



华安证券
HUAAN SECURITIES

证券研究报告

算力新需求下的投资机遇

分析师：胡杨 S0010521090001

日期：2023年5月5日

华安证券研究所

● 主要观点

● AI算力需求放量下，从BOM表看电子产业链机会

ChatGPT用户数量惊人的增速，叠加模型能力的持续提升，模型对算力的需求持续扩张，AI算力需求增长趋势确定。我们预计全球AI服务器出货量将于2027年增长至77万台，CAGR达30%。从AI服务器BOM表的拆解中可以看到，主要受益方向包括GPU、存储、PCB以及电源管理类模拟芯片等

● IC载板或为PCB单机价值增量中最大赢家，关注格局更佳的ABF载板细分领域

按照单颗价格100美金计算，AI服务器IC载板在PCB单机价值量中占比高达46%。此外，由于GPU算力要求的提高，OAM与UBB等GPU模组中的核心部件对PCB板等级要求对应提高，单位面积价格大幅上涨，高层板厂商受益。IC载板行业既有出货量的长期成长性，也有供需影响下的价格周期波动。而GPU等高性能芯片使用的ABF载板在该周期成长赛道中拥有足够大的占比和更好的竞争格局，国产替代的时间节点更早，投资机会更为确定

● 存储周期筑底中，供给端收缩明显，需求端HBM贡献边际增量

历史复盘，存储厂商毛利率与存储行业周期同步，当前处于周期底部。供给端主要IDM玩家同步缩减资本开支，降低稼动率；需求端手机尽管边际复苏偏弱，但单机存储用量有望随去库存周期推进而重新提高；服务器则或在23Q2逐步见底。新需求维度，AI服务器DDR需求远高于普通服务器，到27年将贡献28亿美金需求增量；而AI服务器的HBM需求更为旺盛，有望在27年贡献新需求45亿美金。供给控制下，需求端的复苏有望提高周期拐点的确定性

● 设备是算力的基石，光刻机受限下多重曝光工艺需求增加，驱动设备需求进一步放量

半导体下行周期中代工厂稼动率结构分化，先进制程需求受AI影响稼动率仍在高位。受地缘政治影响，国内或通过多重曝光等工艺解决先进制程需求问题，带来涂胶显影、刻蚀、沉积等设备使用量的提升。结合其他产线扩产加速，我们认为国内扩产预期22年同比下滑-5%，好于市场的悲观预期的-15%。23年仍处于国产化率从20%提升到30%的放量期。因此，23年仍是半导体设备公司的投资窗口期

● 建议关注：兴森科技、华正新材、深南电路、沪电股份、兆易创新、北京君正、东芯股份、芯源微、中微公司、华海清科、拓荆科技

● 风险提示：1) 整体下游需求复苏不及预期；2) AI需求增量边际贡献低于预期



● 主要公司财务情况汇总

		市值 (亿元)	净利润 (亿元)			PE			FY22业绩回顾 (亿元)						
			2023E	2024E	2025E	2023E	2024E	2025E	营业收入	YoY	归母净利润	YoY	毛利率	净利率	ROE
PCB	兴森科技	240	4.92	6.80	9.90	48.8	35.3	24.2	53.54	6.23%	5.26	-15.42%	28.66%	9.10%	8.04%
	华正新材	37	0.98	2.49	3.76	37.8	14.8	9.8	32.86	-9.23%	0.36	-84.85%	12.98%	1.24%	2.15%
	深南电路	396	17.66	21.26	24.65	22.4	18.6	16.1	139.92	0.36%	16.40	10.75%	25.52%	11.72%	13.39%
	沪电股份	386	16.04	18.07	20.43	24.1	21.4	18.9	83.36	12.37%	13.62	28.03%	30.28%	16.33%	16.47%
存储	兆易创新	699	13.28	19.66	26.41	52.6	35.6	26.5	81.30	-4.47%	20.53	-12.16%	47.66%	25.25%	13.52%
	北京君正	397	8.23	10.66	13.77	48.2	37.2	28.8	54.12	2.61%	7.89	-14.79%	38.56%	14.39%	7.03%
	东芯股份	135	2.6	3.91	6.18	51.9	34.5	21.8	11.46	1.03%	1.85	-29.17%	40.58%	18.97%	4.72%
设备	芯源微	244	2.62	3.9	5.84	93.1	62.6	41.8	13.85	67.12%	2.00	158.77%	38.40%	14.45%	9.50%
	中微公司	1052	14.26	18.08	22.26	73.8	58.2	47.3	47.40	52.50%	11.70	15.66%	45.74%	24.64%	7.55%
	华海清科	382	7.43	9.96	12.98	51.4	38.4	29.4	16.49	104.86%	5.02	152.98%	47.72%	30.42%	10.47%
	拓荆科技	496	5.42	8.08	10.95	91.5	61.4	45.3	17.06	125.02%	3.69	438.09%	49.27%	21.35%	9.93%

资料来源: Wind, 华安证券研究所整理



目录

1

总量与BOM表：算力新需求下的供应链机会概览

2

IC载板：PCB单机价值量中容易忽视的大赢家

3

存储：周期底部下，HBM带来需求端边际增量

4

设备：超市场悲观预期下的新工艺需求挖掘

5

新电源解决方案机会：多相控制器/Drmos/芯片电感

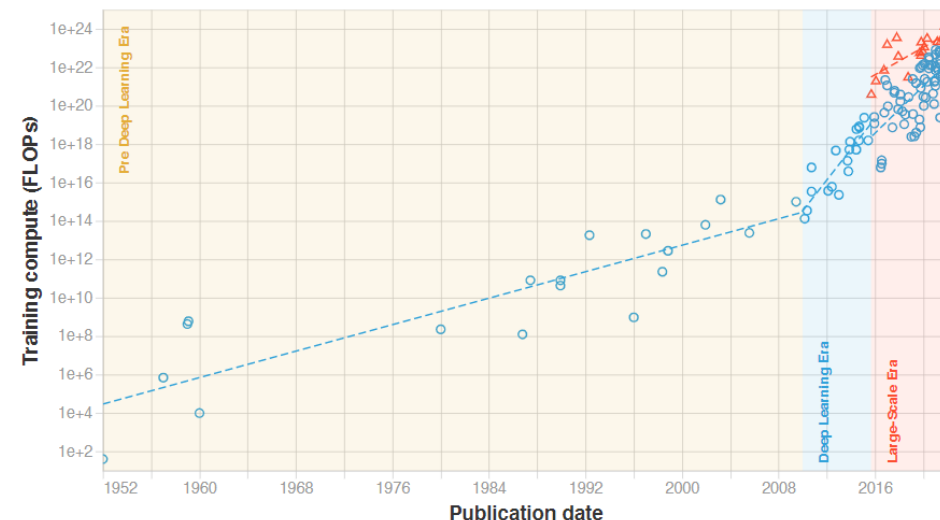
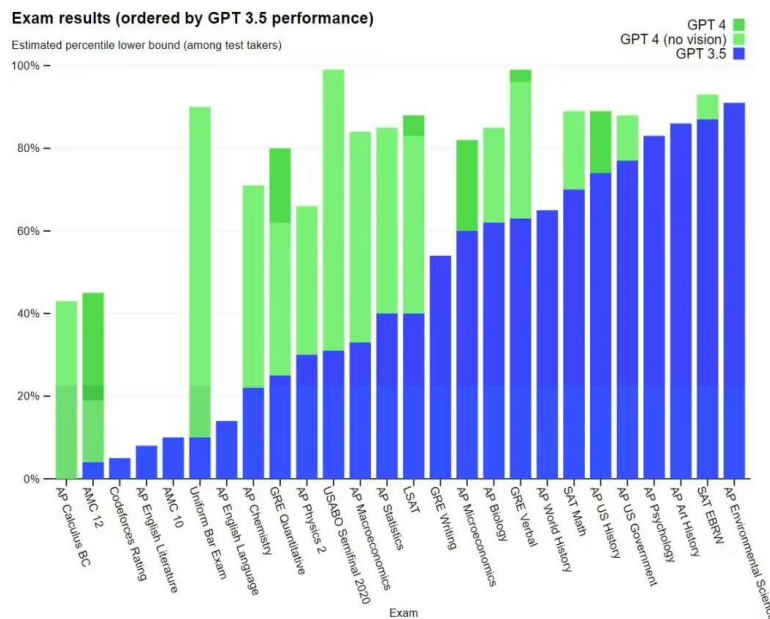
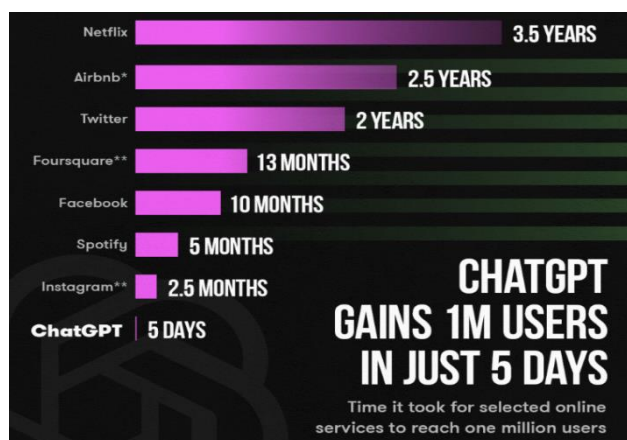
● AI需求持续增加带来算力需求放量

- **ChatGPT用户数量增速惊人：**ChatGPT用户量仅用5天就触及百万用户，约2个月时间完成了月活1亿用户的积累，渗透速度已初现加速态势
- **模型能力持续提升：**3月15日凌晨，OpenAI正式推出GPT-4文本生成AI系统。GPT-4是多模态模型，输入可以是文字，还可以是图像，输出形式为文本。相比GPT-3.5，GPT-4更可靠、更有创意，对于更细微的指令的处理能力更强。在各种专业测试和学术基准上，GPT-4与人类水平相当。它通过了模拟律师考试，且分数在应试者的前10%左右（GPT-3.5的得分在倒数10%左右）
- **伴随模型升级，算力需求持续提高：**2010年之前，训练算力的增长率符合摩尔定律，约每20个月翻一番；2010年深度学习来临，训练算力的增长率大幅度提升，大约每6个月翻一番；2015年末，随着许多公司开始开发大规模的机器学习模型，对训练算力的要求提高了10到100倍

ChatGPT用户数量暴增

ChatGPT 4相较ChatGPT 3.5具有效果表现

新模型需要更高的计算量



资料来源：TRTWorld, CSDN, 华安证券研究所

● 我们预计全球AI服务器出货量将于2027年增长至77万台，2023~2027年AI服务器需求量CAGR约30%

训练阶段所需GPU数量

$$= \frac{\text{总训练计算量}}{\text{训练时长} \times \text{GPU可提供的算力} \times \text{GPU算力利用率}}$$

$$= \frac{\text{模型个数} \times \text{单个模型所需计算量}}{\text{训练时长} \times \text{GPU可提供的算力} \times \text{GPU算力利用率}}$$

$$= \frac{\text{模型个数} \times (\text{模型参数} \times \text{训练集大小} \times \text{单个token所需运算次数})}{\text{训练时长} \times \text{GPU可提供的算力} \times \text{GPU算力利用率}}$$

推理阶段所需GPU数量

$$= \frac{\text{总推理计算量}}{\text{推理时长} \times \text{GPU可提供的算力} \times \text{GPU算力利用率}}$$

$$= \frac{\text{模型个数} \times \text{单个模型所需计算量}}{\text{推理时长} \times \text{GPU可提供的算力} \times \text{GPU算力利用率}}$$

$$= \frac{\text{模型个数} \times (\text{模型参数} \times \text{训练集大小} \times \text{单个token所需运算次数})}{\text{推理时长} \times \text{GPU可提供的算力} \times \text{GPU算力利用率}}$$

注：黄底为假设

训练阶段	2023E	2024E	2025E	2026E	2027E
模型个数 (个)	15	20	25	30	35
模型参数 (亿)	3000	3000	3000	3000	3000
训练集大小 (亿tokens)	4990	4990	4990	4990	4990
每个参数每token需要的FLOPs	6	6	6	6	6
总训练计算量 (Flops)	1.35E+25	1.796E+25	2.246E+25	2.695E+25	3.144E+25
总训练计算量 (TFlops)	1.35E+13	1.80E+13	2.25E+13	2.69E+13	3.14E+13
训练时长 (秒)	2.59E+06	2.59E+06	2.59E+06	2.59E+06	2.59E+06
英伟达A100训练算力 (FP16, TFLOPS)	312	312	312	312	312
GPU算力利用率	40%	40%	40%	40%	40%
GPU需求 (万张)	4.16	5.55	6.94	8.33	9.72
AI服务器需求 (万台, 8路GPU)	0.52	0.69	0.87	1.04	1.21
推理阶段	2023E	2024E	2025E	2026E	2027E
网站单日访问次数 (亿)	40	52	68	88	114
每次访问所进行对话来回的次数	2	2	2	2	2
每次处理的问答所需的token数量 (个)	500	500	500	500	500
模型参数量 (亿)	3000	3000	3000	3000	3000
每个token所需的FLOPs	2	2	2	2	2
总推理计算量 (FLOPs)	2.4E+24	3.12E+24	4.056E+24	5.273E+24	6.855E+24
总推理计算量 (TFLOPs)	2.4E+12	3.12E+12	4.056E+12	5.273E+12	6.855E+12
英伟达T4推理算力 (FP16, TFLOPS)	65	65	65	65	65
GPU算力利用率	40%	40%	40%	40%	40%
推理时长 (秒)	86400	86400	86400	86400	86400
GPU需求 (万张)	106.84	138.89	180.56	234.72	305.14
AI服务器需求 (万台, 4路GPU)	26.71	34.72	45.14	58.68	76.28
总计	2023E	2024E	2025E	2026E	2027E
GPU需求 (万张)	111	144	187	243	315
AI服务器需求 (万台)	27	35	46	60	77

