

HBM: AI的内存瓶颈, 高壁垒高增速

分析师: 刘双锋

liushuangfeng@csc.com.cn

SAC 编号: S1440520070002

分析师: 章合坤

zhanghekun@csc.com.cn

SAC 编号: S1440522050001

分析师: 孙芳芳

sunfangfang@csc.com.cn

SAC 编号: S1440520060001

研究助理: 何昱灵

heyuling@csc.com.cn

发布日期: 2024年3月9日



中信建投证券
CHINA SECURITIES

并购家 免费行业报告网

IPOIPO.CN 每日更新 不用注册会员 无广告 永久免费

摘要

- HBM是限制当前算力卡性能的关键因素，海力士、三星、美光正加大研发投入和资本开支，大力扩产并快速迭代HBM，预计2024年HBM3e 24GB/36GB版本将量产/发布，内存性能进一步提高。HBM供需将持续紧俏，市场规模高速增长。通过分析生产工艺（TSV、键合等）和技术演进方向（先进制程、叠层），我们认为封装测试、前道和后道先进封装的设备和材料将是HBM主要受益方向。
- HBM是当前算力的内存瓶颈。存储性能是当下制约高性能计算的关键因素，从存储器到处理器，数据搬运会面临带宽和功耗的问题。为解决传统DRAM带宽较低的问题，本质上需要对单I/O数据速率和位宽进行提升。HBM由于采用了TSV、微凸块等技术，DRAM裸片、计算核心间实现了较短的信号传输路径、较高的I/O数据速率、高位宽和较低的I/O电压，因此具备高带宽、高存储密度、低功耗等优势。即便如此，当前HBM的性能仍然跟不上算力卡的需求。
- 三大原厂持续加大研发投入，HBM性能倍数级提升。随着技术的迭代，HBM的层数、容量、带宽指标不断升级，目前最先进的HBM3e版本，理论上可实现16层堆叠、64GB容量和1.2TB/s的带宽，分别为初代HBM的2倍、9.6倍和4倍。从Trendforce公布的HBM Roadmap来看，2024年上半年，海力士、三星、美光均会推出24GB容量的HBM3e，均为8层堆叠。2024年下半年，三家厂商将推出36GB版本的HBM3e，或为12层堆叠。此外，HBM4有望于2026年推出。
- HBM制造集成前道工艺与先进封装，TSV、EMC、键合工艺是关键。HBM制造的关键在于TSV DRAM，以及每层TSV DRAM之间的连接方式。目前主流的HBM制造工艺是TSV+Micro bumping+TCB，例如三星的TC-NCF工艺，而SK海力士则采用改进的MR-MUF工艺，在键合应力、散热性能、堆叠层数方面更有优势。目前的TCB工艺可支撑最多16层的HBM生产，随着HBM堆叠层数增加，以及HBM对速率、散热等性能要求的提升，HBM4开始可能引入混合键合工艺，对应的，TSV、GMC/LMC的要求也将提高。

摘要

- **AI刺激服务器存储容量扩充，HBM需求强劲。**随着云计算厂商将更多资本开支投入AI基础设施，AI服务器ODM对2024年展望乐观，预计2024年AI服务器出货量继续大幅增长。相较于一般服务器而言，AI服务器多增加GPGPU的使用，以NVIDIA A100/H100 80GB配置8张计算，HBM用量约为640GB，超越常规服务器的内存条容量，H200、B100等算力卡将搭载更高容量、更高速率HBM。我们测算，随着算力卡单卡HBM容量提升、算力卡出货量提升、技术迭代带来单GB HBM单价提升，2023年HBM市场规模为40亿美元，预计2024年增长至148亿美元，2026年增长至242亿美元，2023~2026年CAGR为82%。
- **目前HBM供应链以海外厂商为主，部分国内厂商打入了海外存储/HBM供应链。**国产HBM正处于0到1的突破期，HBM供应主要为韩系、美系厂商，国内能获得的HBM资源较少。随着国产算力卡需求快速增长，对于算力卡性能至关重要的HBM也有强烈的供应保障诉求和国产化诉求。建议关注：封测、设备、材料等环节。
- **相关标的：**
 - ✓ 封测：通富微电、长电科技、深科技；
 - ✓ 设备：中微公司、北方华创、拓荆科技、芯源微、赛腾股份、华海清科、精智达、新益昌；
 - ✓ 材料：雅克科技、联瑞新材、华海诚科、强力新材、天承科技、飞凯材料、壹石通、兴森科技；
 - ✓ 代理：香农芯创。
- **风险提示：**AI技术落地不及预期；国际环境变化影响；宏观环境的不利因素；市场竞争加剧风险。

一、HBM：算力的内存瓶颈

二、SK海力士HBM工艺分析：TSV、EMC、混合键合成趋势

三、市场测算：未来三年CAGR超80%

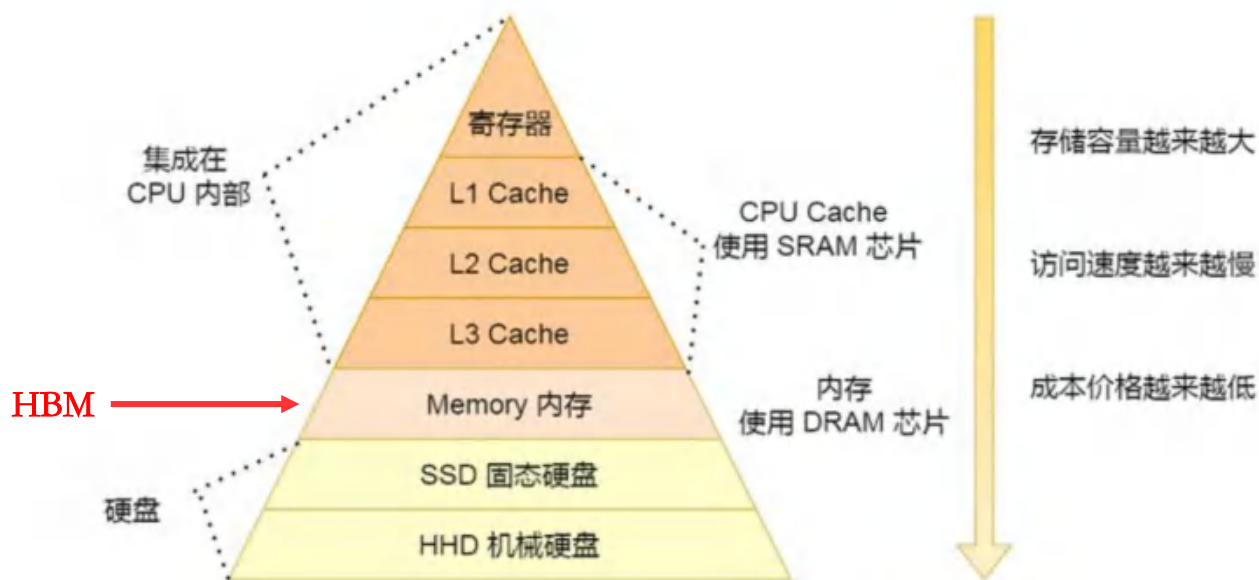
四、相关标的：设备、材料迎来新成长

五、风险提示

CPU拥有多级缓存架构，HBM属于内存环节

- 现代CPU为了提升执行效率，减少CPU与内存的交互(交互影响CPU效率)，一般在CPU上集成了**多级缓存架构**。
- CPU缓存即**高速缓冲存储器**，是位于CPU与主内存间的一种容量较小但速度很高的存储器。由于CPU的速度远高于主内存，CPU直接从内存中存取数据要等待一定时间周期，Cache中保存着CPU刚用过或循环使用的一部分数据，当CPU再次使用该部分数据时可从Cache中直接调用，减少CPU的等待时间，提高了系统的效率。
- 从CPU到用户数据，需经历“寄存器>SRAM>内存>SSD或HDD”的路径，HBM属于内存的一种。

