CPU、GPU、DSP 发展到 ASIC 架构，推动了基于深度学习的语音识别、人脸识别、图文识别、AIGC、目标检测、超分辨率、ADAS 等技术的广泛应用。目前参数在 10 亿至 100 亿之间的部分模型已经可以在端侧顺利运行，如 Stable Diffusion 等参数超过 10 亿的模型已经能够在手机上运行，且性能和精确度达到与云端处理类似的水平。

图 51 数量可观的生成式 AI 模型可从云端分流到终端上运行



资料来源：高通官网，首创证券

## 6.2 生成式 AI 有望促进 XR 加速创新

下一代 AI 渲染工具将赋能内容创作者使用如文本、语音、图像或视频等各种类型的提示，生成 3D 物体和场景，并最终创造出完整的虚拟世界。此外，内容创作者将能够利用文本生成文本的大语言模型，为能够发出声音并表达情绪的虚拟化身生成类人对话。

图 52 生成式 AI 模型将面向 XR 赋能对话式 AI 和全新渲染工具

模态	对话式 AI		AI 渲染工具		
	文本生成文本	文本生成图像	文本生成 3D	图像生成 3D	视频生成 3D
模型示例	ChatGPT	Stable Diffusion	Magic3D	Instant NeRF	Unsolvd
描述	利用大语言模型(LLM)生成类人回复	利用 2D 扩散模型将文本转化为逼真的图像	利用扩散 + NeRF (或类似技术) 将文本转化为 3D 模型	利用 NeRF 将图像转化为逼真的 3D 模型	将视频转化为逼真的 3D 模型
执行	语音 ↓ ASR* ↓ 文本 ↓ ChatGPT ↓ 文本 ↓ TTS** ↓ 语音	语音 ↓ ASR ↓ 文本 ↓ Stable Diffusion ↓ 图像 ↓ 游戏引擎* ↓ 3D 纹理	语音 ↓ ASR ↓ 文本 ↓ Magic3D ↓ 3D ↓ 游戏引擎 ↓ 3D 物体	图像 (单/多张) ↓ NeRF ↓ 3D ↓ 游戏引擎 ↓ 3D 物体 3D 场景 3D 虚拟化身	视频 ↓ 生成式 AI ↓ 3D ↓ 游戏引擎 ↓ 3D 场景 3D 世界
在 XR 中的应用	为能够发音并表达情绪的虚拟化身生成类人对话	为 3D 物体/虚拟化身生成新纹理或颜色	生成逼真的 3D 物体以推动虚拟世界普及	利用手机摄像头生成 3D 场景或用户的 3D 虚拟化身	生成 3D 场景并最终生成整个 3D 虚拟世界

\* ASR = 自动语音识别 \*\* TTS = 文本生成语音 \*游戏引擎 = 将生成式 AI 模型引入图形渲染管线

资料来源：高通官网，首创证券

## 6.4 国产数字 SoC 厂商积极布局端侧算力

### 6.4.1 瑞芯微：紧跟 AI 快速发展趋势，提供多种算力平台。

边缘计算领域已有多年积累，头部客户合作关系良好。公司边缘计算是公司的重要的产品线之一，公司在边缘计算领域已有多年积累。从 RK3288、RK3399、RK3568 到 RK3588 等，公司多款产品与众多头部企业已有合作。公司十分看好边缘计算市场的发展前景，积极主动拥抱 AI 发展大势。2023 年，公司持续拓展更多边缘计算领域的客户，已取得一定成绩。

公司可提供 0.5T-6T 的多算力芯片，满足下游客户边缘计算需求。公司的芯片在靠近物或数据源头的一侧就近提供比云计算更快、更稳定的服务响应，满足各个行业在实时性、可靠性、应用智能、安全与隐私保护等方面的关键需求。公司提供从 0.5T 到 6T 的不同算力芯片，在不同芯片之间可实现算法快速移植、更新迭代，从而帮助客户缩短研发周期、节省研发投入，快速推出不同定位、不同算力的产品。

既有平台开发 AI 技术满足客户需求，持续升级人工智能 NPU IP 和相关工具链。在既有的芯片平台上，公司从算法和软件层面出发，持续研发各种应用场景需要的 AI 技术，更好地匹配客户产品的需求。在芯片设计上，公司持续升级人工智能 NPUIP 和相关工具链，更好地支持基于 Transformer 底层技术的各种大模型，同时针对不同定位的平台，设计不同的算力性能以满足客户的差异化需求。