基于现有部分宣布将使用AI大模型的公司数量进行假设，假设每家公司使用1个模型。根据我们统计，截至23年4月，已宣布推出AI模型的公司包括阿里、商汤、腾讯、百度、华为、Google、Microsoft等至少15家。假设未来越来越多的科技公司将推出AI模型

20年发布的GPT-3拥有1750亿参数，22年谷歌大部的PaLM拥有5400亿参数。取二者折中，假设23年AI模型平均参数约3000亿

根据Cornell大学所发布的论文《Language Models are Few-Shot Learners》，其作者在训练模型时使用的数据集合计约4990亿tokens，假设目前企业所使用的训练集与其持平

根据OpenAI官网数据，每参数每token的训练成本通常约6 FLOPs

1TFLOPs=10¹²FLOPs

假设训练需要30天，30天=30天*24小时*60分钟*60秒=2.59* (10⁶)秒

根据Virtaitech，GPU算力利用率通常在10~30%。但英伟达联合发布的论文中表示通过使用训练的并行技术可以将A100 GPU算力利用率优化到52%，因此，取折中，假设GPU算力约40%

根据英伟达，NVIDIA DGX-1（英伟达所推出首个专为数据中心中的企业打造的AI系统）配备了8个英伟达A100 GPU。假设每台AI服务器搭载一个NVIDIA DGX-1系统，则每台AI服务器需要8个英伟达A100 GPU

根据Semrush，2023年3月全球网页访问次数最大的是Google，23年3月Google每天访问次数约37.63亿次。由于不止Google一家推出AI模型，因此，假设23年网站单日访问次数约40亿并逐年增长

假设每次用户在访问时需要与AI进行2次问答来回才能得到想要的结果

在英语语境下，一般1000个token等于750个单词。假设每次处理的问答字数约375个单词（即500个token）

20年发布的GPT-3拥有1750亿参数，22年谷歌大部的PaLM拥有5400亿参数。取二者折中，假设23年AI模型平均参数约3000亿

根据OpenAI官网数据，每参数每token的推理成本通常约2 FLOPs

官网参数

根据Virtaitech，GPU算力利用率通常在10~30%。但英伟达联合发布的论文中表示通过使用训练的并行技术可以将A100 GPU算力利用率优化到52%，因此，取折中，假设GPU算力约40%

假设推理需要花一天，1天=1天*24小时*60分钟*60秒=86400秒

根据英伟达，NVIDIA DGX Work Station配备了4个GPU，假设推理阶段每台AI服务器搭载的系统与NVIDIA DGX Work Station类似，均需搭载4颗GPU

●

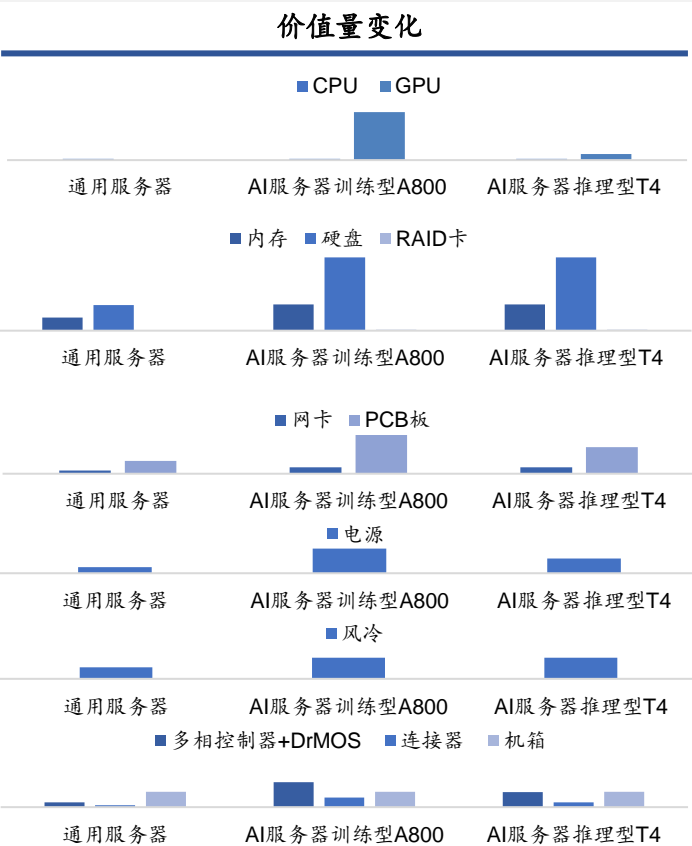
拆解AI服务器BOM表寻求产业链受益方向

- 主要受益方向：GPU、存储、PCB、电源及部分电源管理类模拟芯片
- 由于存储相关的DRAM以HBM模式封装于GPU中，价值量核算到GPU分项，DRAM实际增量更高；同理，GPU使用的IC载板核算在GPU中，其巨大的增量使得PCB行业受益程度比BOM表表观增量更为可观

		通用服务器	AI服务器训练型A800	AI服务器推理型T4
高性能计算资源	CPU	2张 英特尔5318 18526元	2张 英特尔5318 18526元	2张 英特尔5318 18526元
	GPU		8张 英伟达A800 764568元	4张 英伟达T4 96600元
内存&存储	内存	12根 内存条 12420元	24根 内存条 24840元	24根 内存条 24840元
	硬盘	5~10块 SSD（500美元/块） 24150元	20块 SSD 69000元	20块 SSD 69000元
	RAID卡	低配RAID卡 500元	低配RAID卡 878元	低配RAID卡 878元
网络连接	网卡	1张 双口万兆网卡 870元	2张 双口万兆网卡 1740元	2张 双口万兆网卡 1740元
	PCB板	8~10层 M6板为主 3400元	18~20层 M8 10350元	14~16层 M6 7140元
电源	电源	2颗 800~1200W电源 3100元	4颗 1800W高功率电源 12400元	4颗 1600~1800W电源 7440元
散热	风冷	6组 风扇 300元	风扇数量增多，扇叶面积做大 550元	风扇数量增多，扇叶面积做大 550元
其它	多相控制器+DrMOS	467元	新增GPU所需用量8套 2292元	新增GPU所需用量4套 1380元
	连接器	200元	900元	440元
	机箱	1400元	1400元	1400元
合计价值量		65333元	907444元	229934元

资料来源：华安证券研究所整理

敬请参阅末页重要声明及评级说明





目录

1 总量与BOM表：算力新需求下的供应链机会概览

2 IC载板：PCB单机价值量中容易忽视的大赢家

3 存储：周期底部下，HBM带来需求端边际增量

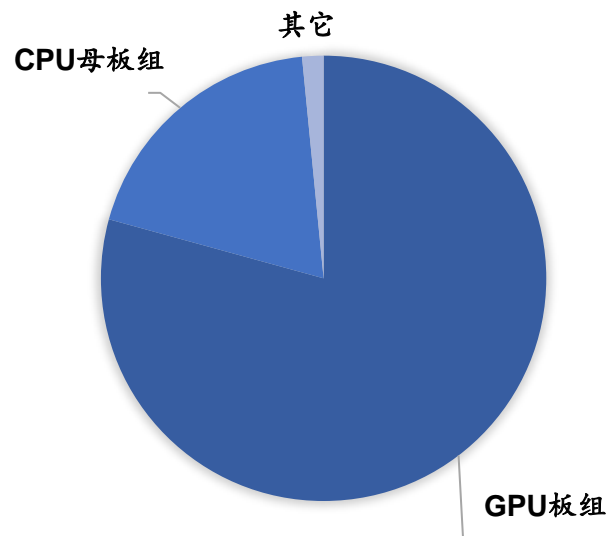
4 设备：超市场悲观预期下的新工艺需求挖掘

5 新电源解决方案机会：多相控制器/Drmos/芯片电感

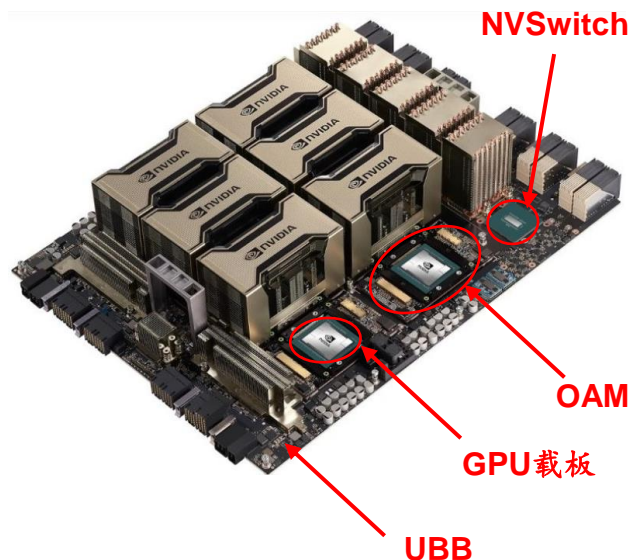
● AI服务器PCB板结构分析

- 以英伟达DGX A100服务器为例，整机PCB按照价值量进行粗分，主要由CPU母版组和GPU板组构成。其中，前者价值量占比19%；后者占比高达79%。GPU板组是相对普通服务器的主要增量
- GPU板组的结构：**UBB底板上承载8块GPU加速卡和6颗NVSwitch芯片；每块GPU加速卡需要一块承载板OAM，同时，承载的GPU芯片本身也需要对应的IC载板，共8颗；同理，每颗NVSwitch芯片各需要一块对应的IC载板，共6颗
- CPU母版组结构：**CPU母版价值量主要集中在主板以及CPU载板，其他功能性配板包括内存卡、网卡、拓展卡、存储操作系统驱动板等。其中，以内存卡和网卡的面积相对较大

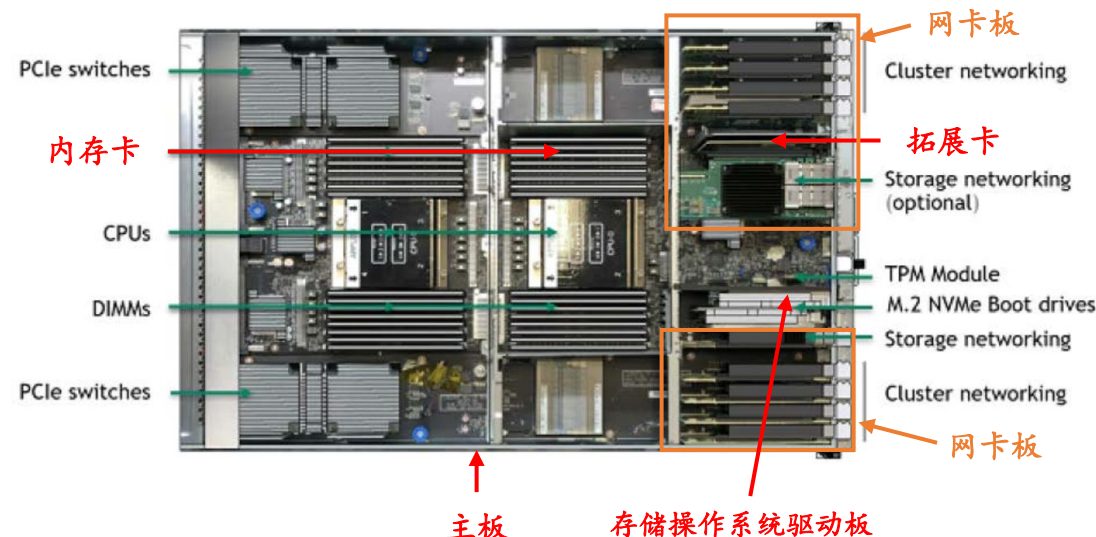
英伟达DGX A100 PCB价值量分布



英伟达DGX A100 GPU板组



英伟达DGX A100 CPU母版组

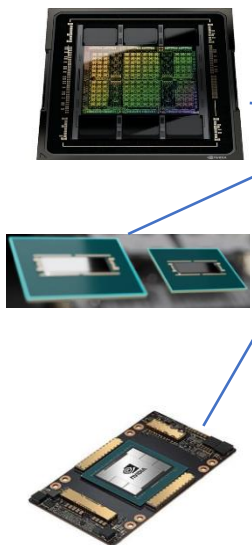


资料来源：I THome、英伟达官网，华安证券研究所整理

● AI服务器PCB板价值量分析—IC载板或为最大赢家

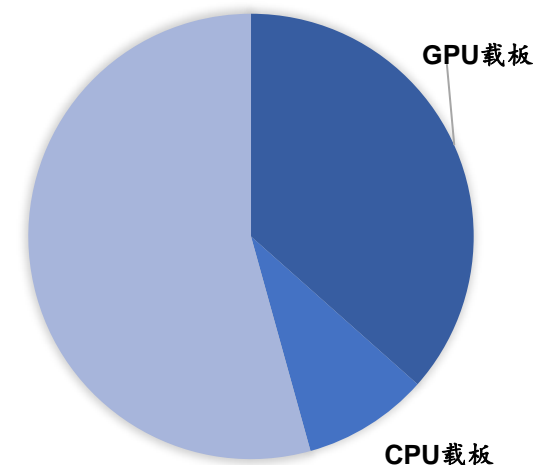
- **载板价值量占比高：**GPU载板与CPU载板单价高达100美元/颗，单机按照8颗GPU，2颗CPU进行计算，合计在PCB单机价值量中占比高达46%。考虑到GPU载板价值量占比37%为AI服务器相对普通服务器的纯增量，更凸显IC载板在算力供应链中的重要性
- **OAM与UBB高单价印证PCB需求等级提高：**GPU板组中的OAM与UBB面积大，其合计价值量与GPU载板接近。从单价来看，UBB高达132美金/平方英尺，价格是普通配件PCB板的6-10倍；而OAM板分中高端配置与高端配置，对应单价93~160美金/平方英尺不等。具体对应的PCB板等级从M6升级大M8，层数从8~10层升级到18~20层，带动价格持续提高
- **PCB层数及型号都持续上升，主要原因是AI运算的复杂性：**高性能服务器通常具有更复杂的设计，需要处理更多的信号和电源路径，多层板有利于路径布局。更高的信号完整性要求方面，高层PCB板可以提供更多的屏蔽层，减少信号干扰和反射，在高速数据传输中提高信息传递质量。其他还有电源管理及散热需求，高层PCB板提供更多的电源平面和散热通道，更好实现电源分布和管理，并将热量传输到散热器。以及电磁兼容性的需求，多层PCB可以更好地控制电磁干扰和射频噪声，从而提高设备的EMC性能