图表：CPU的多级缓存架构



存储性能是当下制约高性能计算的关键因素

- 内存墙（传输带宽慢或容量有限）是算力提升的重要瓶颈。
- 从存储器到处理器，数据搬运会面临2个问题：（1）数据搬运慢；（2）搬运能耗大。

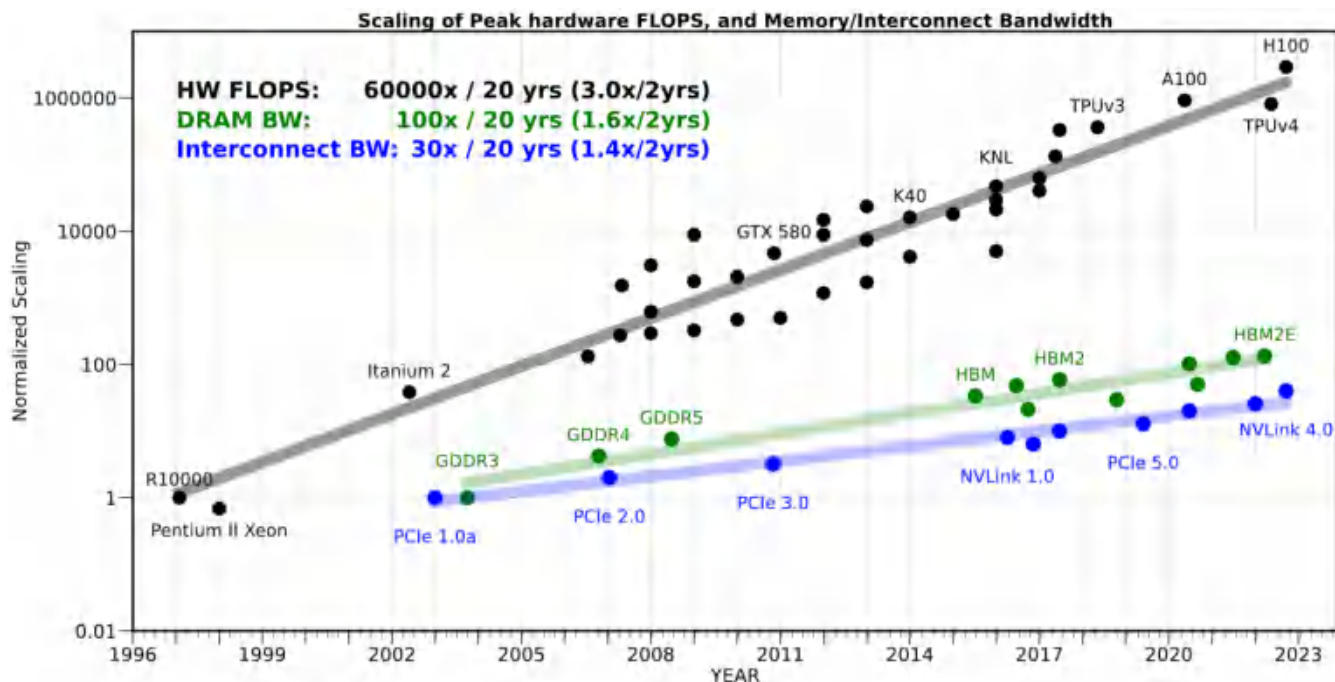
图表：数据传输的速率及能耗限制算力性能发挥



存储性能是当下制约高性能计算的关键因素

- 算力的增速远大于存储的增长速度。目前绝大部分的计算系统基于冯诺依曼计算机体系，而该体系下存储与计算单元分离，需要通过总线不断在存储与处理器之间传输数据，因此存储的带宽制约了算力的利用效率。此外，从外部处理器到内存之间不断进行数据的搬运，搬运时间往往是运算时间的成百上千倍，产生的无用能耗过多。