### 6.4.2 全志科技：各下游应用领域积极拥抱 AI。

随着 AI 算法在市场的持续渗透，各类场景对 AI 专用算力的需求日益增长，公司深入分析 AI 算法的应用场景，主动积极推进 NPU 和 DSP 的 AI 专用算力在终端产品应用



的落地,同时在大算力&大数据总线架构、多核异构、AI 专用处理器、低功耗系统架构、先进工艺平台等方向持续研究和突破。

**智能音箱: 搭载公司 R329 芯片的天猫精灵已训练出“鸟鸟分鸟”。**公司早在 2020 年就已推出主打 AI 语音专用的 R329 芯片,这是全志科技首款搭载 Arm 中国全新 AI 处理单元(AIPU)的高算力、低功耗 AI 语音专用芯片。通过集成高性能的 AIPU、DSP、CPU,将为智能音箱、智能家居带来崭新的 AI 交互体验,2023 年基于公司 R329 芯片的天猫精灵 SOUND 系列已接入标杆客户大模型终端操作系统。

图 53 阿里利用 15 天时间训练出“鸟鸟分鸟”



资料来源: 智东西公众号, 首创证券

图 54 全志科技 AI 语音专用芯片 R329



资料来源: 全志科技公众号, 首创证券

**智能机器人: 面向中高端推出新一代 MR 系列高性能八核+AI 专用算力芯片。**针对扫地机器人避障不够智能的用户痛点,下游客户推出了集成 AI 物体识别及避障功能的中高端产品,公司面对这一产品趋势,推出了新一代 MR 系列高性能八核+AI 专用算力芯片及解决方案,满足高端扫地机产品对新功能和体验的需求。此外,伴随 AI 的快速发展,公司已提前布局商用服务机器人和割草机等市场,实现机器人领域的产业全覆盖。

#### 6.4.3 炬芯科技: 从智能音频入局, 打造低功耗端侧 AI 算力。

AI 模型在音频领域有许多应用场景,部分场景下对能耗比提出较高要求。AI 模型在音频领域的应用场景包括语音识别、噪音抑制、语音翻译、语音/音频关键词识别、人声增强、人声分离等,具有广阔的市场前景。同时满足更高的算力与低功耗将成为端侧 AI 处理器芯片的重要需求,低功耗装置对于端侧人工智能算力提出新的挑战和需求。

**发挥公司传统技术优势,专攻低功耗端侧 AI 算力。**炬芯科技是国内领先的低功耗 AIoT 芯片设计厂商,凭借在低功耗音频等领域的多年积累,面对 AI 热潮,公司从智能音频入局,率先发力。公司将基于多核异构 AI 计算架构,在产品中逐步整合 AI 加速引擎,打造低功耗端侧 AI 算力,致力于提供高能耗比、高集成度、高性能和高安全性的端侧 AIoT 芯片产品。

公司最新一代三核异构高端 AI 音频芯片预计将于 2024 年开始 Sample。公司将推出最新一代升级为 CPU+DSP+NPU 三核异构的高端 AI 音频芯片 ATS286X,预计在 2024 年开始 Sample。其中 NPU 将采用 CIM (Compute-in Memory) 架构达成高能效比,为便携式产品提供低功耗下的大算力,将为 AI 降噪、人声分离、人声隔离等应用带来高品质的提升,未来将广泛应用于智能音频、智能办公、智能教育、智能陪护等多个市场领域。

图 55 炬芯科技最新一代高端 AI 音频芯片将于 2024 年 Sample



资料来源：炬芯科技公众号，首创证券

#### 6.4.4 乐鑫科技：低成本领域布局 AI SoC，高端方案第三方合作。

**高端方案领域：**使用自身 Wi-Fi 产品进行数据传输，搭配第三方复杂 AI 算法应用。生成式 AI 的蓬勃发展下，未来 AI 的应用领域将逐渐多元，面对高算力复杂应用领域，公司选择与第三方积极合作，各取所长，互利共赢。目前公司现有产品线 ESP32-S3 已可对接 OpenAI 的 ChatGPT 或百度的“文心一言”等云端 AI 应用。在高端方案领域，公司提供连接方案搭配第三方复杂 AI 算法共同实现 AI 应用。