英伟达DGX A100 PCB价值量拆解



	单台DGX A100服务器用量	单价	单机价值量
GPU板组			
GPU载板	8颗	100美元/颗	800美元
NVSwitch	6颗	30美元/颗	180美元
OAM (OCP Accelerator Module)	2.58平方英尺	中低端配置: 93美元/平方英尺 高端配置: 160美元/平方英尺	中低端配置: 240美元 高端配置: 413美元
UBB (Unit Baseboard)	3.23平方英尺	132美元/平方英尺	426美元
合计			1646~1819美元
CPU motherboard			
CPU载板	2颗	100美元/颗	200美元
CPU主板	4.09平方英尺	40美元/平方英尺	164美元
功能性配板			
CPU内存卡	1.38平方英尺		
网卡	1.29平方英尺		
拓展卡	0.11平方英尺		
存储操作系统驱动板	0.11平方英尺		
功能性配板总计	2.89平方英尺	20美元/平方英尺	58美元
合计			422美元
其它配件			
电源	1.23平方英尺	13~20美元/平方英尺	16~25美元
硬盘	0.69平方英尺	13~20美元/平方英尺	9~14美元
前控制台板	0.12平方英尺	13~20美元/平方英尺	1.6~2.4美元
合计			26.6~41.4美元
共计 (GPU板组+CPU motherboard+其它)			2094.6~2282.4美元

GPU载板和CPU载板的价值量占比

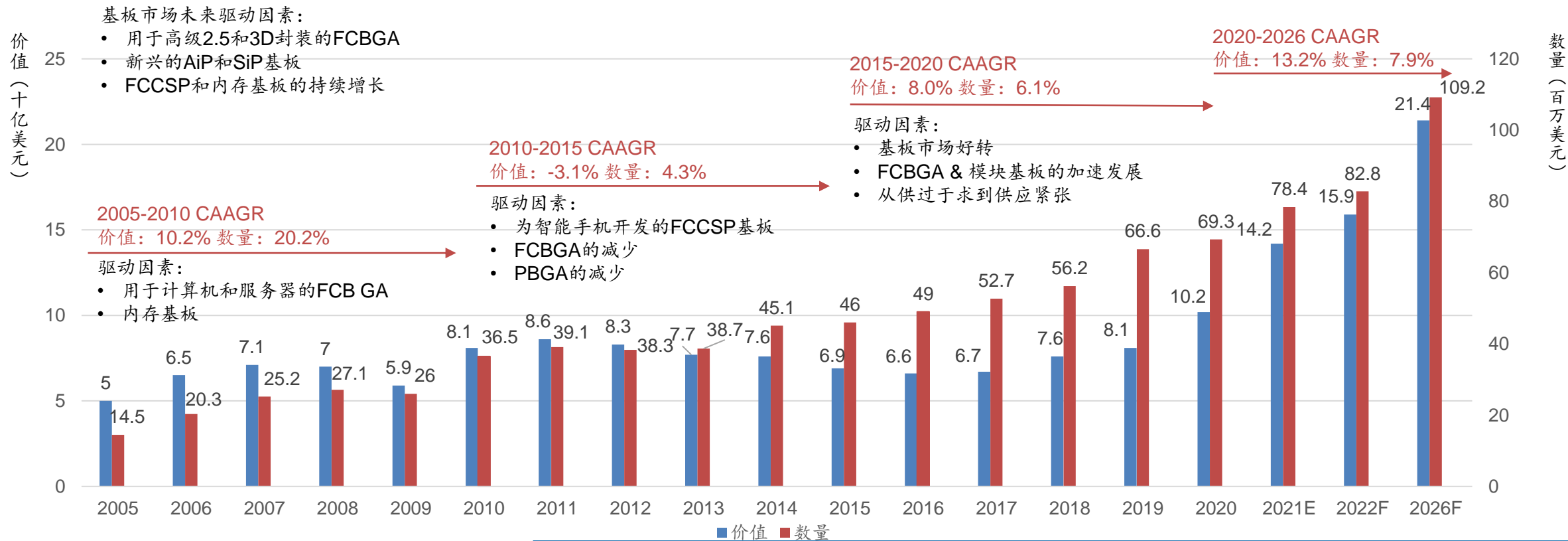




● 载板行业复盘—供需情况带来价格扰动，长期成长确定性高

- 纵观05年至今，载板行业出货量整体呈现持续增长的态势，主要驱动因素包括计算机和服务器带动的FCBGA载板放量，智能手机带动的FCCSP载板放量，以及模块化带来的基板需求增加等。未来，随着下游产品周期的放量以及芯片集成度不可逆的持续提高，载板行业将维持持续的增长态势。
- 过去在10-15年智能手机爆发期，市场规模曾出现过负增长的阶段，主要是供需周期带来的价格波动影响。因此，行业的跟踪仍需关注供需情况变化

载板市场规模及出货量复盘



资料来源：Prismark，华安证券研究所

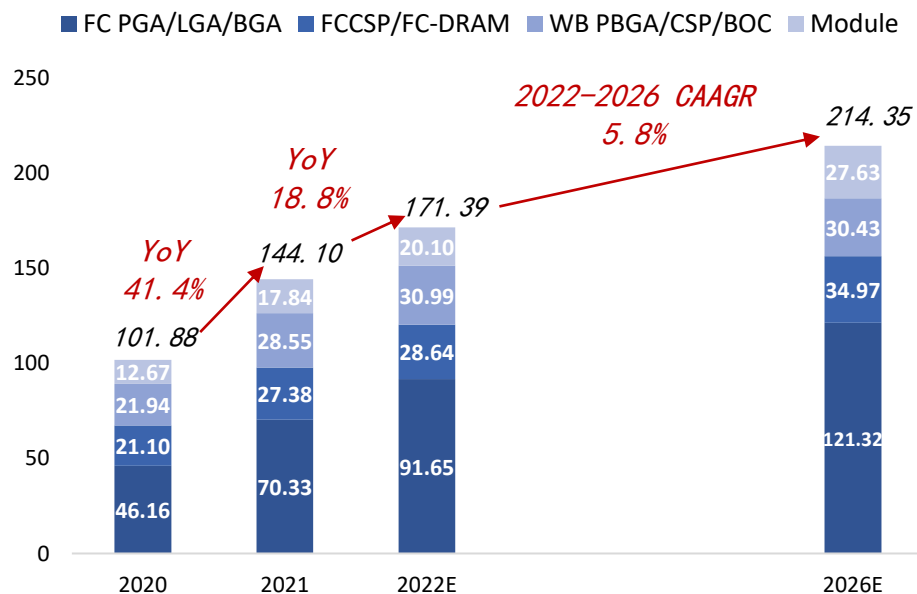
敬请参阅末页重要声明及评级说明



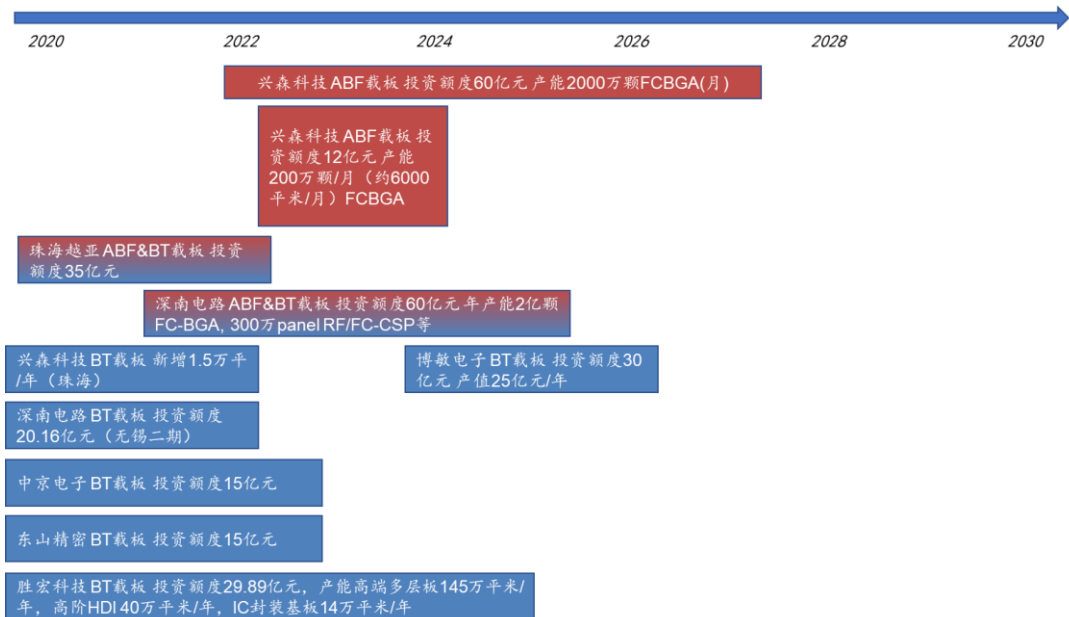
● 载板市场规模与国内进度

- **市场规模较大：**载板市场全球规模22年预计达约170亿美金，未来4年CAGR 5.8%。其中，以FC PGA/LGA/BGA载板(从基材分类角度属于ABF载板类)占比最大，占约90亿美金，未来四年CAGR 7.2%，最为可观。ABF载板的高速成长主要驱动力是高级2.5D和3D封装的持续成长带动，主要应用范畴为GPU/CPU/ASIC等。
- **竞争格局良好：**ABF载板由于生产集中在台湾（占比40%），日韩（占比50%）和德系厂家，生产环境需要80%无尘室，国产替代刚开始，使得产品毛利率较高（40%左右，高阶可以做到80%）
- **国内主要厂商投资进度：**兴森科技ABF载板两期投资额度总额高达72亿元，珠海越亚35亿和深南电路60亿元投资含ABF和BT载板。相对而言，国内BT载板整体投资额度更大，进展更快；ABF载板格局与前景更优