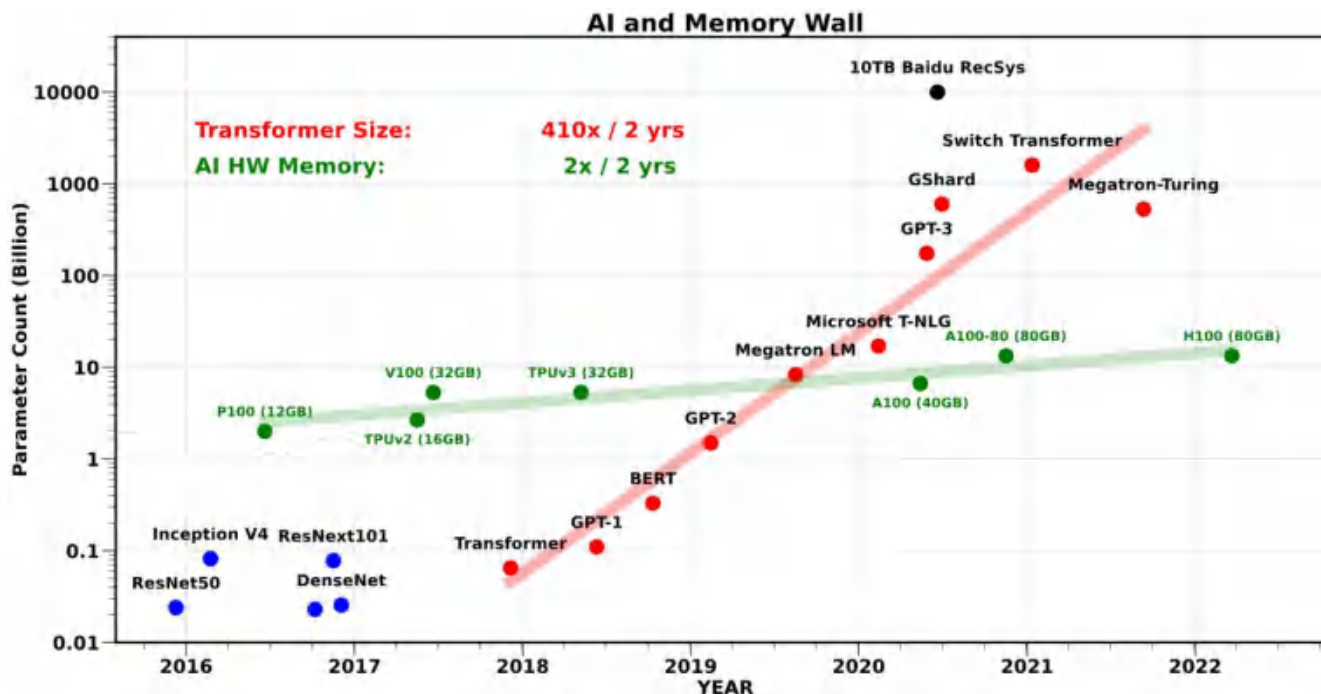
图表：算力的增速远超存储与互连带宽的增速



存储性能是当下制约高性能计算的关键因素

- 模型体量的增速远大于算力卡存储容量的增速。随着 Transformer 模型的大规模发展和应用，模型大小每两年平均增长了240倍，而单个GPU内存容量仅以2年2倍的速度扩大。为了摆脱单一算力芯片内存有限的问题，可以将模型部署于多颗GPU上运行，但在算力芯片之间移动数据，仍然比单一芯片内部移动数据低效，因此算力芯片内存容量的缓慢增速制约了更大规模的模型应用。

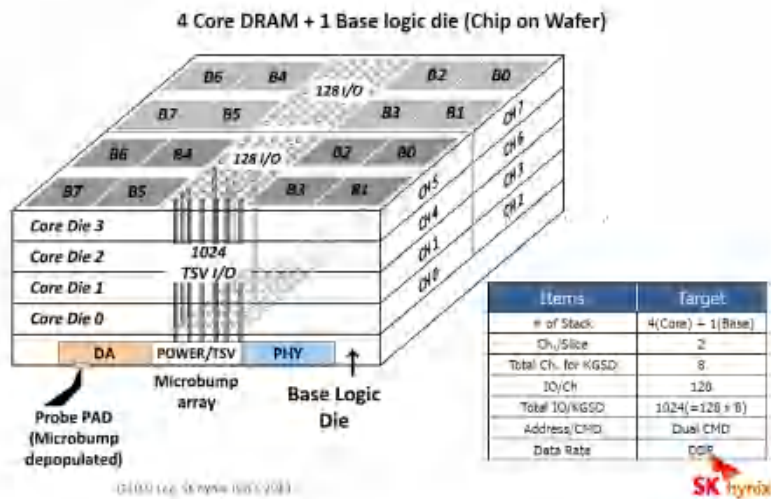
图表：大模型体积的增速远高于算力芯片存储容量的增速



HBM：基于TSV技术获得的高带宽内存，已成为高性能计算的首选

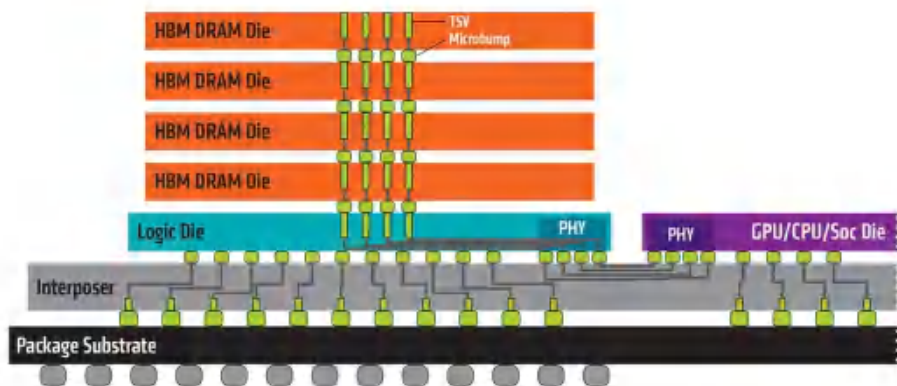
- 如同闪存从2D NAND向3D NAND发展，DRAM也正在从2D向3D技术发展，HBM（High Bandwidth Memory，高带宽存储）为主要代表产品。从结构上看其特点有：（1）3D堆叠结构并由TSV互连：HBM由多颗DRAM die堆叠成3D结构，使用TSV技术实现信号的共享与分配；（2）高I/O数量带来高位宽：HBM的每颗DRAM Die包含多个通道，可独立访问。每个通道又包含多个I/O口，位宽64/128bit，使HBM的总位宽高达1024bit。
- HBM性能优异，主要用于高性能计算芯片。HBM的结构特点为其带来存储密度更大、功耗更低、带宽更高的优势。HBM通常采用CoWoS等先进封装技术与计算核心进行互连，多用于与数据中心算力芯片GPU/FPGA/ASIC等配合工作。