**低成本方案：**研发 AI MCU 与无线连接功能集成的 SoC。面向低成本的下游终端应用领域，公司在自有的 SoC 中加入 AI 功能使得下游客户可以直接应用。如公司已在 ESP32-S3 中集成了加速神经网络计算和信号处理等工作的向量指令。AI 开发者们将可以使用指令优化后的软件库，实现本地的图像识别、语音唤醒和识别等应用。目前公司已同步研发基于 ESP32-S3 芯片的离线语音唤醒/识别的技术，可实现多达 200 条离线命令词，可被广泛应用于智能家居设备，目前已有客户开始定制唤醒词和命令词。

图 56 乐鑫科技 ESP32-S3-BOX 功能示意图



资料来源：乐鑫信息科技公众号，首创证券

## 7 模拟芯片：消费电子新机发布拉动需求，关注新品新领域扩展

### 7.1 平台厂商发布 AI 主控芯片，有望拉动换机需求

高通推出首款专为生成式 AI 打造的移动平台骁龙 8 Gen 3。2023 年 10 月 24 日，

请务必仔细阅读本报告最后部分的重要法律声明



高通发布第三代骁龙 8 移动平台，是高通首个支持多模态生成式 AI 模型的平台，该平台能够在终端侧运行高达 100 亿参数的 AI 模型，面向 70 亿参数的大语言模型每秒生成高达 20 个 token。这意味着，各类虚拟助手、GPT 聊天机器人、Stable Diffusion 等 LLM 等都可以在手机等终端正常运行。

图 57 高通骁龙 8Gen3 芯片概况

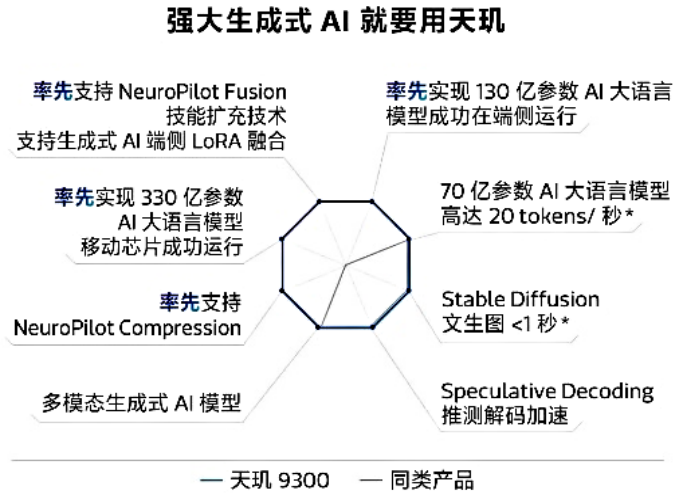


资料来源：高通中国公众号，首创证券

小米 14 首发搭载高通骁龙 8 Gen 3，开售前五分钟首销销量是上代手机的 6 倍。目前小米自研的 60 亿参数大模型已经在第三代骁龙 8 上流畅运行，小米 14 Pro 也通过将本地 AI 大模型植入到小米澎湃 OS 操作系统中，支持 AI 创作、AI 扩图、AI 魔法消除 Pro、AI 写真和 AI 搜图等一系列功能，为用户带来全新生成式 AI 体验。开售仅 5 分钟，小米 14 系列首销销量已猛增至上代首销总量的 6 倍。

联发科发布天玑 9300，率先搭载基于硬件的生成式 AI 引擎。2023 年 11 月 6 日，联发科正式发布天玑 9300，基于 MediaTek 第 7 代 APU 架构内建硬件级的生成式 AI 引擎，能够实现更快速且安全的边缘 AI 计算，率先实现 330 亿参数 AI 大预言模型移动芯片成功运行，70 亿参数 AI 大语言模型高达 20 tokens/秒，Stable Diffusion 文生图速度少于 1 秒。