载板市场空间及增速（亿美元）



本土载板厂进度





目录

1

总量与BOM表：算力新需求下的供应链机会概览

2

IC载板：PCB单机价值量中容易忽视的大赢家

3

存储：周期底部下，HBM带来需求端边际增量

4

设备：超市场悲观预期下的新工艺需求挖掘

5

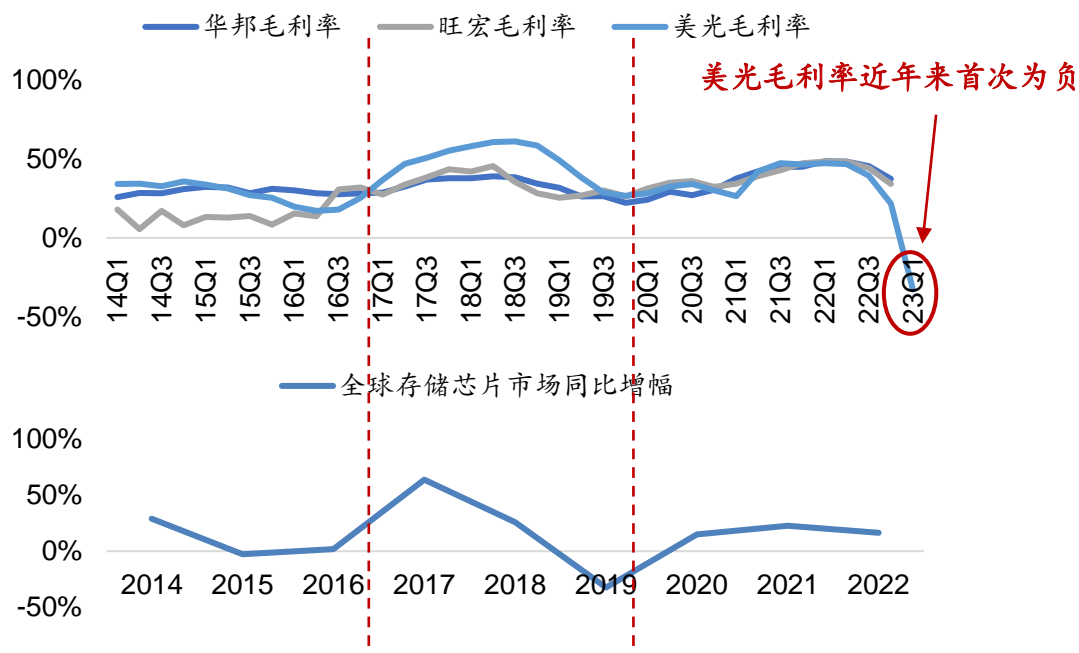
新电源解决方案机会：多相控制器/Drmos/芯片电感



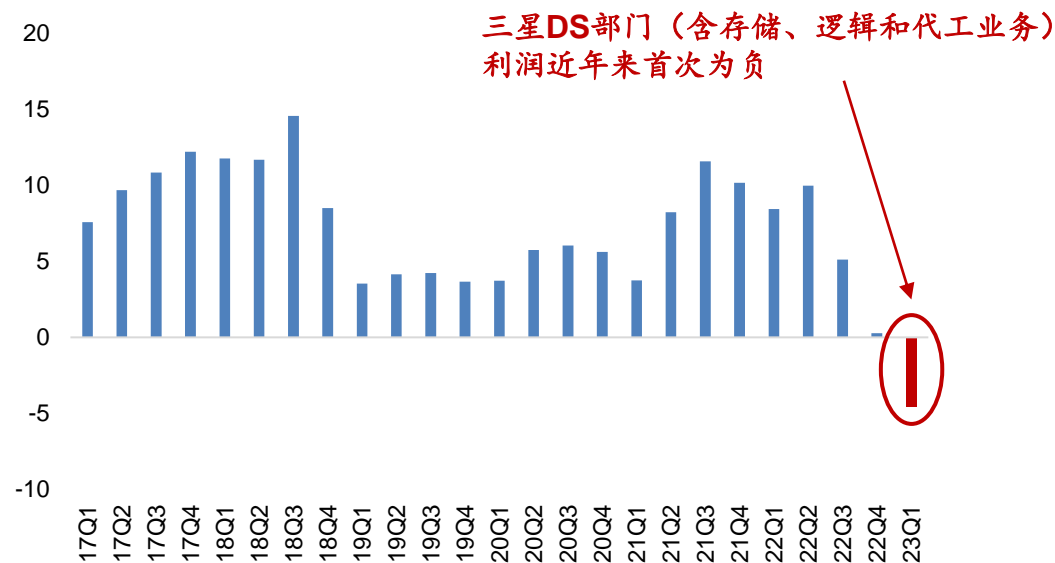
● 盈利能力揭示存储行业当前处于周期底部的位置

- 通过对比存储厂商历史毛利率变化与存储芯片市场增速变化，我们发现，存储厂商毛利率与存储芯片周期波动一致。2022年，受下游市场需求疲软影响，存储厂商库存水位逐渐堆高，为应对库存问题，存储厂商纷纷选择通过降价以在一定程度上刺激库存的消耗。从当前毛利率来看，美光CY23Q1毛利率为-32.66%，同比下滑79.87pct，环比下滑54.52pct，已处于其历史低点；同时，美光预计其CY23Q1的库存将为峰值。此外，根据三星4月27日发布的公告，公司DS部门（含存储、逻辑和代工业务）营业亏损4.58万亿韩元，创近5年来最低

存储厂商毛利率走势与存储周期同步



三星DS部门营业利润近6年来首次亏损（万亿韩元）

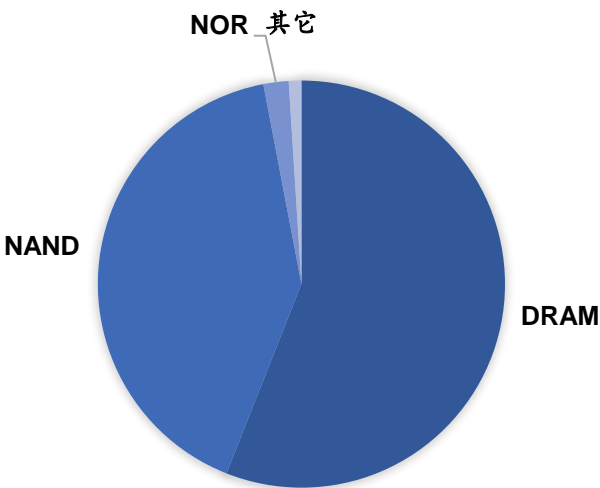


资料来源：Wind、华经产业研究院、公司公告，华安证券研究所

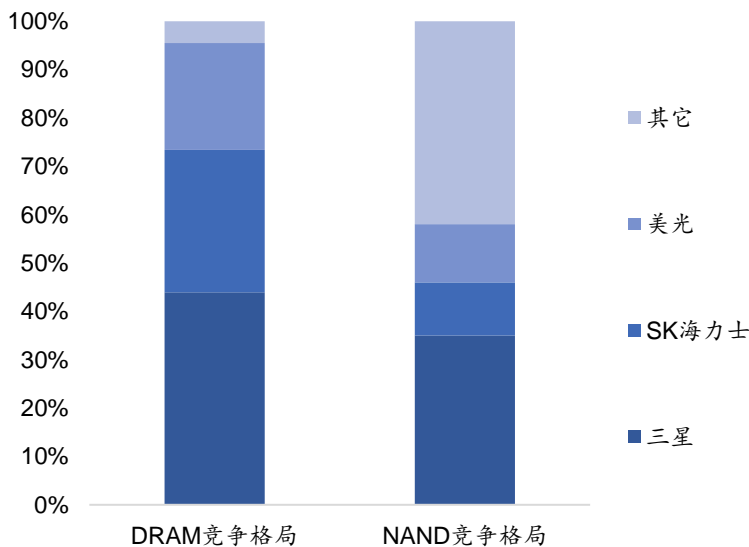
● 供给：边际改善，龙头厂商缩减资本开支并进减产

- DRAM和NAND是最大的存储市场（根据市场规模，DRAM/NAND合计占存储市场的97%），三星、美光、SK海力士又是在这两个市场占据最大份额的厂商
- 2022年，由于市场需求疲软，美光与海力士纷纷计划削减23年的资本支出40%-50%不等，同时进行存储产品减产。三星虽未宣布削减其23年的资本支出，但已于23年4月宣布减产。我们认为，供给端的边际改善有望为夯实周期底部，减缓存储价格下降提供必要条件

DRAM和NAND是最大的存储市场



三星/美光/海力士是市占率最大的厂商



龙头厂商的减产和缩减资本开支规划

	减少产出	降低投资
三星	于23年4月宣布减产	可能灵活调整2023年设备方面的资本支出
SK海力士	将围绕收益较低的存储产品进行减产	2023年的投资规模从15至20万亿韩元减少50%以上
美光	减少20%的NAND Flash 和 DRAM的晶圆产出	2023年资本开支将减少到70-75亿美元（原120亿）。2024年也预期减少

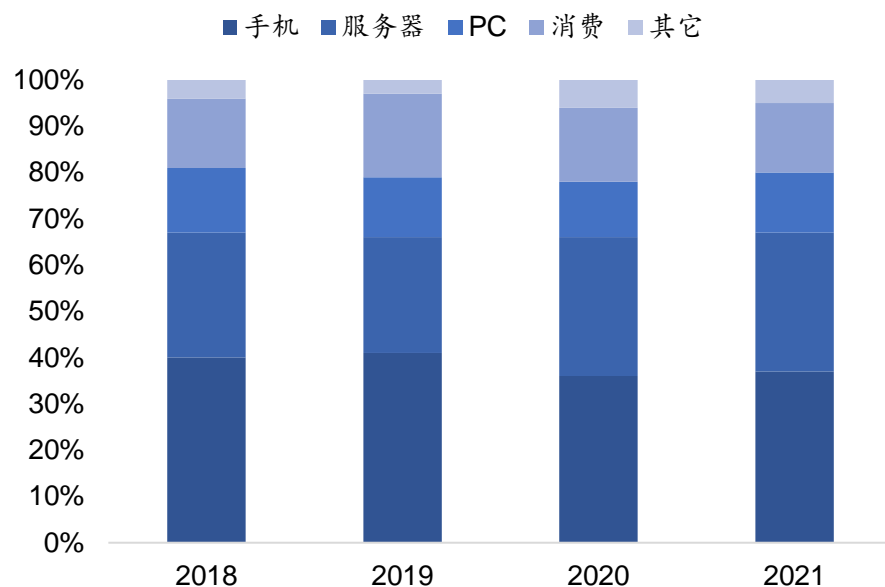
资料来源：DRAMExchange、CFM闪存市场，华安证券研究所



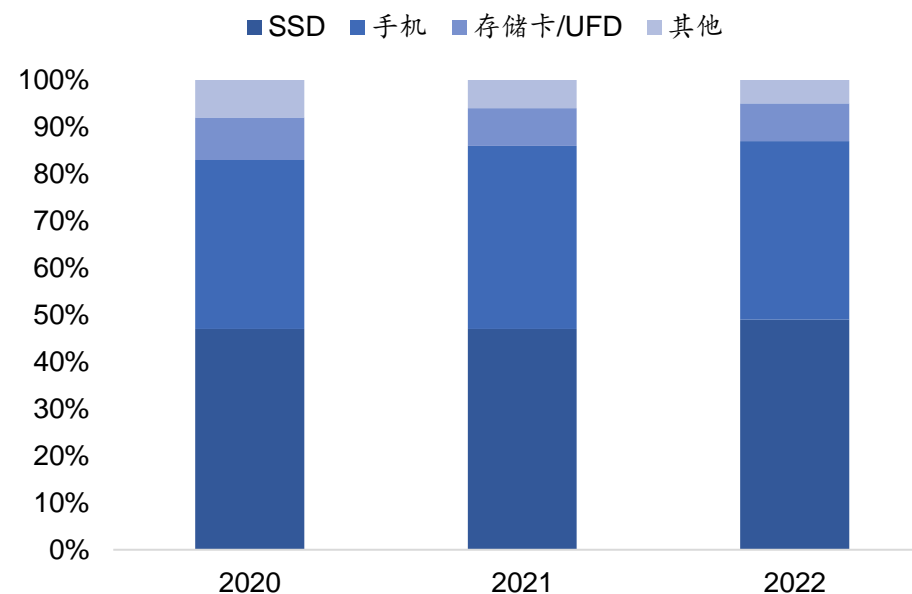
● 需求：主要观察手机和服务器两大下游需求复苏情况

- 从下游应用来看，手机和服务器是在DRAM和NAND市场中占比最大的应用
- DRAM角度，手机占比37%，服务器占比30%。NAND角度，手机占比39%，服务器与PC的SSD整体占比高达47%。因此，手机与服务器的需求恢复情况对存储芯片景气度有着较大的影响。而AI服务器对存储的新需求也会对整体需求改善有边际贡献

DRAM下游应用占比



NAND下游应用占比



资料来源：华经产业研究院、CFM闪存市场，华安证券研究所



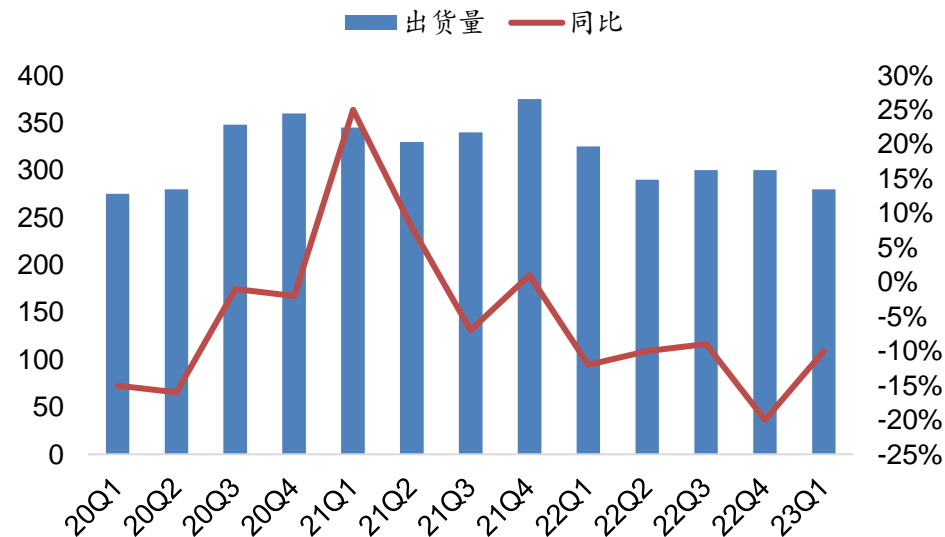
● 手机：出货量未见明显复苏，单机容量提升是主要驱动要素

虽然目前手机出货量没有看到明显的增长，但Q1仍出货量同比跌幅环比略有收窄。随着经济逐渐复苏，手机需求将逐步好转。并且，手机单机DRAM和NAND容量的增长也在一定程度上能够冲抵手机的疲软需求：

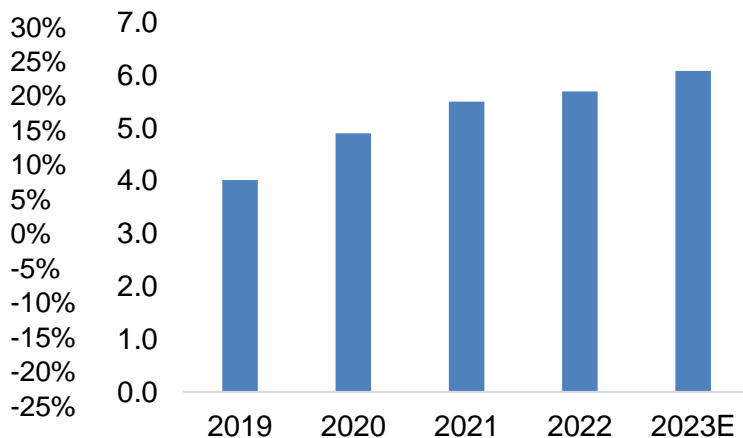
①DRAM方面，2022年手机DRAM容量增长不足的主要原因在于手机品牌厂背负巨大的库存压力，导致新产品导入多以沿用既有库存规格为主，限制手机用DRAM容量增长。但展望23年，随着手机库存去化逐渐成效，以及受益于下一代iPhone规格容量升级带动，trendforce预计23年手机DRAM平均搭载容量同比增长6.7%，优于22年的3.9%