图表：HBM结构示意图



数据来源：SK Hynix，中信建投

图表：HBM采用CoWoS封装与计算核心互连

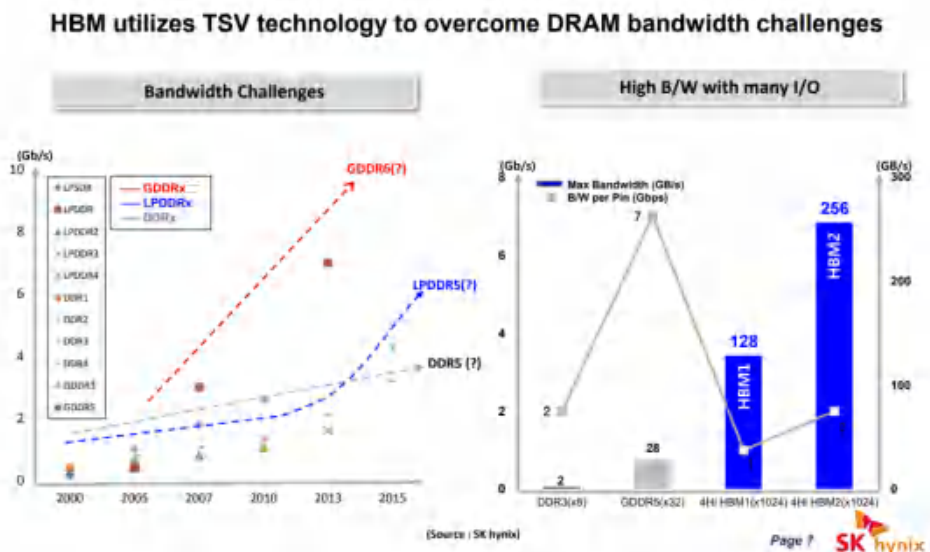


数据来源：AMD，中信建投

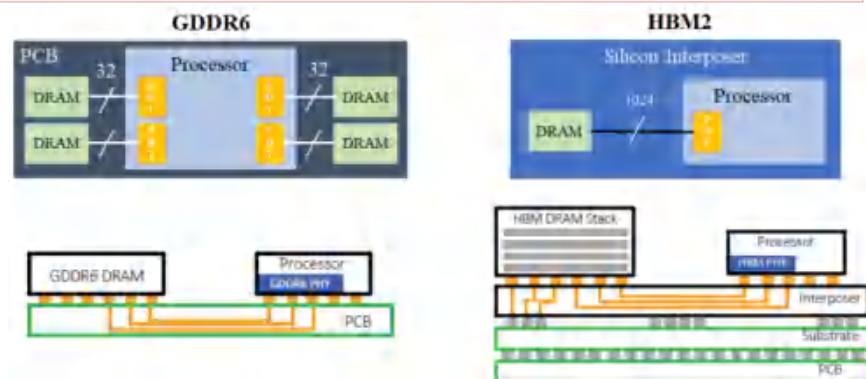
HBM优势：相比GDDR，大幅增加计算核心可用的带宽

- 总带宽=I/O数据速率 (Gb/s) *位宽/8。为解决DDR带宽较低的问题，本质上需要对单I/O的数据速率和位宽 (I/O数*单I/O位宽) 进行提升，着重于不同方向发力的方案分别为GDDR和HBM。
- GDDR提升单I/O的数据速率，总带宽提升有限：GDDR采取大幅提升单I/O数据速率的手段来改善总带宽，GDDR5和GDDR6的单I/O数据速率已达到7 Gb/s到16Gb/s，超过HBM3的6.4 Gb/s。虽然GDDR的位宽相比DDR也有提升，但由于GDDR仍然是通过PCB与计算核心进行互连，总位宽的提升受到限制。
- HBM利用TSV技术提升I/O数，总带宽提升明显：HBM利用TSV技术在维持较低的单I/O数据速率的情况下，大幅提升了位宽进而获得了远优于GDDR的总带宽表现。

图表：HBM利用高位宽克服带宽的限制



图表：决定GDDR、HBM带宽差异的核心因素对比

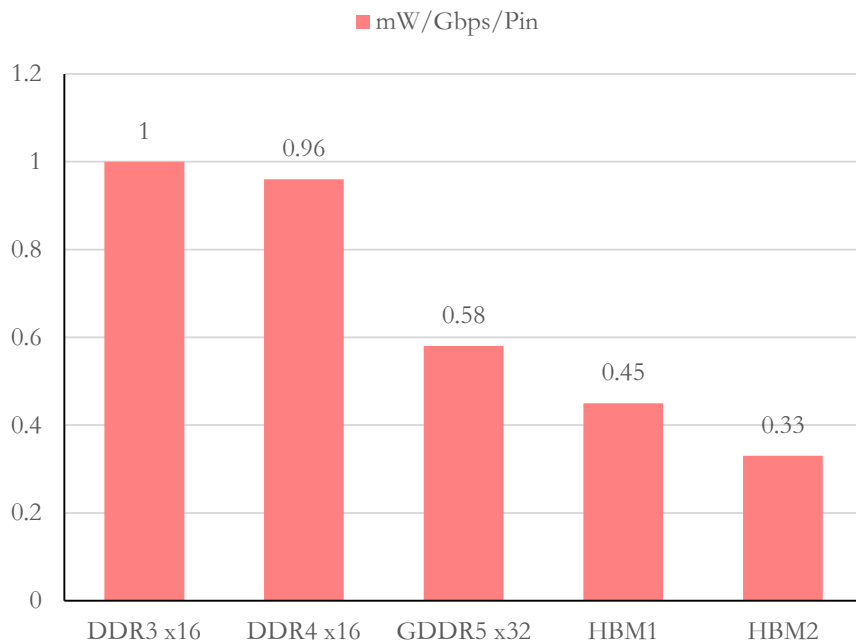


最新标准	GDDR6	HBM2e	HBM3
最高I/O速率(Gb/s)	16	3.6	6.4
每颗DRAM通道数	2	8	16
单通道位宽 (bit)	16	128	64
最高位宽 (bit)	32	1024	1024
最高带宽 (GB/s)	64	460	819

HBM优势：相比GDDR，功耗、集成度更具优势

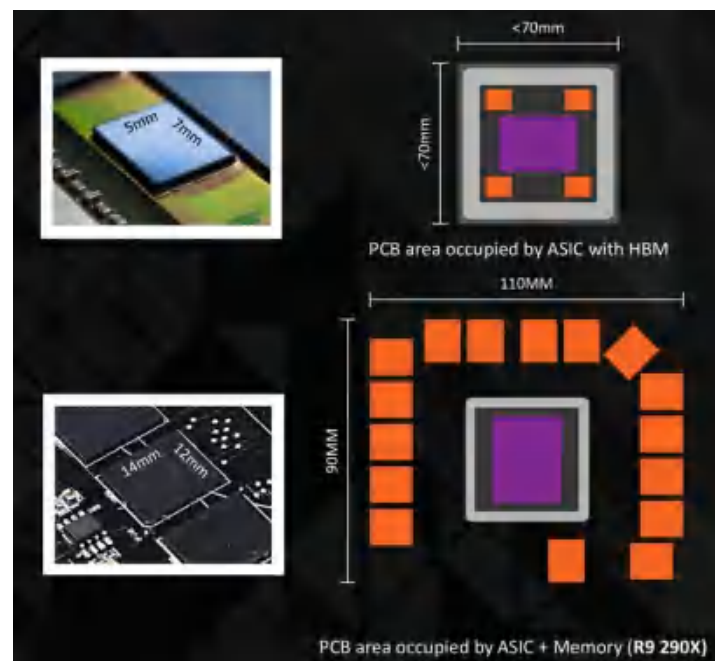
- HBM的高I/O速率和低电压使其获得优异的能效比。由于采用了TSV、微凸块等技术，DRAM裸片、计算核心间实现了较短的信号传输路径、较高的数据速率和较低的I/O电压，使HBM的能效比优于DDR和GDDR。
- HBM采用DRAM堆叠获得高存储密度，有利于提高集成度。相比于1GB GDDR5（4颗256M颗粒），1GB HBM1的表面积节省达到了94%。对比AMD使用4GB HBM1的GPU R9 Fury X与使用4GB GDDR5的R9 290X，HBM方案的面积可以控制在4900 mm²以内，而使用GDDR5方案的面积达到9900 mm²。

图表：HBM能效比优于GDDR与DDR



数据来源：Sk hynix, 中信建投

图表：HBM有效减少芯片的面积占用



数据来源：AMD, 中信建投

HBM标准历经多次升级，I/O速率、带宽获得明显提升

- **三大DRAM原厂推动技术迭代，性能表现持续提升。**HBM高带宽的特点使其适合数据中心GPU等高性能算力芯片，全球三大DRAM原厂陆续加强HBM的技术投入。随着技术的迭代，HBM的层数、容量、带宽指标不断升级，目前最先进的HBM3e版本，理论上可实现16层堆叠、64GB容量和1.2TB/s的带宽，分别为初代HBM的2倍、9.6倍和4倍。
- **容量（GB）：**由单颗DRAM颗粒的容量、DRAM颗粒的堆叠层数共同决定。
- **总带宽（GB/s）：**从HBM1到HBM3e，尽管不同代际之间的I/O数量和单I/O位宽存在差异，但总位宽均保持为1024bit，因此I/O速率的提升推动了总带宽提升。根据DigiTimes援引Seoul Economy的消息，HBM4有望将总位宽提升至2048bit，即使I/O速率维持不变，总带宽也将有翻倍增长。