图 58 联发科天玑 9300 芯片概况



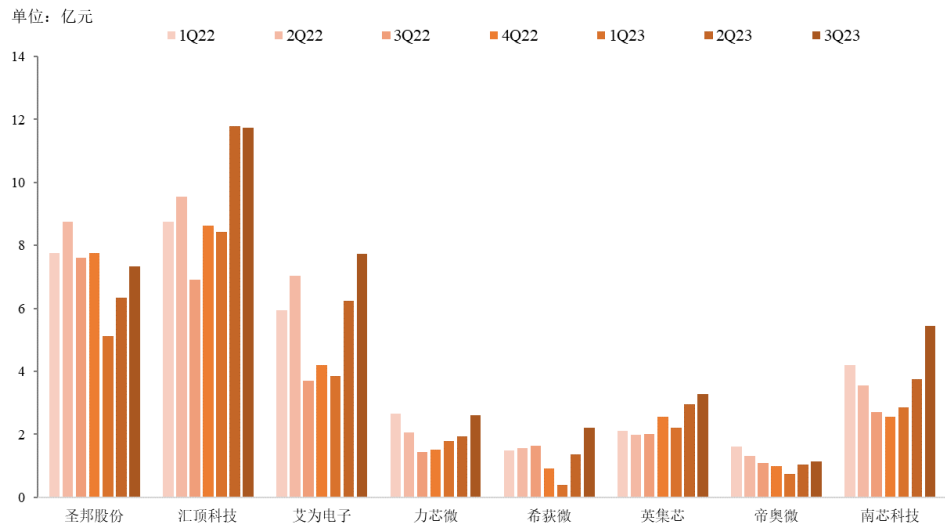
资料来源：联发科技公众号，首创证券

## 7.2 终端库存去化尾声，消费电子需求率先复苏

伴随消费电子终端客户库存去化逐步进入尾声，消费类占比高的模拟公司率先复苏。尽管手机等消费电子终端设备出货量在 2023Q3 同比仍在下滑，但经过几个季度的消化，各终端厂商库存水位也逐渐接近正常，受三四季度消费电子旺季需求的拉动，终端厂商于 Q2 开始逐步加大备货。

下游对应手机等消费电子占比高的模拟公司营收在 Q2 开始环比呈现大幅增长，其中，希荻微、艾为电子、汇顶科技、帝奥微、英集芯和南芯科技 2023Q2 收入环比增速分别高达 240.45%、62.37%、39.58%、38.95%、33.77%和 31.20%。

图 59 2022Q1-2023Q3 下游消费类收入较多的模拟芯片公司营业收入



资料来源：Wind，首创证券

## 7.3 关注国产模拟公司新产品和新领域扩展

模拟芯片是一种处理连续性模拟信号的集成电路芯片，广泛应用于消费电子、通讯、

工控、医疗、汽车等各下游终端。全球龙头 TI 的产品种类多达 8 万种以上。国内模拟芯片公司持续向新料号和新的下游应用领域持续拓展。

#### **圣邦股份：车规模拟深入布局，车规料号已超 100 颗。**

公司在快速增长的汽车芯片市场领域积极布局，不断扩大产品品类，提升产品性能品质，增强国际竞争力。自主研发多品类的符合汽车应用标准的车规级电源管理及信号链类模拟芯片，包括各类驱动芯片、隔离芯片、DC/DC 电源转换芯片、LDO、运放及比较器、电压基准芯片、小逻辑芯片等。在现已披露的 30 个产品大类中，每个品类下都有车规级的产品在研发、验证阶段或已进入销售阶段，总计 100 余颗。

#### **帝奥微：重点发力汽车电子，围绕马达驱动和车灯完善产品系列。**

公司推出了国内首款用于汽车车身控制的 15A 的 H 桥直流电机驱动产品，围绕马达驱动，还将推出步进马达、域区直流马达等用于汽车热管理，头灯，空调风门、水泵，油泵，风扇等领域的产品。车灯方面，公司已经布局了包括 12 路/16 路/24 路的贯穿式尾灯；多拓扑架构头灯驱动、BUCK-BOOST 头灯驱动和矩阵管理器（ADB）等头灯产品；以及具有控制单元的氛围灯等产品。

#### **必易微：DC/DC 产品稳步放量，支持级联式 BMS AFE 芯片加速研发。**

公司于 2022 年推出了 4.5-40V 电压段、0.6-6A 电流范围 DC-DC 芯片全系列产品并实现量产，广泛应用于电工照明、家用电器、网络通讯、安防监控等领域，目前正处于稳步快速放量阶段。BMSAFE 产品方面，针对 200-800V 高电压段应用，公司积极研发支持菊花链级联式 BMS AFE 芯片，同时推进 ISO26262 功能安全标准的认证，目标领域为大型储能系统、新能源汽车。