②NAND方面，为满足高像素拍摄需求，需要配置更大的储存容量，为智能手机单机NAND Flash搭载量提升带来基本动能。iPhone产品组合仍全线往更高容量靠拢；Android高端机种也跟进将512GB做为标准配备，中低端机种储存空间则随硬件规格持续升级而提高，因此单机容量仍有增长空间。据Trendforce预估，2023年智能手机NAND Flash单机搭载容量年成长仍能维持22.1%

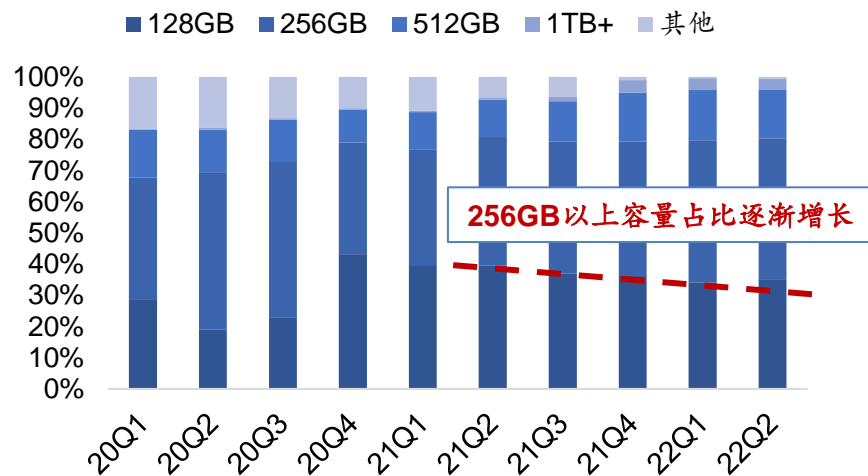
智能手机单季度出货量（亿部）



单机DRAM容量逐年提升 (GB)



800美元以上高端手机的NAND容量逐步增长

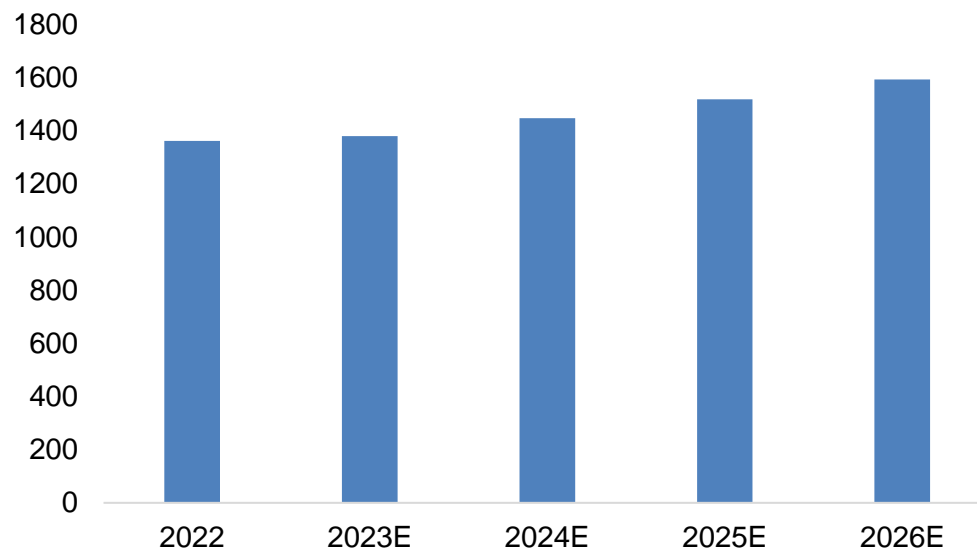




● 普通服务器：出货情况有望于23H2出现好转

- 虽然服务器厂商的库存去化尚未结束，但根据美光CY23Q1法说会，预计数据中心将于23Q2触底，且库存将于23年底达到相对健康的水平。英业达也在23年3月的法说会中表示，预计服务器第二季度、第三季度能见度有望提高，且库存水位已达到健康水平。因此，预计全球服务器出货量有望于23H2出现好转

全球服务器出货量（万台）

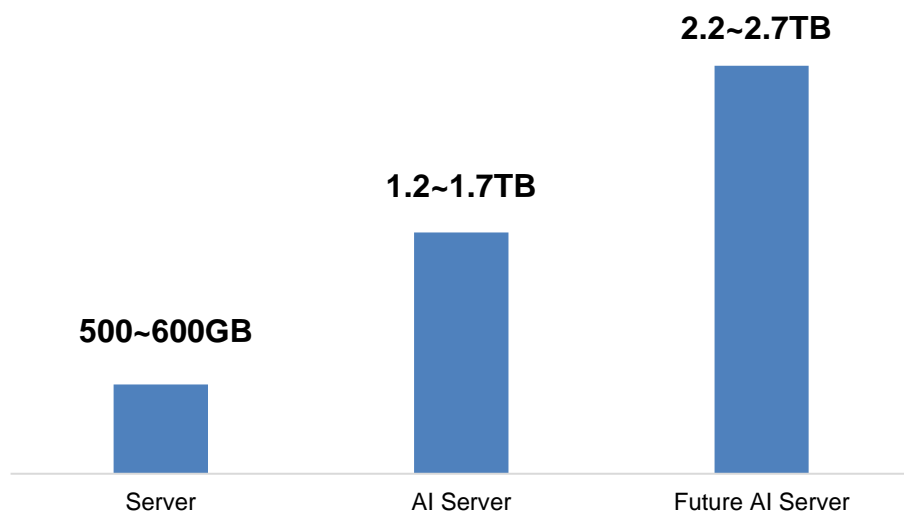


资料来源：Trendforce、Research and Markets，华安证券研究所

● AI服务器带来的价值增量—DDR：预计至27年市场规模约27.82亿美元，2023~2027 CAGR约29.89%

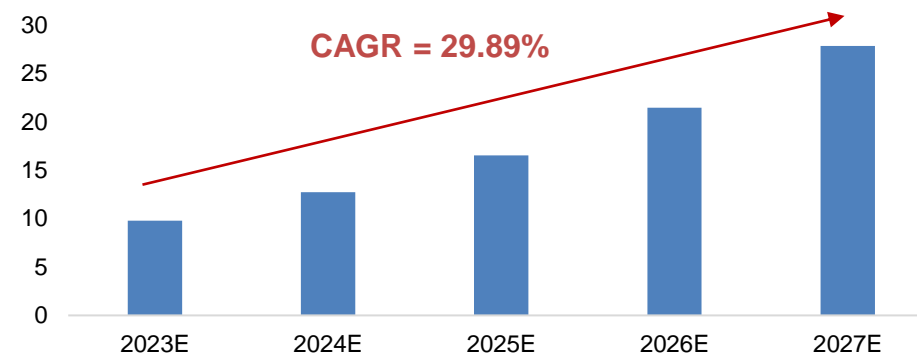
- 根据Trendforce，AI服务器所需DRAM容量相比普通服务器具有明显提升。根据我们测算，单个普通服务器所需DRAM（不含HBM）价格约1.24万元，单个训练型AI服务器和推理型AI服务器所需DRAM（不含HBM）价格均约2.48万元。基于我们此前对AI服务器需求量的测算，我们预计27年AI服务器用DRAM（不含HBM）市场约27.82亿美元，2023~2027年市场规模CAGR约29.89%

AI服务器所需DRAM容量（不含HBM）远高于普通服务器



注：黄底为假设

AI服务器所需DRAM容量（不含HBM）市场规模测算（亿美元）

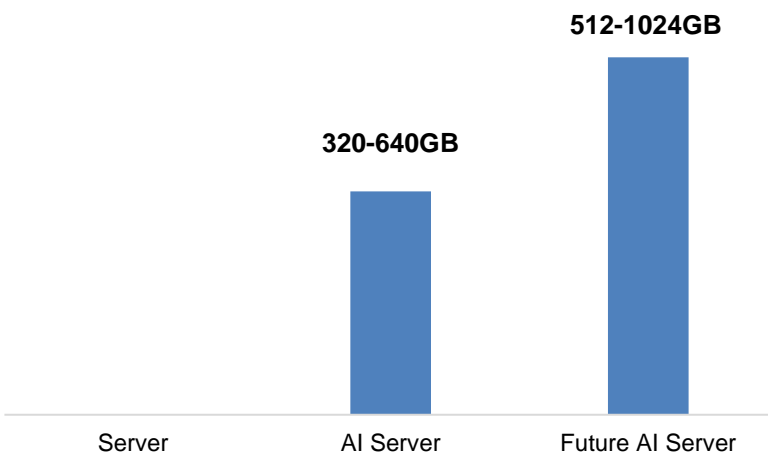


	2023E	2024E	2025E	2026E	2027E
①训练端					
训练端AI服务器出货量（万台）	0.52	0.69	0.87	1.04	1.21
单台训练端AI服务器所使用的DRAM价格（万元/台）	2.48	2.48	2.48	2.48	2.48
训练端AI服务器DRAM市场规模（亿元）	1.29	1.72	2.16	2.59	3.02
②推理端					
推理端AI服务器出货量（万台）	26.71	34.72	45.14	58.68	76.28
单台推理端AI服务器所使用的DRAM价格（万元/台）	2.48	2.48	2.48	2.48	2.48
推理端AI服务器DRAM市场规模（亿元）	66.35	86.25	112.13	145.76	189.49
合计AI服务器DRAM市场规模（亿元）	67.64	87.97	114.28	148.35	192.51
合计AI服务器DRAM市场规模（亿美元）	9.77	12.71	16.51	21.44	27.82

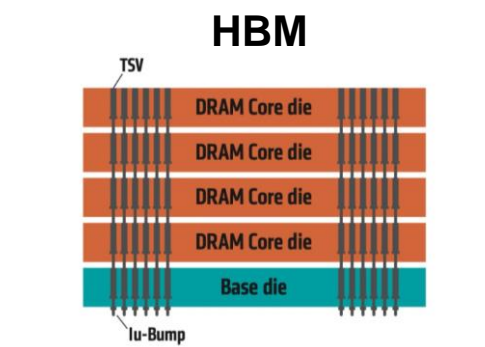
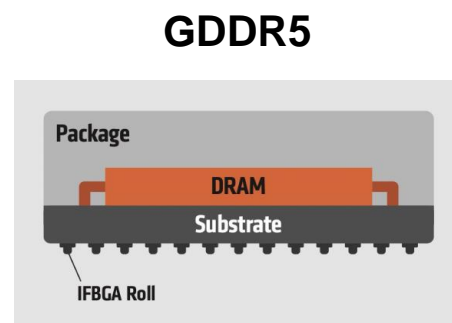
● AI服务器带来的价值增量—HBM：HBM已成为高端GPU标配，是AI服务器所带来的全新增量

- **HBM的高带宽更能满足AI服务器的需求，已在云端高性能服务器领域成为高端GPU的标配，是AI服务器所带来的全新增量。**AI服务器需要在短时间内处理大量数据，对高带宽需求大幅提升。而HBM（High Bandwidth Memory）的特点之一便是相比DDR/GDDR，可以以更高的效率实现更高的传输带宽，因此，HBM已成为高端GPU标配
- 从结构上看，HBM属于3D结构，每个HBM封装内部都堆叠了多层DRAM die，因此，HBM的互联宽度相比采用2D结构的DDR更大。从传输位宽的角度来看，HBM的每层DRAM die是2个128bit通道，4层DRAM die高度的HBM内容总共就是1024bit位宽。作为对比，GDDR5内存每通道位宽32bit，16通道则合计512bit，远低于HBM的位宽。而HBM现在已升级至HBM3，可堆叠的DRAM die层数进一步增加，在容量和速度方面又有了提升

HBM是AI服务器带来的全新增量



HBM采用3D结构，可堆叠更多DRAM die，可实现更小尺寸



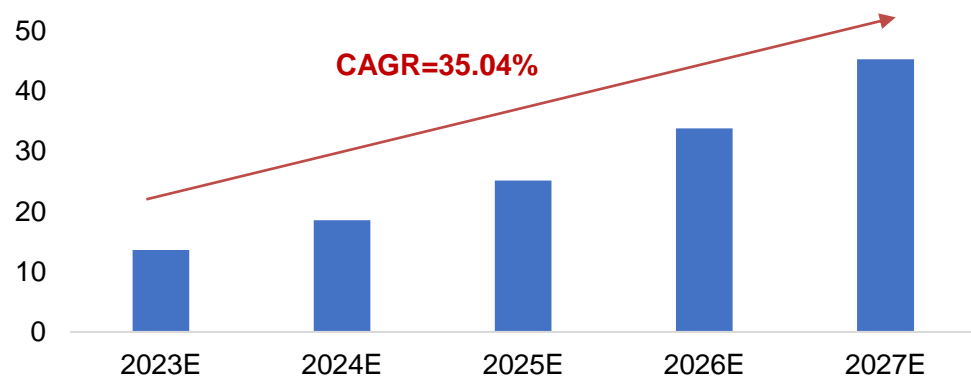
相比DDR等2D结构DRAM，HBM具有更高带宽

	LPDDR4	LPDDR5	DDR4	DDR5	GDDR5	HBM2
结构	2D	2D	2D	2D	2D	3D
Max I/F BW	34GB/s	51GB/s	25.6GB/s	51GB/s	32GB/s	307GB/s
应用	汽车，消费电子	汽车，消费电子	服务器，PC，消费电子	服务器，PC，消费电子	图像	图像，HPC