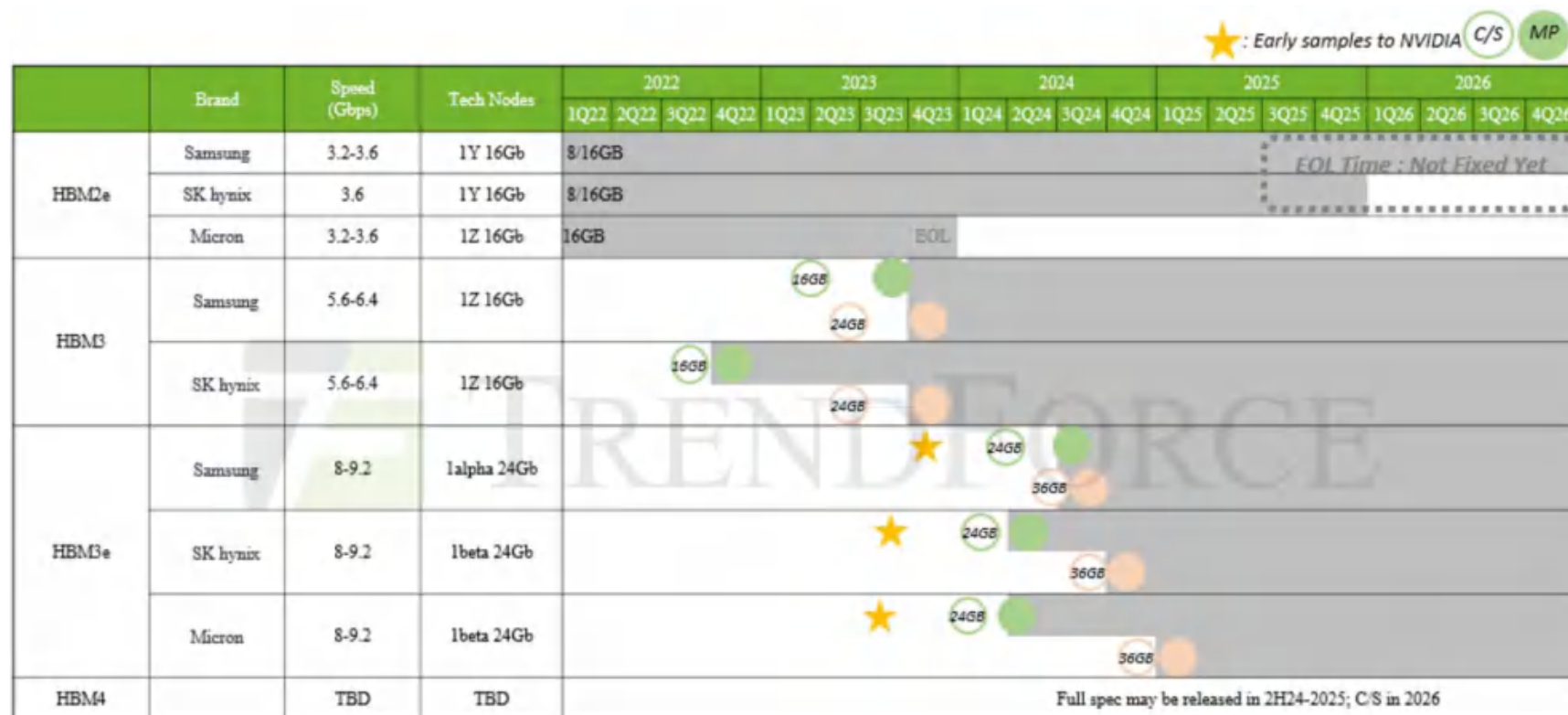
图表：各代HBM参数对比

代际	制程 (nm)	最大I/O 速率 (Gb/s)	通道数	单通道位 宽 (bit)	总位宽 (bit)	最大带宽 (GB/s)	可堆叠高 度	DRAM颗 粒最大容 量 (Gb)	HBM最 大容量 (GB)	Sk Hynix	Samsung	Micron
HBM	20nm	1	8	128	1024	128	8	16	16	√		
HBM2	20nm	2	8	128	1024	256	8	16	16	√	√	√
HBM2e	1y/1z 16Gb	3.6	8	128	1024	461	12	24	36	√	√	
HBM3	1z 16Gb	6.4	16	64	1024	819	16	32	64	√	√	
HBM3e	1α/1β 24Gb	9.6	16	64	1024	1229	16	32	64	√	√	√

2024年HBM3e 24GB版本和36GB版本将量产

- 自2023年ChatGPT发布以来，AI服务器的强劲需求正推动HBM快速迭代。从Trendforce公布的HBM Roadmap来看，2024年上半年，海力士、三星、美光均会推出24GB容量的HBM3e，按照1 alpha/beta 24Gb的工艺计算，均为8层堆叠。2024年下半年，三家厂商将推出36GB版本的HBM3e，或为12层堆叠。

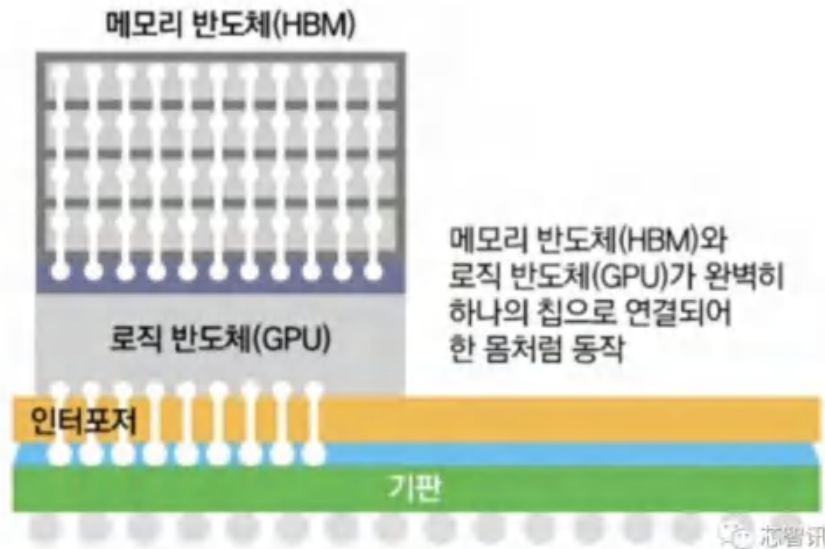
图表：各家HBM Roadmap



HBM4展望：除配置升级外，还可能朝客制化发展

- 当前供应商聚焦于迭代HBM3e，单die容量和堆叠层数仍有提升空间，但HBM4已经列入研发计划：
- 1、**堆叠层数**：HBM4在堆叠层数上，除了现有的8/12层外，将再往16层发展，更高层数也将带动新堆叠方式Hybrid bonding的需求。HBM4 12层将于2026年推出，而16层则预计于2027年问世。
- 2、**逻辑die**：受到GPU/HBM规格更往高速发展带动，Trendforce预计HBM最底层的Logic die（Base die）将采用12nm制程，该部分将由晶圆代工厂提供，使得单颗HBM需要结合晶圆代工厂与存储器厂的合作。
- 3、**客制化**：根据韩国中央日报报道，SK海力士正在招募CPU、GPU等逻辑芯片的设计人员，目标是将未来的HBM4以3D堆叠的形式堆叠在英伟达、AMD等公司的逻辑芯片上，预计该HBM4内存堆栈将采用2048位接口。该方案未定，只是在讨论中。