#### **南芯科技：手机充电端新产品快速导入，围绕汽车电子重点布局。**

2023 年公司推出了高集成度的 AC/DC 产品，其集成了原边、副边、协议、隔离等技术，全整合方案有望加速公司产品在客户端导入。汽车电子方面，公司 USB 及无线充电方案在客户端实现大规模量产；高性能 DC-DC 电源芯片、高边开关等新产品在客户端实现规模送样，部分客户已经进入项目定点设计阶段。

#### **晶丰明源：多款多相控制器已完成客户送样，1 颗 DrMOS 已经通过客户测试。**

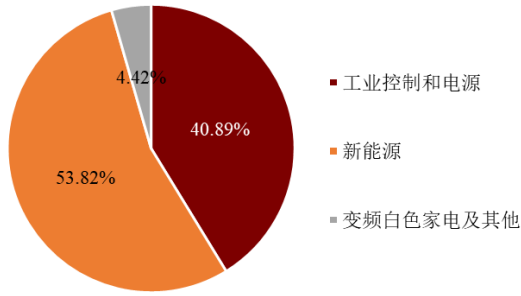
公司从大电流 DC/DC 多相方案入手，研发主要为 CPU 和 GPU 供电的多相控制器和 DrMos。2023H1 公司 DC/DC 产品线共 27 款产品进入研发阶段、14 款产品完成研发；公司推出的 16 相多相控制器产品已完成客户送样，应用于个人电脑和服务器的 40A、50A、70A 及 90A 等多款 DrMOS 产品均已完成内部评估，进入客户送样阶段，1 颗 DrMOS 已经通过客户测试。

## **8 功率半导体：国产替代持续推进，关注 SiC 渗透率拐点**

### **8.1 国产替代持续推进，新能源占比大幅提升**

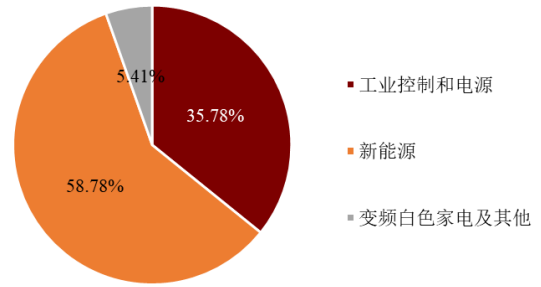
经过前期的验证导入，2023 年国内 IGBT 公司收入的下游结构从以工控为主转向以新能源领域为主要收入来源。其中，斯达半导 2023H1 在新能源领域实现收入 9.92 亿元，占收入比重从 2022 年的 53.82% 提升至 58.78%；宏微科技 2023 年 1-9 月新能源发电领域收入占比约 40%，电动汽车领域收入占比约 20%。