● AI服务器带来的价值增量—HBM：预计至27年市场规模约45.28亿美元，2023~2027 CAGR约35.04%

- 根据yole，HBM单GB价格约10美元。基于Trendforce对于AI服务器所需HBM容量的预测，结合我们对于AI服务器需求量的测算，我们预计27年AI服务器用HBM市场约45.28亿美元，2023~2027 CAGR约35.04%

全球HBM市场规模测算（亿美元）



	2023E	2024E	2025E	2026E	2027E
AI服务器需求量 (万台)	27	35	46	60	77
单台AI服务器所需HBM用量(GB)	500	540	580	620	660
HBM单价 (美元/GB)	10.0	9.7	9.4	9.1	8.9
合计HBM市场规模 (亿美元)	13.62	18.55	25.11	33.79	45.28

假设1：根据Trendforce，目前AI服务器平均所需HBM容量约320~640GB，未来将提升至512~1024GB，假设当前AI服务器搭载的容量约500GB，且逐年提升

假设2：根据yole，HBM 23年单价约10美元/GB。假设因规模效应，成本逐渐下降

注：黄底为假设

全球HBM市场规模敏感性分析（亿美元）

基于推理端变量，对全球AI服务器（训练+推理）用HBM市场规模所做的敏感性分析		推理端-网站单日访问次数（亿次）				
		30	35	40	35	50
推理端-每次访问所进行对话来回的次数（次）	1	5.27	6.10	6.94	6.10	8.61
	2	10.28	11.95	13.62	11.95	16.95
	3	15.28	17.79	20.29	17.79	25.30
	4	20.29	23.63	26.97	23.63	33.65
	5	25.30	29.47	33.65	29.47	41.99

- 我们预计全球服务器用存储芯片市场规模将于2027年达到372.98亿美元，HBM占其12%
- 根据我们测算，全球服务器用DRAM市场规模将于27年增长至372.98亿美元，HBM将占其12%。若按照全球DRAM市场约700亿美元来估算，则27年AI服务器用HBM市场将占全球DRAM市场的6.5%。

服务器用DRAM（不含HBM）市场规模					
	2023E	2024E	2025E	2026E	2027E
普通服务器					
普通服务器出货量（万台）	1380	1448	1518	1593	1671
单台普通服务器所使用的DRAM价格（万元/台）	1.24	1.24	1.24	1.24	1.24
普通服务器DRAM市场规模（亿元）	1713.71	1797.80	1885.85	1978.26	2075.19
合计普通服务器DRAM市场规模（亿美元）	247.65	259.80	272.52	285.88	299.88
AI服务器					
①训练端					
训练端AI服务器出货量（万台）	0.52	0.69	0.87	1.04	1.21
单台训练端AI服务器所使用的DRAM价格（万元/台）	2.48	2.48	2.48	2.48	2.48
训练端AI服务器DRAM市场规模（亿元）	1.29	1.72	2.16	2.59	3.02
②推理端					
推理端AI服务器出货量（万台）	26.71	34.72	45.14	58.68	76.28
单台推理端AI服务器所使用的DRAM价格（万元/台）	2.48	2.48	2.48	2.48	2.48
推理端AI服务器DRAM市场规模（亿元）	66.35	86.25	112.13	145.76	189.49
合计AI服务器DRAM（不含HBM）市场规模（亿美元）	9.77	12.71	16.51	21.44	27.82
合计服务器DRAM（不含HBM）市场规模（亿美元）	257.42	272.51	289.04	307.31	327.70
服务器用HBM市场规模					
	2023E	2024E	2025E	2026E	2027E
合计HBM市场规模（亿美元）	13.62	18.55	25.11	33.79	45.28
HBM占服务器DRAM市场					
	2023E	2024E	2025E	2026E	2027E
服务器DRAM市场规模（HBM+DDR）（亿美元）	271.04	291.06	314.14	341.11	372.98
HBM占服务器DRAM（HBM+DDR）比	5%	6%	8%	10%	12%
HBM占DRAM市场比例 （按照全球DRAM市场700亿美元来测算）	1.9%	2.7%	3.6%	4.8%	6.5%

- 基于Trendforce和Research and Markets对服务器出货量进行的预测数据
- 基于BOM表中的数据
- CAGR ≈ 4.90%
- 基于此前我们对于AI服务器的测算
- 基于BOM表中的数据
- 基于此前我们对于AI服务器的测算
- 基于BOM表中的数据
- CAGR ≈ 29.89%
- CAGR ≈ 6.22%
- CAGR ≈ 35.04%
- CAGR ≈ 8.31%

- “AI服务器带来的新增量+厂商减产”有望进一步促进供需平衡，市场预计存储芯片价格跌幅将于23Q2开始收窄
- 2023年，随着经济的复苏，下游需求有望逐渐恢复。并且，受益于ChatGPT的出现所催生对于HBM的全新需求，三星与SK海力士的HBM订单自2023年初以来大幅增加，叠加三星23年4月宣布的减产消息，我们认为，需求上的增长与供给上的减少将进一步造成供需差缩窄，进一步强化促进存储芯片景气度拐点到来的确定性
- 第三方角度，根据Trendforce，预计DRAM价格跌幅将于23Q2开始收窄

市场预计DRAM价格跌幅将于23Q2收窄

	23Q1E	23Q2F
PC DRAM	DDR4:down 15-20% DDR5:down 18-23% Blended ASP:down 15-20%	DDR4:down 8-13% DDR5:down 10-15% Blended ASP:down 10-15%
Server DRAM	DDR4:down 20-25% DDR5:down 23-28% Blended ASP:down 20-25%	DDR4:down 13-18% DDR5:down 15-20% Blended ASP:down 13-18%
Mobile DRAM	Down 13-18%	Down 10-15%
Graphics DRAM	Down 18-23%	Down 10-15%
Consumer DRAM	Down 18-23%	Down 10-15%
合计 DRAM	Down ~20%	Down 10-15%

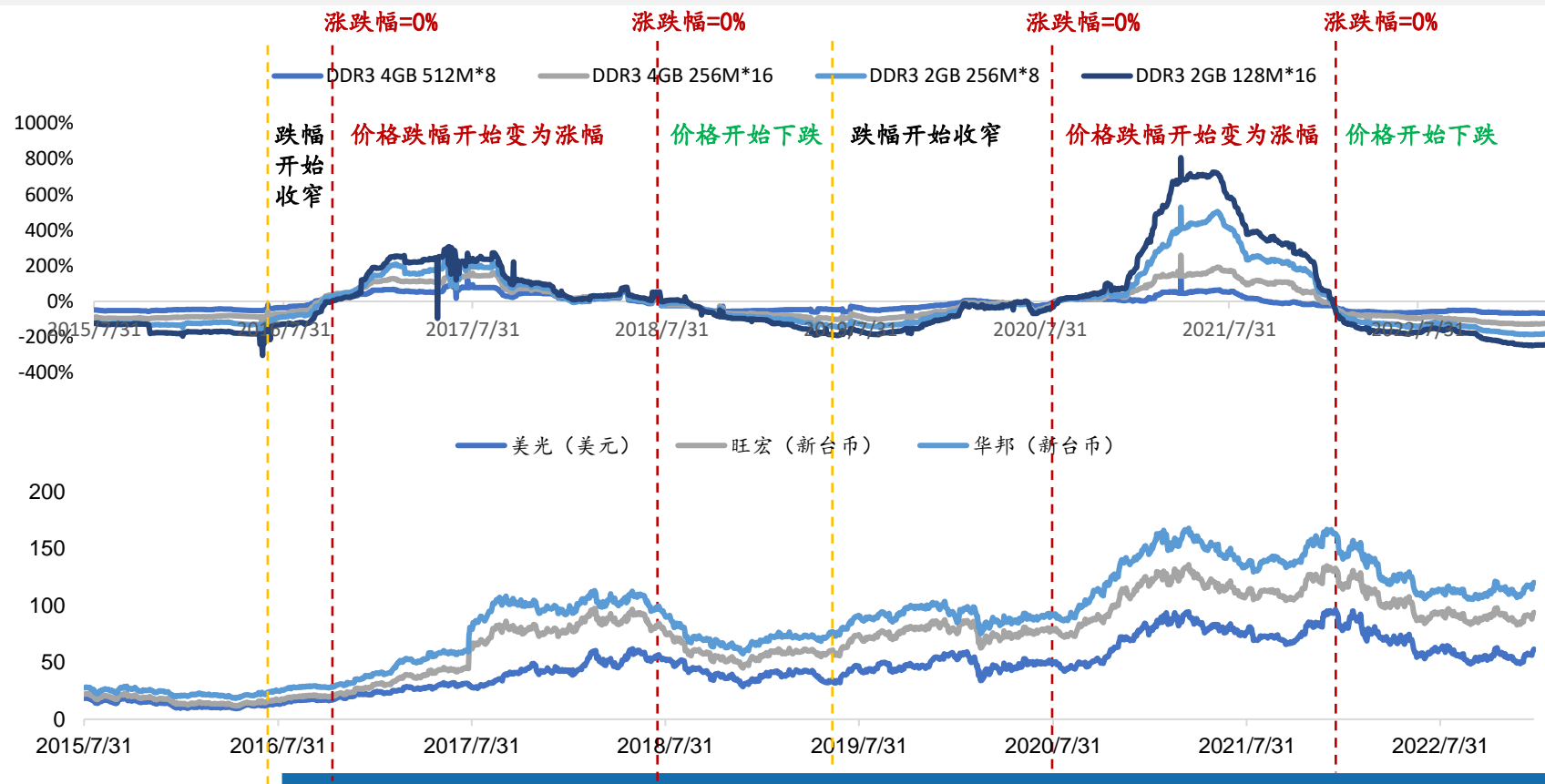
资料来源：Trendforce，华安证券研究所



● DRAM合约价格与股价走势相似，价格收窄意味着存储板块投资机会有望到来

- 从DRAM现货价格走势来看，DRAM现货价格走势与股价呈一定关联性，走势相似。当现货价格跌幅开始收窄时，股价便开始筑底反弹；当现货价格实现上涨时，股价也会同步上涨；在价格涨幅边际弱化时股价开始见顶
- 此外，DRAM/NAND厂商的价格走势与NOR厂商方向一致但幅度有所差异。因此，随着DRAM现货价格跌幅收窄，预计存储板块也将逐步复苏，迎来投资机会

DRAM现货价同比增幅





目录

1

总量与BOM表：算力新需求下的供应链机会概览

2

IC载板：PCB单机价值量中容易忽视的大赢家

3

存储：周期底部下，HBM带来需求端边际增量

4

设备：超市场悲观预期下的新工艺需求挖掘

5

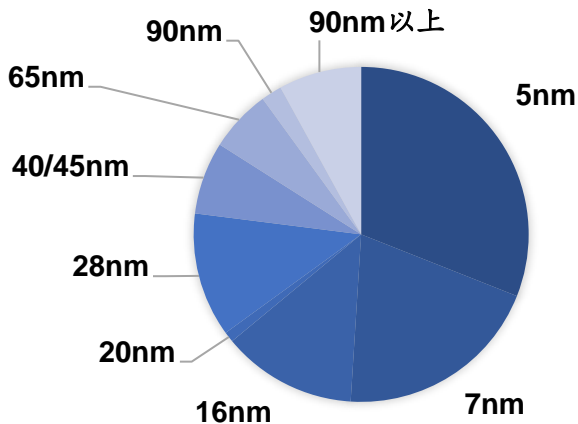
新电源解决方案机会：多相控制器/Drmos/芯片电感



● 晶圆厂工艺与半导体设备是支撑算力的地基

- 以GPU, CPU, FPGA, ASIC为代表的算力芯片, 极大程度上仰仗14/7nm以下先进制程的晶圆产能, 在全球半导体周期并未明显转好的前提下, 14nm及以下制程产能仍然维持着相对紧张, 如台积电高雄工厂28nm制程扩产暂停同时将转为先进制程。国内来看, 主流fab厂的产能利用率仍在下降通道中(22年Q4中芯国际产能利用率不足80%), 但中芯南方等14/28nm节点产线依然维持着相对较高的稼动率
- 算力芯片中, SOC, CPU, GPU等芯片对于先进制程工艺的需求较高, AI加速到来, 算力需求在云端, 边缘端以及终端多环节同步激增, 台积电的几大重点客户英伟达, Intel, 苹果等对应了大量的先进制程算力需求, 而随着国际环境的复杂, 国内AI芯片公司在上游环节受到限制, 大陆算力芯片需要很大程度上依靠国产fab产能和工艺。因此, 晶圆厂工艺配合国产设备将为国产算力芯片公司提供坚实的基础

台积电7nm以下制程已经占据50%以上收入构成 (23Q1)



国内算力GPU制程等参数与英伟达A100对比

厂商	英伟达	海光信息	摩尔线程	壁仞科技	天教智芯
型号	A100	深算一号	MTT S3000	壁砺100P	天垓100
制程	7nm	7nm FinFET	NA	7nm	7nm
核心数目	6912	4096	4096	NA	NA
时钟频率	0.77-141GH2	15-1.7GH2	1.9GH2	NA	NA
显存容量	40GB/80GB	32G8	32G8	64GB	32G8
显存类型	HBM2e	HBM2	GDDR6	HBM2E	DRAM HBM2
FP32 运算性能	19.5 TFLOPS	NA	15.2 TFLOP5	240TFLOPS (峰值)	37 TFLOPS
总线接口	PCIe 4.0 x16	PCIe Gen4 X 16	PCIe Gen5 x16	PCIe 5.0 X16	PCIe Gen4.0 x 16
TDP	250W	350W	s35W	450-55W	250W