图表：HBM4概念图

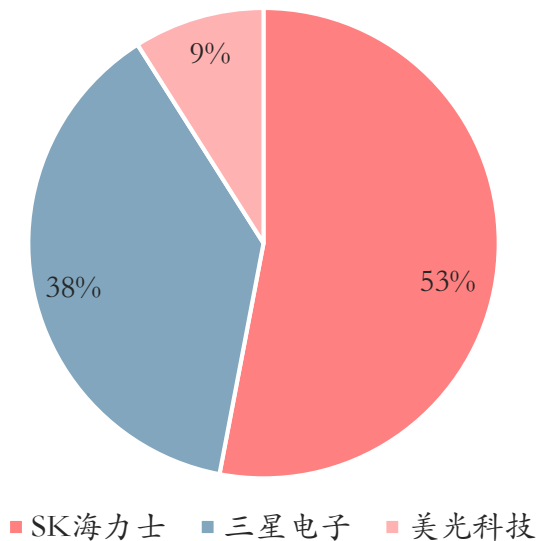


数据来源：芯智讯，韩国中央日报，中信建投

HBM市场寡头垄断，SK海力士占过半份额

- HBM市场格局集中，SK海力士占有主导地位。根据TrendForce，2022年三大原厂HBM市占率分别为SK海力士53%、三星电子约38%、美光约9%。
- HBM市场竞争白热化。2023年市场主要HBM代际是HBM2、HBM2e和HBM3，算力卡性能提升刺激HBM产品更迭，2023年下半年伴随NVIDIA H100与AMD MI300的搭载，HBM3渗透率提升。2024年伊始，SK海力士完成HBM3e开发，并送样英伟达测试，有望于上半年量产。预计三星电子和美光科技即将送样HBM3e，也有望于上半年量产，其中美光科技跳过了HBM3，直接研发HBM3e。三家原厂在HBM领域的竞争日趋白热化。

图表：2022年HBM市场竞争格局

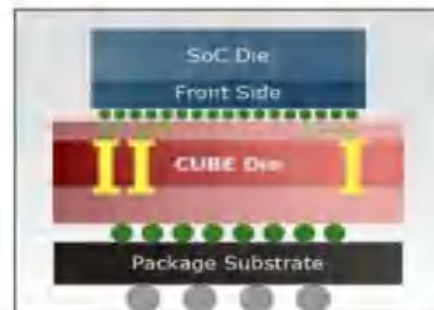
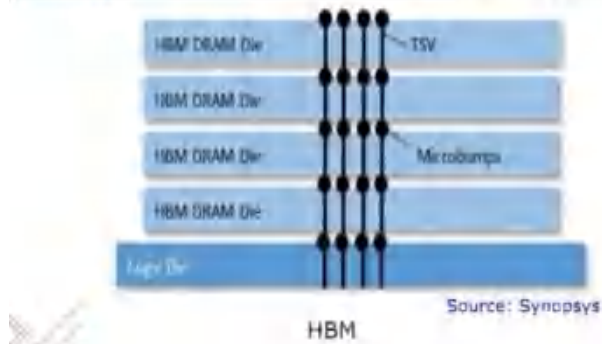


更多DRAM厂商正切入HBM赛道，国产HBM有望突破

- 二线、三线DRAM厂商也正在切入HBM赛道。华邦电于2023年8月介绍了其类HBM高带宽产品CUBEx，采用1~4层TSV DRAM堆叠，I/O速度500M~2Gbps，总带宽最高可达1024GB/s，颗粒容量为0.5~4GB，功耗低至不足1pJ/bit。这种比常规HBM拥有更高带宽的CUBEx可用于AR、VR、可穿戴等领域。
- 国产DRAM厂商有望突破HBM。目前一线厂商DRAM制程在1alpha、1beta水平，国产DRAM制程在25~17nm水平，中国台湾DRAM制程在25~19nm水平，国内DRAM制程接近海外。且国内拥有先进封装技术资源和GPU客户资源，有强烈的国产化诉求，未来国产DRAM厂商有望突破HBM。

图表：华邦电的类HBM高带宽产品CUBEx

Item	HBM1	HBM2	HBM2E	HBM3	CUBEx
Stack height	4Hi	8Hi	8Hi/12Hi	8Hi/12Hi/16Hi	1 - 4Hi
I/O Transfer rates	1 Gbps	2.4 Gbps	3.6 Gbps	6.4 Gbps	500M-2 Gbps
Overall package bandwidth	128 GB/s	256 GB/s	460 GB/s	819 GB/s	Up to 1024 GB/s
Package capacity	4GB	8GB	16GB/24GB	16GB/24GB/32GB	0.5 - 4GB
I/O number(R+W)	1024	1024	1024	1024	256-1024
Power / bit	6pJ/ bit	<5pJ/ bit	<5pJ/ bit	<4pJ/ bit	<1pJ/bit
Application	AI / Machine Learning in data center, HPC, accelerator system				SoC L4, AI NMC, AR/VR,...