图 60 斯达半导 2022 年各下游领域收入占比



资料来源：斯达半导 2022 年年报，首创证券

图 61 斯达半导 2023H1 各下游领域收入占比



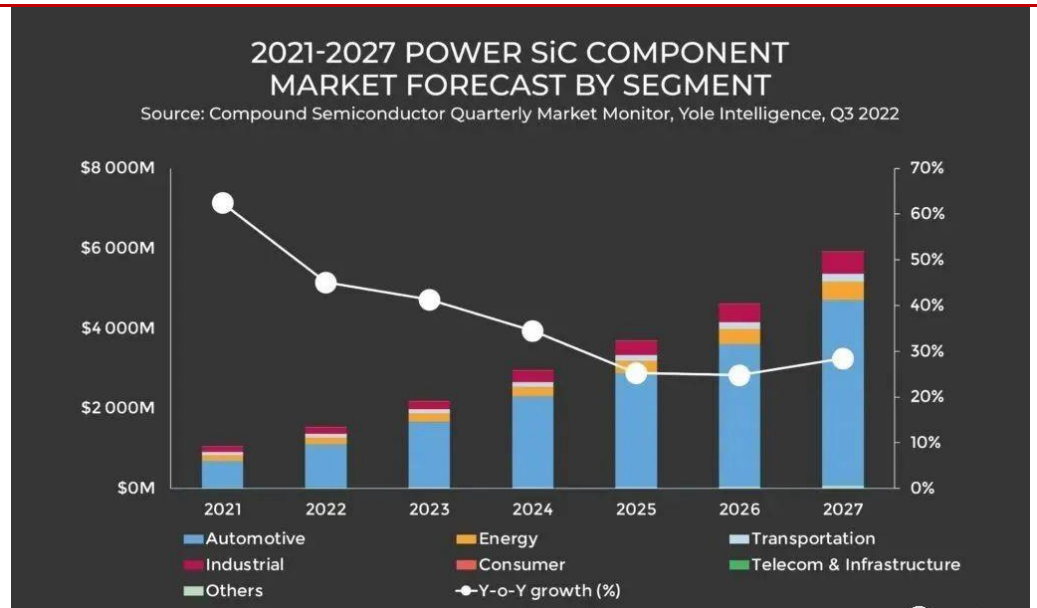
资料来源：斯达半导 2023 年半年报，首创证券

光伏 IGBT 模块大幅放量，国产替代向深水区迈进。2023H1 斯达半导基于第七代第七代微沟槽 Trench Field Stop 适用于光储行业的 IGBT 芯片以及模块开始批量供货，预计下半年市场份额将迅速提高。2023H1 宏微科技光伏用 400A/650V 三电平定制模块开发顺利，已开始批量交付，目前产能稳定，终端表现良好。扬杰科技针对光伏领域，成功研制了 650V/400A 三电平 IGBT 模块并投放市场，同时着手开发下一代 950V/600A 三电平 IGBT 模块。

## 8.2 碳化硅深入布局，关注渗透率拐点

根据 IHS 数据，受新能源汽车行业庞大的需求驱动，以及光伏风电和充电桩等领域对于效率和功耗要求提升的影响，预计到 2027 年碳化硅功率器件的市场规模将超过 100 亿美元，2018-2027 年的复合增速接近 40%。由于碳化硅材料具有宽禁带、高饱和漂移速度、高临界击穿电场等优异的性能，能够提升光伏风电和充电桩等领域对于效率和功耗要求，国内功率半导体公司积极布局。

图 62 2021 年-2027E 碳化硅各下游市场规模



资料来源：Yole，首创证券



**斯达半导：自建产线布局 SiC 芯片，车规 SiC MOSFET 模块持续放量。**碳化硅领域，斯达半导除拥有 SiC MOSFET 模块封装产线外，公司还通过募投项目自建 SiC 芯片产线，IDM 模式抢先布局 SiC 市场。项目建成后，公司车规级 6 吋 SiC MOSFET 芯片产能将达 6 万片/年。2023H1 公司应用于新能源汽车主电机控制器的车规级 SiC MOSFET 模块持续放量，在 800V 系统的主电机控制器领域，公司新增多个车规级 SiC MOSFET 模块项目定点，预计将拉动公司未来 SiC 领域收入。

**宏微科技：SiC 混合模块已批量出货，SiC MOS 预计 2024Q1 交付客户。**2022 年公司几款 SiC 混合模块已批量出货，2022 年底资产 SiC SBD 单管样品在客户端测试，目前公司 SiC 二极管已在客户端开展验证工作，SiC MOS 实现样品产出，预计于 2024Q1 将成品交付客户使用。

**扬杰科技：车规全国产主驱 SiC 模块计划于 2025 年批量上车。**车载模块方面，公司自主开发的 HPD 以及 DCM 全碳化硅主驱模块将在 2023 年完成 A 样试制，目前已经获得多家 Tier1 和终端车企的测试及合作意向，计划于 2025 年完成全国产主驱碳化硅模块的批量上车。第三代半导体产品的持续推出，为公司实现半导体功率器件全系列产品的一站式供应奠定了坚实的基础。

**新洁能：部分 SiC MOSFET 已批量销售，GaN HEMT 部分产品已开发完成。**1200V SiC MOSFET 产品方面，公司已完成 23mohm~62mohm 系列产品开发。650V SiC MOSFET 产品方面，公司新开发了部分产品应用于新能源汽车 OBC、光伏储能、工业及自动化等行业，相关产品通过客户验证，并实现小规模销售。GaN HEMT 部分产品开发完成并通过可靠性测试。