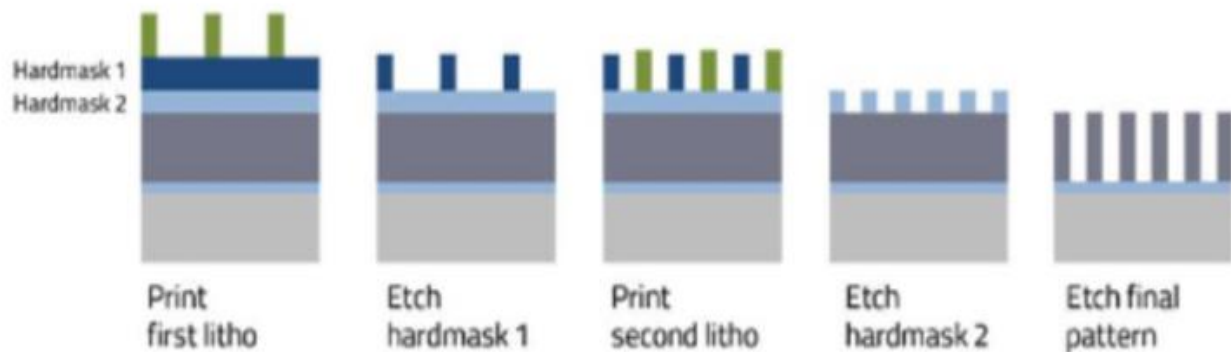
资料来源: 台积电财报, 各公司官网, 华安证券研究所



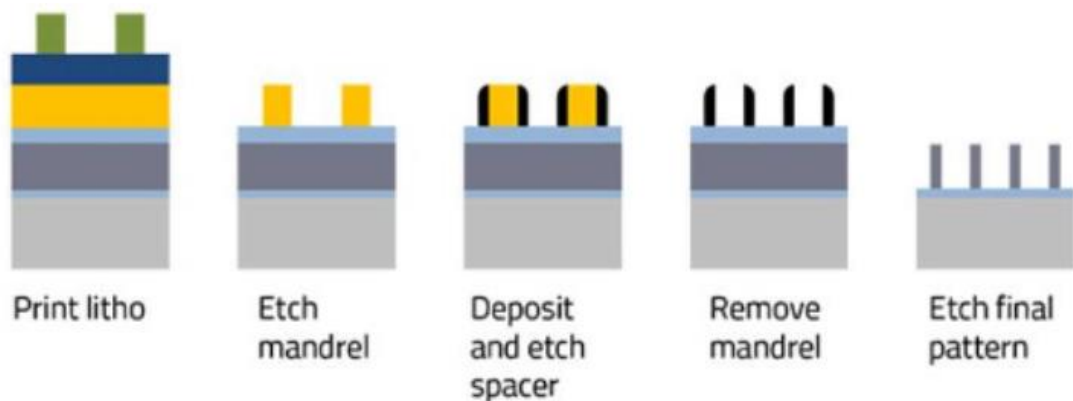
● 新需求：多重曝光与先进制程带来设备需求增加

- 算力芯片以及各类终端对于芯片性能需求的大幅提升，**国内在光刻机波长精度受限的情况下，光刻多重曝光技术带来了新的技术解决方案**。先利用浸没式光刻机形成节距较大的线条，再利用侧墙图形转移的方式形成1/2节距的线条，该技术适合线条排列规则的图形层，如FinFET工艺中的Fin或后段金属线条。**光刻多重曝光技术通常需要反复进行涂胶—光刻—显影—刻蚀等工艺流程，因此光刻多重曝光技术的发展大大增加了浸没式DUV光刻机，涂胶显影机，刻蚀机，薄膜沉积等设备的潜在需求**
- 对于国产先进工艺来说，光源波长为193nm的DUV光刻机搭配浸没式技术是海外可以供应的最先进机台，根据光刻机瑞利公式，DUV理论上用来制造28nm线宽芯片，而通过多重曝光技术，实际可以做到7nm的极限线宽，多重曝光如果广泛应用，将同步扩大中国市场对浸没式光刻机，涂胶显影机，刻蚀机，薄膜设备的需求

光刻-刻蚀两步骤重复 (Lam research)



多重曝光技术图解——获得更小线宽 (Lam research)



资料来源：LAM RESEARCH，华安证券研究所



● 需求端：晶圆厂扩产受地缘政治冲击可控

晶圆厂内资	尺寸	产线	现有产能	总规划产能	22新增	23新增	24新增
中芯国际	12寸	中芯北方	9	10	2	0	0
	12寸	中芯北京	5.2	6	0	0	0
	12寸	中芯京城	< 2	10	1	1	3
	12寸	中芯东方	0	10	0	0	2
	12寸	中芯深圳	0	4	0	2	2
	12寸	中芯西青	0	10	0	0	0
	12寸	中芯南方	2.5	7	1	1	0
	8寸	中芯上海	11.5	13.5	0	0	0
	8寸	中芯深圳	7	7	0	0	0
	8寸	中芯天津	11.5	15	3.5	0	0
	8寸	中芯绍兴	7.5	16	2.5	3	3
	8寸	中芯宁波	1.5	8	2	2	2
长江存储	12寸	长存一期	10	10	0	0	0
	12寸	长存二期	0	20	2 → 0	4 → 0	0
合肥长鑫	12寸	长鑫一期	7	12	2	2	1
	12寸	长鑫二期	0	12	进度延缓	2	3
	12寸	长鑫北京	3	10	0	0	0
合肥晶合集成	12寸	晶合一期	12	12	0	0	0
	12寸	晶合二期	0	25	0	0	3
华虹半导体	12寸	ICRD一期	2	2	0	0	0
	12寸	ICRD二期	1	2	1	1	0
	12寸	ICRD三期	0	4	0	0	2
	12寸	华虹无锡	6.5	9.5	1	2	0
	8寸	上海fab1-3	17.8	17.8	0	0	0
其他厂商	12寸				12	10	8
	8寸				12	10	5
12寸增量					20	21	24
8寸增量					20	15	10
约当8寸					65	62.25	64

算力2022年10月7日美国BIS新规扩大半导体设备进出口限制范围后，对大陆半导体上游产业链影响较大。虽然28nm以上不受制裁的产线占据fab厂的大多数，但是存储厂商，中芯京城等偏先进制程的fab厂进度出现延缓

但随着2023年初各大厂商对国产设备工艺平台的持续推进，以及2022年以来二三线晶圆厂扩产的加速，我们测算了部分厂商进度推延状态下的前道设备市场状况。虽然制裁确实影响了2022年的扩产幅度，但2023年fab扩产实际同比下滑程度并不明显，而随着23年下半年中芯京城等fab扩产招标的开启，24年大陆晶圆厂资本开支将会同比提升

1. 2022年相较于制裁前的乐观预期扩产需求下滑了约20%

2. 2023年相较于2022年同比预计下滑约5%以内

3. 中芯多个项目同时开工，预计2023年的资本开支将有很大一部分投入到土地建设工程上，24年之后随着周期反转整体，以及前期土地建设的完成，多个产线同时扩产迎接下一波上行周期，预计设备投入将会在24年之后大幅度回暖

4. 四大厂预计将是未来大陆60%-70%的产能贡献，也是80%以上的资本开支贡献



● 需求端：“-15% VS -5%”，中国大陆fab厂扩产或好于市场悲观预期

- 除了大陆市场自身状况以外，全球半导体市场需求也受到了周期性因素的扰动。从资本开支的角度看，2022年末全球部分重要企业均下调了投入预期，台积电将其2022年资本支出从400-440亿美元降至约360亿美元。SK海力士预计故决定削减明年资本支出50%以上，并对收益较低的存储产品进行减产。美光因需求下滑，放慢生产速度并削减资本支出，预计2023会计年度资本支出将减少约80亿美元，即至少减少30%，晶圆厂设备支出减少50%
- SEMI对2023年全球fab扩产趋势给出了较为悲观的预期：**其预计，2023年中国大陆地区扩产量约减少47%，全球fab扩产量将同比减少15%。
- 但根据我们测算将好于市场预期：2022年以来二三线晶圆厂（中小fab）扩产的加速，2023年fab扩产实际同比下滑程度并不明显（约同比下滑5%以内），而随着23年下半年中芯京城等fab扩产招标的开启，24年大陆晶圆厂资本开支将会同比提升**

各地区fab扩产变动幅度

	2022	2023E	2024E
美国	22%	3%	43%
中国	-6%	-47%	41%
欧洲&中东	140%	19%	14%
日本	-11%	-14%	55%
韩国	-11%	-25%	59%
SE Asia	17%	-15%	42%
中国台湾	20%	6%	6%
合计	6%	-15%	31%

SEMI预计中国大陆地区23年扩产量减少约47%，
但根据我们测算，23年FAB扩产实际同比下滑约在5%以内

资料来源：semi，华安证券研究所



● 供给端：设备的2023仍处于国产化率20%到30%的快速放量期

按照稳态市场空间计算的国产化率						
1000	300		(单位：亿美元)	(单位：亿元)		
半导体设备	分类	占比	全球市场规模	国内市场规模	国产体量	国产化率
晶圆制造前端设备 1010亿美元	光刻机	20%	200	420	-	<1%
	刻蚀设备	22%	220	462	54	12%
	ICP	11%	105	220.5	30	14%
	CCP	11%	105	220.5	24	11%
	薄膜沉积设备	22%	220	462	53	11%
	PECVD	7%	72.6	152.46	18	12%
	溅射PVD	4%	41.8	87.78	13	15%
	管式CVD	3%	26.4	55.44	5	9%
	ALD	2%	24.2	50.82	5	10%
	LPCVD	2%	24.2	50.82	5	10%
	EPI	1%	8.8	18.48	<1	0%
	MOCVD	1%	6.6	13.86	7	51%
	ECD/ECP镀铜	1%	6.6	13.86	0.5	4%
	其他	1%	6.6	13.86	2	14%
	量测设备	10%	100	210	7	3%
	清洗设备	6%	60	126	50	40%
	涂胶显影	4%	35	73.5	3.5	5%
	热处理	3%	25	52.5	20	38%
	离子注入	10%	100	210	5	2%
	CMP	3%	25	52.5	14	27%
	其他	2%	15	31.5	15	48%
总国产化率						16%
国产化（除光刻机）						20%

注：统计为2022预估销售额，从订单增速来看，预计23年整体国产化率会提高到25%以上

资料来源：公开资料整理，华安证券研究所

- 2022年，国产设备国产化率持续增加，根据我们测算，2022年整体国产化率达到16%。
- 离子注入，涂胶显影处在国产化起量的第一年**：这两类设备出货量在21/22两年从零到一的突破的同时，订单量也迅速攀升，具备基数低，国产化率提升快的特点，预计23年订单层面和出货量层面仍然将保持较快增长，国产化率覆盖度有望迅速提升
- 刻蚀机和薄膜设备的整体国产化率均在10%-15%之间**。光刻机，刻蚀机，薄膜设备是三类最主要的前道设备环节，这几类设备体量大，分别占据前道设备资本开支20%以上市场空间。国产设备正在匹配不同薄膜材料的刻蚀和沉积工艺，并在28nm/14nm的环节逐步提高覆盖度，在部分细分领域如EPI，ALD等领域逐步补全产品品类，长期替代空间依旧广阔
- 热处理设备，清洗设备，CMP设备等环节国产化覆盖率较高。热处理（各类炉管）及清洗设备国产化率在30-40%左右
- 光刻机核心设备是国产化能力相对薄弱的环节，目前上海微已有i-line/90nm光刻机，以及后道先进封装光刻机的出货，但我国28nm国产光刻机量产仍需时间。
- 展望2023，除去光刻机的国产化率正在从20%的节点上向上提升，24年随着存储线的拉通国产化率有望较23年加速提升