一、HBM：算力的内存瓶颈

二、SK海力士HBM工艺分析：TSV、EMC、混合键合成趋势

三、市场测算：未来三年CAGR超80%

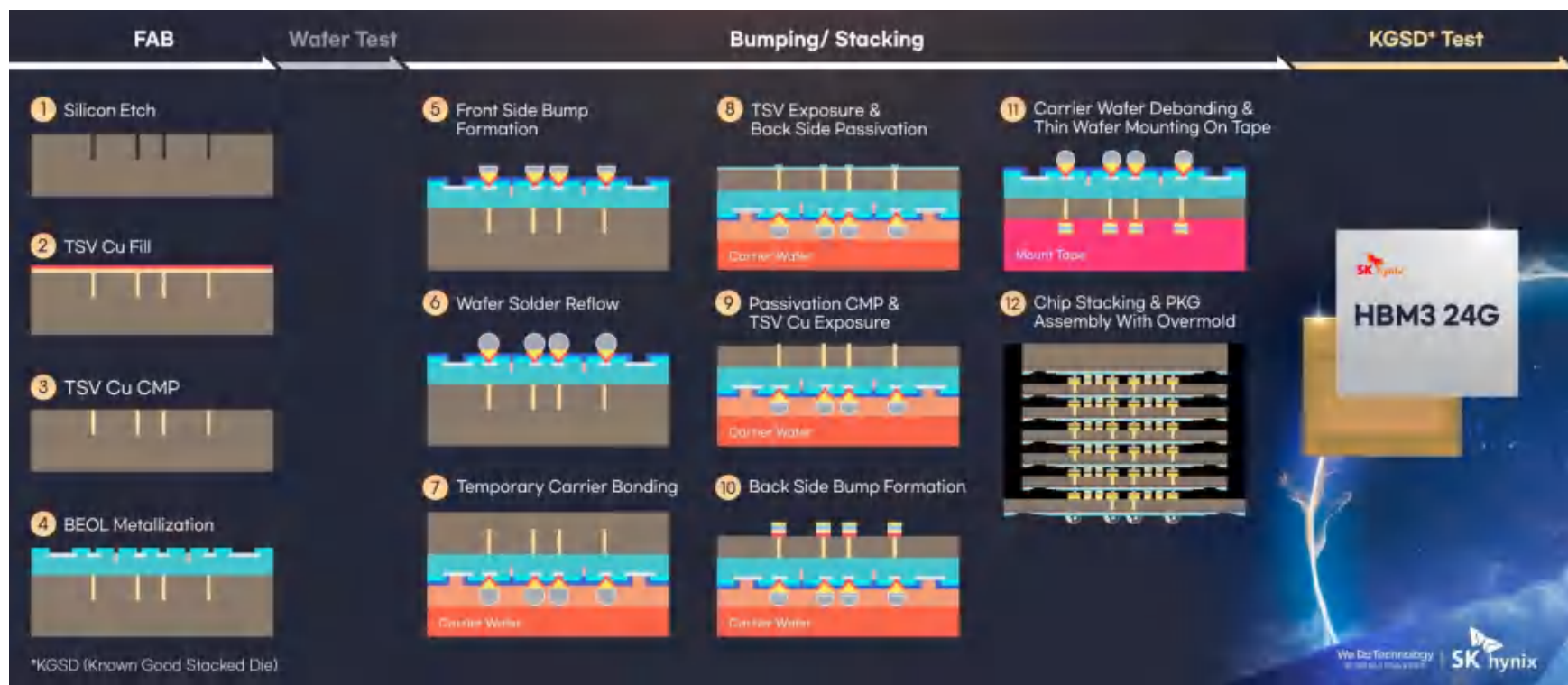
四、相关标的：设备、材料迎来新成长

五、风险提示

SK海力士HBM工艺分析

- SK海力士在HBM领域的成功归因于：布局较早、绑定大客户英伟达、先进的封装工艺。
- SK海力士的HBM采用定制的DRAM颗粒（TSV DRAM），核心工艺是MR-MUF工艺和EMC材料。

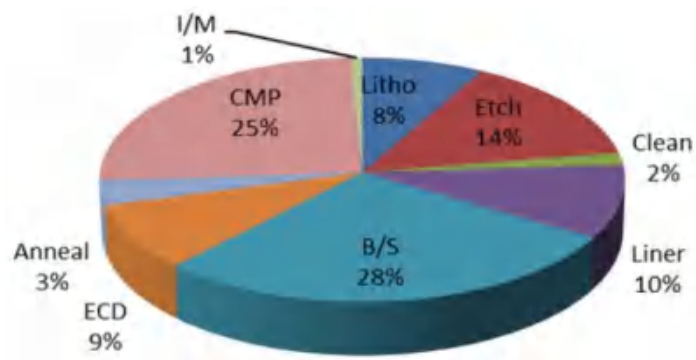
图表：SK海力士HBM生产工艺全流程



TSV DRAM

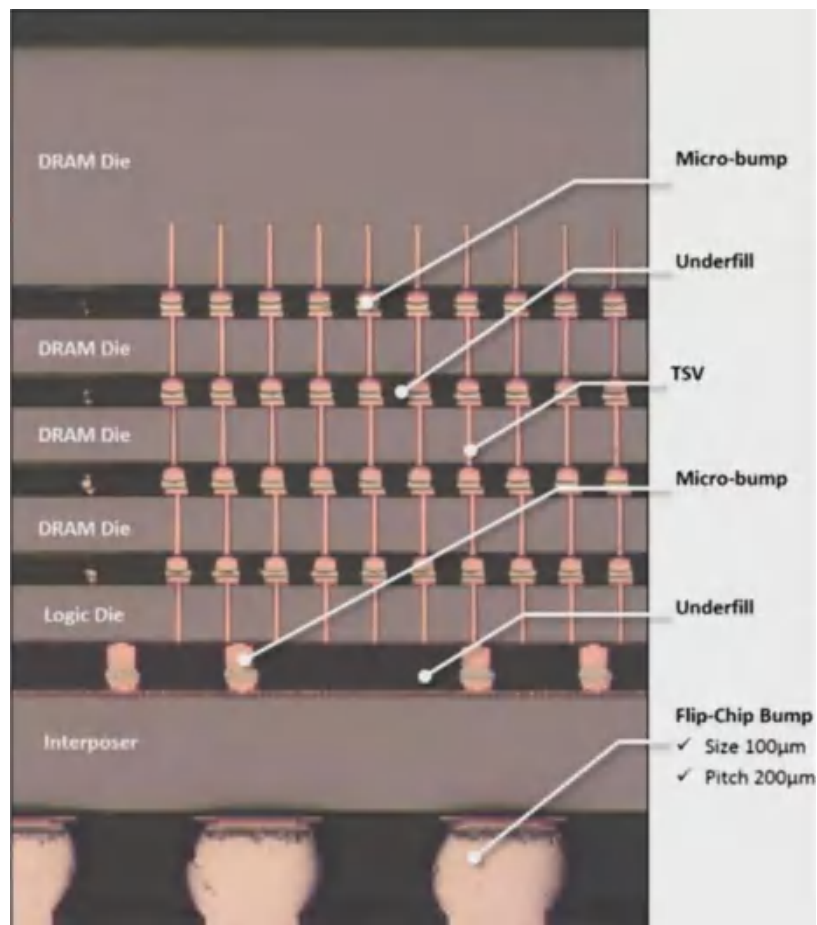
- **定制DRAM颗粒：**用于生产HBM的DRAM颗粒需要进行TSV打孔以便堆叠，因而需要在生产DRAM颗粒时，预留不含存储电路的位置用于TSV打孔。电路布局涉及存储厂商的商业机密，一般的封测厂商难以获得，也就难以参与HBM后道的生产，因此存储IDM在前道后道工艺磨合方面具备先天优势。
- TSV属于前道工艺，由存储厂商完成，主要用到刻蚀设备、薄膜沉积设备、电化学沉积设备ECD、CMP、清洗设备、退火设备、减薄机、划片机等，相关供应商包括：AMAT、TEL、Lam Research等。

图表：TSV的成本结构



数据来源：《Cost Analysis of TSV Process and Scaling Options》，
中信建投。B/S指Barrier and seed