## 9 投资建议

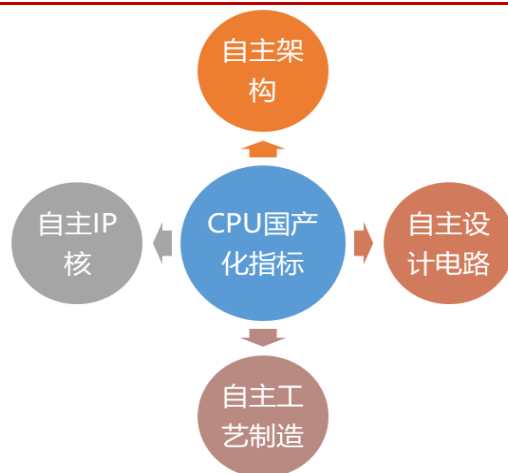
在终端出货量见底回升的基本面见底的大背景下，我们认为“乐观想象”将是市场的内在逻辑，看好“空间大、渗透率低、创新产品”相关的板块估值提升。

### 9.1 国产化纯度较高的 CPU 公司

国产化指标主要有四个：架构国产化、IP 核国产化、电路设计国产化、工艺制造国产化。拥有国产化指标较多的公司，国产化的程度越高。

推荐关注龙芯中科，龙芯中科是上市公司里面满足四大国产化指标的公司。

图 63 国产化重要指标



资料来源：首创证券

### 9.2 关注 AI 终端的机会

主芯片厂商发布支持 AI 计算 SoC 芯片，AI 手机有望拉动换机需求。同时，行业去

库存完成，消费芯片有望率先复苏带动行业复苏。

### 9.3 关注苹果 MR 产品 Vision Pro 正式量产的机会

2023 年 6 月 6 日，苹果发布首款 MR 头显 Apple Vision Pro，等该产品量产上市后，有望带动 MR 产业发展，以及二级市场 MR 相关投资机会。

## 10 风险提示

基本面见底，但底部时间较长。短期手机、电脑等终端出货量降幅收窄，中期出货量或保持稳定。

下游终端销售不及预期，向上游电子元器件厂商压价，影响毛利率。终端销售不及预期，终端厂商为了维持自身的盈利能力，向上游芯片供应商压价，导致芯片供应商的毛利率承压。

主动去库存，导致产业链各环节产能闲置。各公司 2023 年通过减少订单保持自身存货处于合理水平，若需求不能持续增加，产业链各环节出现产能闲置。



## 分析师简介

何立中，电子行业首席分析师，北京大学硕士，曾在比亚迪半导体、国信证券研究所、中国计算机报工作。

## 分析师声明

本报告清晰准确地反映了作者的研究观点，力求独立、客观和公正，结论不受任何第三方的授意或影响，作者将对报告的内容和观点负责。

## 免责声明

本报告由首创证券股份有限公司（已具备中国证监会批复的证券投资咨询业务资格）制作。本报告所在资料的来源及观点的出处皆被首创证券认为可靠，但首创证券不保证其准确性或完整性。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。投资者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、税收等方面咨询专业财务顾问的意见。对依据或者使用本报告所造成的一切后果，首创证券及/或其关联人员均不承担任何法律责任。投资者需自主作出投资决策并自行承担投资风险，任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

本报告所载的信息、材料或分析工具仅提供给阁下作参考用，不是也不应被视为出售、购买或认购证券或其他金融工具的要约或要约邀请。该等信息、材料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，首创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。

首创证券的销售人员、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。首创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。首创证券的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。

在法律许可的情况下，首创证券可能会持有本报告中提及公司所发行的证券头寸并进行交易，也可能为这些公司提供或争取提供投资银行业务服务。因此，投资者应当考虑到首创证券及/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信赖依据。

本报告的版权仅为首创证券所有，未经书面许可任何机构和个人不得以任何形式转发、翻版、复制、刊登、发表或引用。

## 评级说明

	评级	说明
1. 投资建议的比较标准 投资评级分为股票评级和行业评级  以报告发布后的 6 个月内的市场表现为比较标准，报告发布日后的 6 个月内的公司股价（或行业指数）的涨跌幅相对同期的沪深 300 指数的涨跌幅为基准	股票投资评级	买入 相对沪深 300 指数涨幅 15% 以上
		增持 相对沪深 300 指数涨幅 5%-15% 之间
		中性 相对沪深 300 指数涨幅 -5% -5% 之间
		减持 相对沪深 300 指数跌幅 5% 以上
2. 投资建议的评级标准 报告发布日后的 6 个月内的公司股价（或行业指数）的涨跌幅相对同期的沪深 300 指数的涨跌幅为基准	行业投资评级	看好 行业超越整体市场表现
		中性 行业与整体市场表现基本持平
		看淡 行业弱于整体市场表现