目录

1

总量与BOM表：算力新需求下的供应链机会概览

2

IC载板：PCB单机价值量中容易忽视的大赢家

3

存储：周期底部下，HBM带来需求端边际增量

4

设备：超市场悲观预期下的新工艺需求挖掘

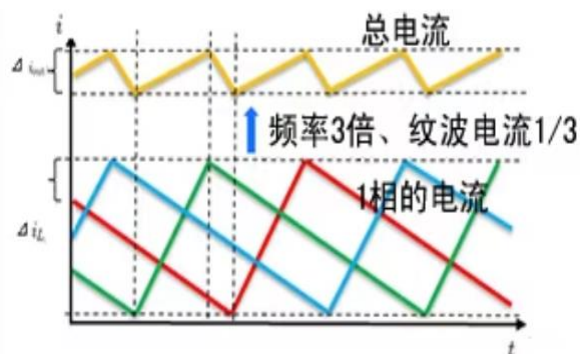
5

新电源解决方案机会：多相控制器/Drmos/芯片电感

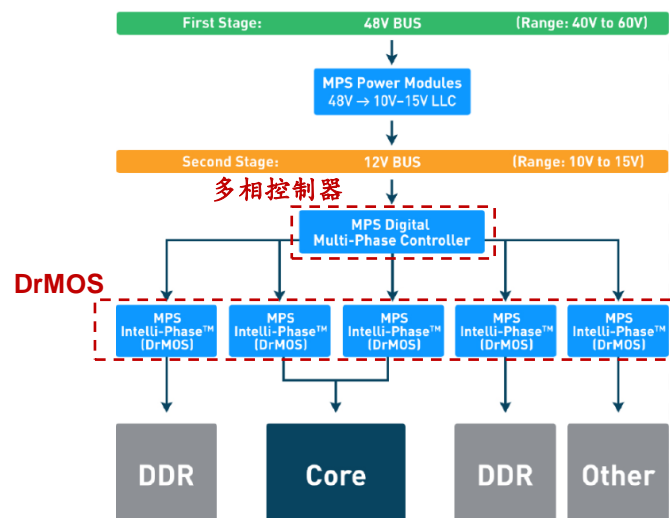
● 服务器用CPU/GPU的最佳电源解决方案驱动多相控制器+DrMOS市场增长

- 多相BUCK电源是服务器用CPU/GPU的最佳电源解决方案。随着大数据、云计算、人工智能概念的兴起，通信基站、数据中心等基础建设终端应用都需要功率更大的CPU\GPU来支持更为强劲的算力需求。而普通单相BUCK电源难以满足CPU/GPU大负荷瞬间的大电流需求，会有过大的应力导致器件发热严重、使用寿命缩短、效率降低。而多相BUCK电源通过并联的方式，能够同时提供m个频率相同而初相位互异的电压的电源。多个相位并联在一起共用输入和输出电容，可以有效减少输出和出入电容，在高负载电流下能提高热性能和效率，并且可以改善负载瞬态期间的过冲和下冲，因此被认为是这些应用场景的最佳解决方案
- DrMOS方案将驱动IC与功率MOSFET集成，是目前常用的多相电源供电方案。目前，常说的多相电源包含控制器和DrMOS，是一种多路交错并联的同步BUCK拓扑。DrMOS将功率MOSFET和驱动IC集成在一个芯片中，采用芯片封装工艺，再内部优化功率回路和驱动回路，可以最大程度减少寄生电感和电容影响。使用DrMOS方案可以将功率回路和控制回路完全分开互不干扰，设计简洁灵活，系统更为紧凑可靠，更有利于数字电源的设计

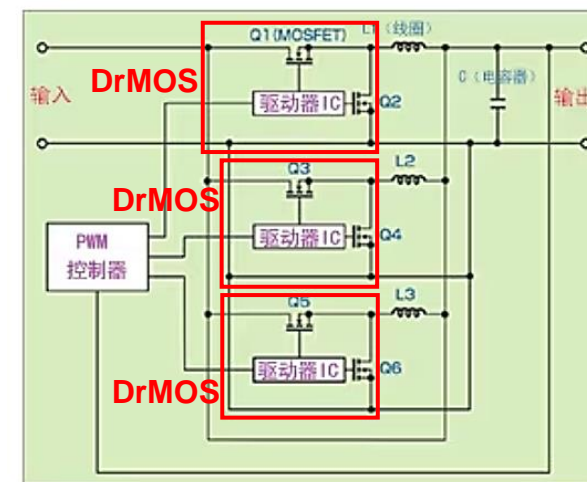
多相电路的各电流波形



48V数据中心应用中的芯源系统产品框图



DrMOS集成了驱动IC和功率MOSFET

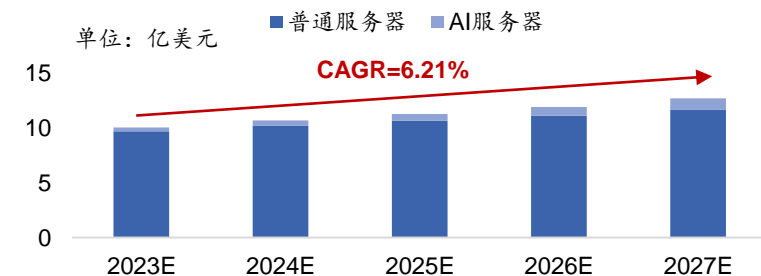


● 我们预计服务器用多相控制器+DrMOS市场规模将于27年达到12.32亿美元，2023~2027 CAGR约6.21%

单台普通服务器中CPU对应价值量	
①多相控制器	
单台服务器用CPU所需的多相控制器相数（相）	10
单台服务器用DDR所需的多相控制器相数（相）	10
单台服务器所需多相控制器的相数（相）	20
多相控制器单价（美元/相）	0.58
单台普通服务器对应的多相控制器价值量（美元）	11.6
②DrMOS	
单台普通服务器所对应的DrMOS电流（A）	1300
DrMOS单价（美元/10A）	0.43
单台普通服务器所对应的DrMOS价值量（美元）	55.9
单台普通服务器所对应的多相控制器+DrMOS价值量（美元）	67.5

- 根据产业链调研，目前普通服务器2颗CPU通常合计需配十相以上的多相控制器
- 根据英飞凌、MPS等厂商提供的解决方案，普通服务器中所需DDR通常合计需要十相左右的多相控制器
- 根据TI所公布的多相控制器价格，多相控制器相数越高，平均单相的价格越小，目前TI的12相控制器单相价格约0.58美元，故假设多相控制器的单相价格也为0.58美元/相
- 根据产业链调研，普通服务器CPU用DrMOS通常合计约1000A，DDR所需DrMOS安培数低于CPU，假设单台服务器所需DrMOS约1300A
- 根据TI所公布的DrMOS价格，单颗DrMOS安培越高，单价越高，但其每10A的单价越低，目前70A的DrMOS中每10A的DrMOS单价约0.43美元，故假设DrMOS每10A单价约0.43美元

服务器用多相控制器+DrMOS市场规模



单台AI服务器中GPU对应价值量	
单个GPU对应的多相控制器+DrMOS价值量	
①多相控制器	
一个GPU所需多相控制器数量（颗）	1
多相控制器单价（美元/颗）	9
一个GPU对应的多相控制器价值量（美元）	9
②DrMOS	
一个GPU所对应的DrMOS数量	16
DrMOS单价（美元/颗）	1.5
一个GPU所对应的DrMOS价值量（美元）	24
一个GPU所对应的多相控制器+DrMOS价值量（美元）	33
训练端	
单台AI服务器所需GPU数量	8
单台AI服务器所用GPU对应的多相控制器+DrMOS价值量（美元）	264
推理端	
单台AI服务器所需GPU数量	4
单台AI服务器所用GPU对应的多相控制器+DrMOS价值量（美元）	132

- 根据产业链调研，一颗GPU配一颗多相控制器
- V100需要1颗16相多相控制器，假设A100与V100类似，也需要一颗16相多相控制器。TI的8相控制器价格约7美元，12相价格约8美元，假设16相控制器单价约9美元
- 一颗16相多相控制器需要配16颗DrMOS
- 根据TI，一颗DrMOS单价约1.5美元
- 假设一台训练型AI服务器需要8颗GPU
- 假设一台推理型AI服务器需要4颗GPU

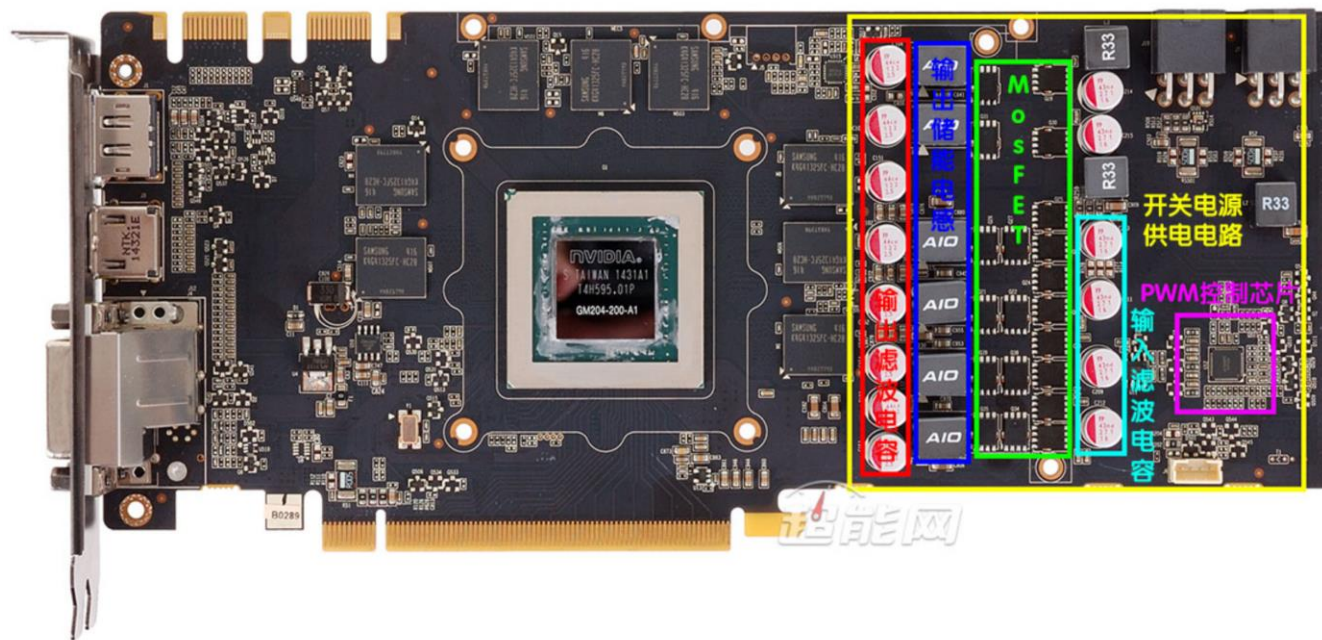
服务器用多相控制器+DrMOS市场规模					
	2023E	2024E	2025E	2026E	2027E
普通服务器					
普通服务器出货量（万台）	1380	1448	1518	1593	1671
多相控制器+DrMOS单价（美元/台）	67.5	67.5	67.5	67.5	67.5
普通服务器多相控制器+DrMOS市场规模（亿美元）	9.31	9.77	10.25	10.75	11.28
AI服务器					
①训练端					
训练端AI服务器出货量（万台）	0.52	0.69	0.87	1.04	1.21
多相控制器+DrMOS价值量（美元/台）	264	264	264	264	264
训练端AI服务器用多相控制器+DrMOS市场规模（亿美元）	0.014	0.018	0.023	0.027	0.032
②推理端					
推理端AI服务器出货量（万台）	26.71	34.72	45.14	58.68	76.28
多相控制器+DrMOS单价（美元/台）	132	132	132	132	132
推理端AI服务器用多相控制器+DrMOS市场规模（亿美元）	0.35	0.46	0.60	0.77	1.01
AI服务器用多相控制器+DrMOS市场规模（亿美元）	0.37	0.48	0.62	0.80	1.04
服务器用多相控制器+DrMOS市场规模（亿美元）	9.68	10.25	10.87	11.55	12.32



● 其他：AI服务器的GPU增量带动芯片电感用量增长

- 电感起到为芯片前端供电的作用，单台普通服务器所需芯片电感价值量约15美元。而AI服务器相较于普通服务器，增加了GPU用量，也带动了电感的需求量提升。以英伟达GM204芯片为例（用于PC），英伟达GPU至少需要6颗电感，假设AI服务器所用GPU所需的电感数量与GM204类似，则若AI服务器搭载4颗GPU，便需要增加至少24颗电感；若搭载8颗GPU，则需要增加至少48颗电感
- 根据产业链调研，芯片电感根据工艺复杂程度、功率大小、重量等因素，单价在0.43~1.45美元/颗左右。因此，若AI服务器搭载4颗GPU，单台所需电感价值量便至少增加10.32美元；若搭载8颗GPU，单台所需电感价值量便至少增加20.64美元

英伟达GM204的开关电源供电电路



资料来源：超能网，华安证券研究所



重要声明

分析师声明

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格，以勤勉的执业态度、专业审慎的研究方法，使用合法合规的信息，独立、客观地出具本报告，本报告所采用的数据和信息均来自市场公开信息，本人对这些信息的准确性或完整性不做任何保证，也不保证所包含的信息和建议不会发生任何变更。报告中的信息和意见仅供参考。本人过去不曾与、现在不与、未来也将不会因本报告中的具体推荐意见或观点而直接或间接接收任何形式的补偿，分析结论不受任何第三方的授意或影响，特此声明。

免责声明

华安证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。本报告中的信息均来源于合规渠道，华安证券研究所力求准确、可靠，但对这些信息的准确性及完整性均不做任何保证，据此投资，责任自负。本报告不构成个人投资建议，也没有考虑到个别客户特殊的投资目标、财务状况或需要。客户应考虑本报告中的任何意见或建议是否符合其特定状况。华安证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

本报告仅向特定客户传送，未经华安证券研究所书面授权，本研究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，或以任何侵犯本公司版权的其他方式使用。如欲引用或转载本文内容，务必联络华安证券研究所并获得许可，并需注明出处为华安证券研究所，且不得对本文进行有悖原意的引用和删改。如未经本公司授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。本公司并保留追究其法律责任的权利。

投资评级说明

以本报告发布之日起6个月内，证券（或行业指数）相对于同期沪深300指数的涨跌幅为标准，定义如下：

行业评级体系

增持：未来 6 个月的投资收益率领先沪深 300 指数 5%以上；

中性：未来 6 个月的投资收益率与沪深 300 指数的变动幅度相差-5%至 5%；

减持：未来 6 个月的投资收益率落后沪深 300 指数 5%以上；

公司评级体系

买入：未来6-12个月的投资收益率领先市场基准指数15%以上；

增持：未来6-12个月的投资收益率领先市场基准指数5%至15%；

中性：未来6-12个月的投资收益率与市场基准指数的变动幅度相差-5%至5%；

减持：未来6-12个月的投资收益率落后市场基准指数5%至15%；

卖出：未来6-12个月的投资收益率落后市场基准指数15%以上

无评级：因无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使无法给出明确的投资评级。 市场基准指数为沪深300指数。



谢谢！