图表：HBM中的TSV结构

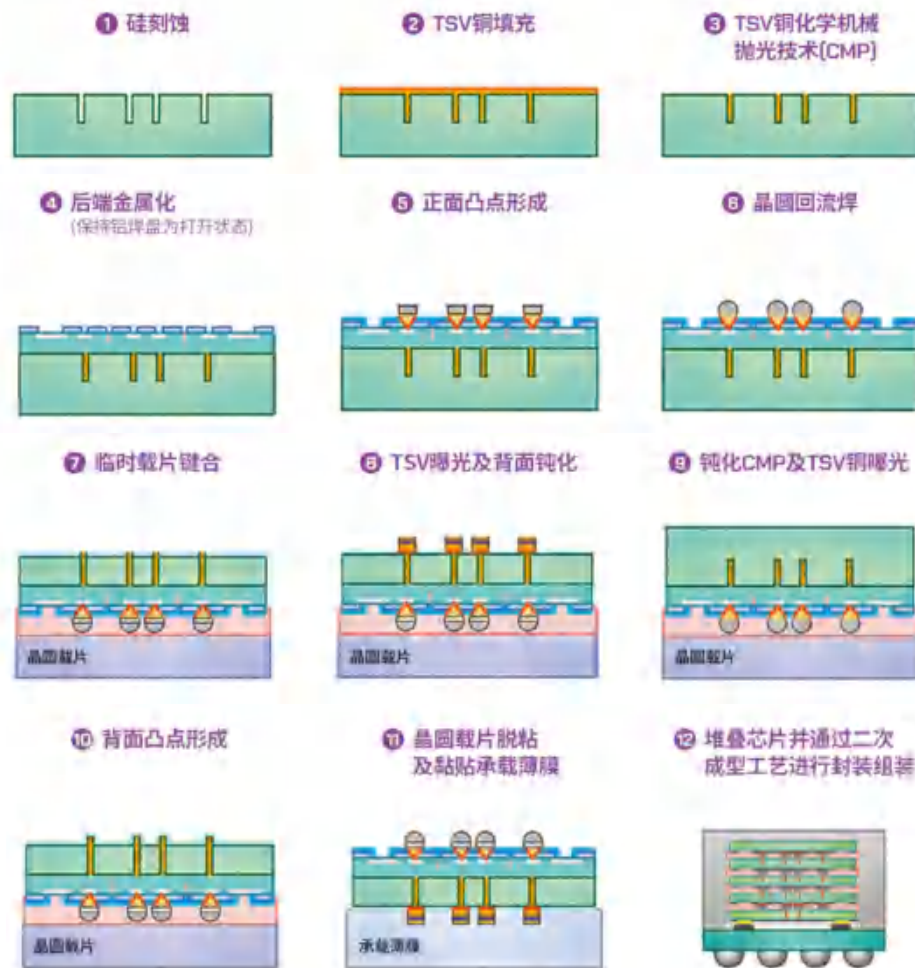


数据来源：半导体行业观察，中信建投

TSV DRAM

- TSV是目前唯一的垂直电互连技术，是实现2.5D/3D先进封装的关键技术之一。
- 1、**高密度集成**：通过先进封装，可以大幅度地提高电子元器件集成度，减小封装的几何尺寸，和封装重量。
- 2、**提高电性能**：由于TSV技术可以大幅度地缩短电互连的长度，从而可以很好地解决出现在SOC技术中的信号延迟等问题，提高电性能。
- 3、**多种功能集成**：通过TSV互连的方式，可以把不同的功能芯片(如射频、内存、逻辑、数字和MEMS等)集成在一起实现电子元器件的多功能。
- 4、**降低制造成本**：TSV三维集成技术虽然目前在工艺上的成本较高，但是可以在元器件的总体水平上降低制造成本。

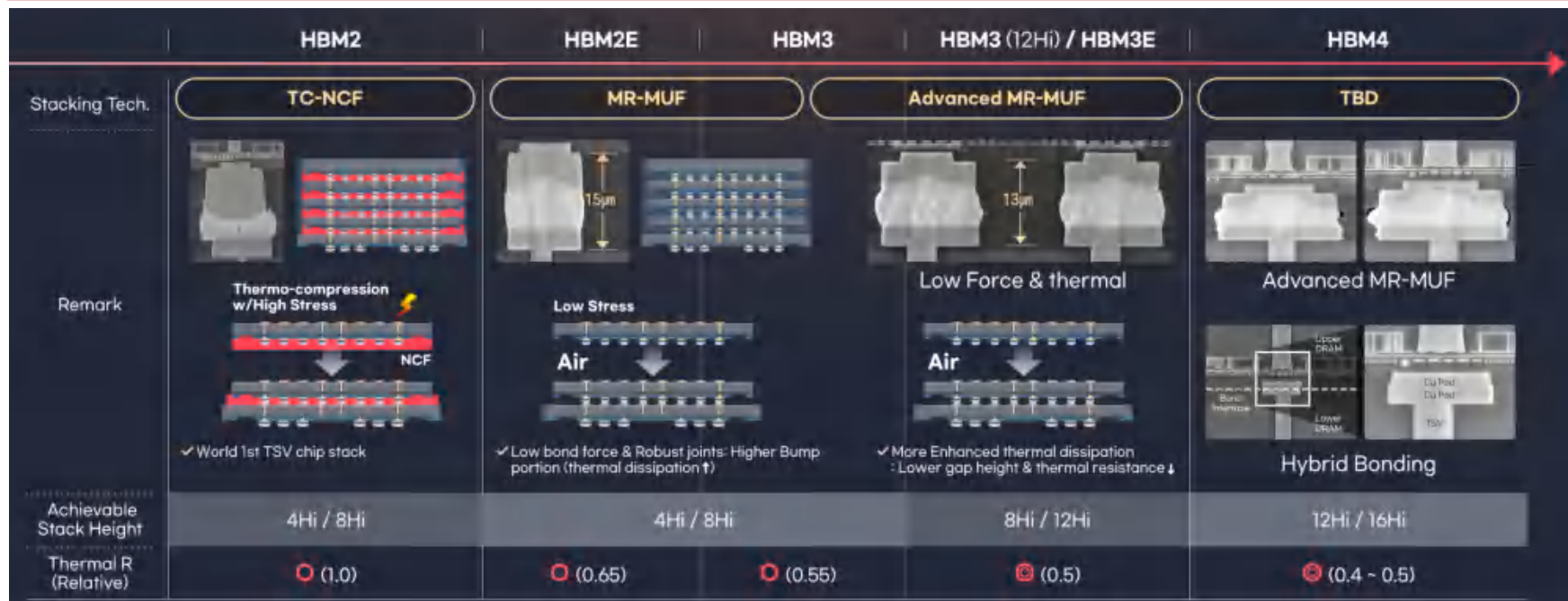
图表：TSV的工艺流程



MR-MUF工艺

- SK海力士HBM领先三星电子和美光的关键在于MR-MUF工艺。HBM制造的关键在于每层TSV DRAM之间的连接方式，三星电子从HBM生产之初就一直采用热压缩非导电薄膜（TC-NCF）工艺，而SK海力士采用的则是批量回流模制底部填充（MR-MUF）工艺。
- 技术路线图显示，从HBM2e开始，海力士放弃了TC-NCF工艺，改用MR-MUF工艺，实现了更低的键合应力和更优的散热性能。海力士HBM3e将采用改进的MR-MUF工艺，降低键合应力，提升散热性能，增加堆叠层数。HBM4有可能采用混合键合（Hybrid bonding）工艺，进一步降低TSV DRAM层与层之间的间隙，实现更高层数堆叠。

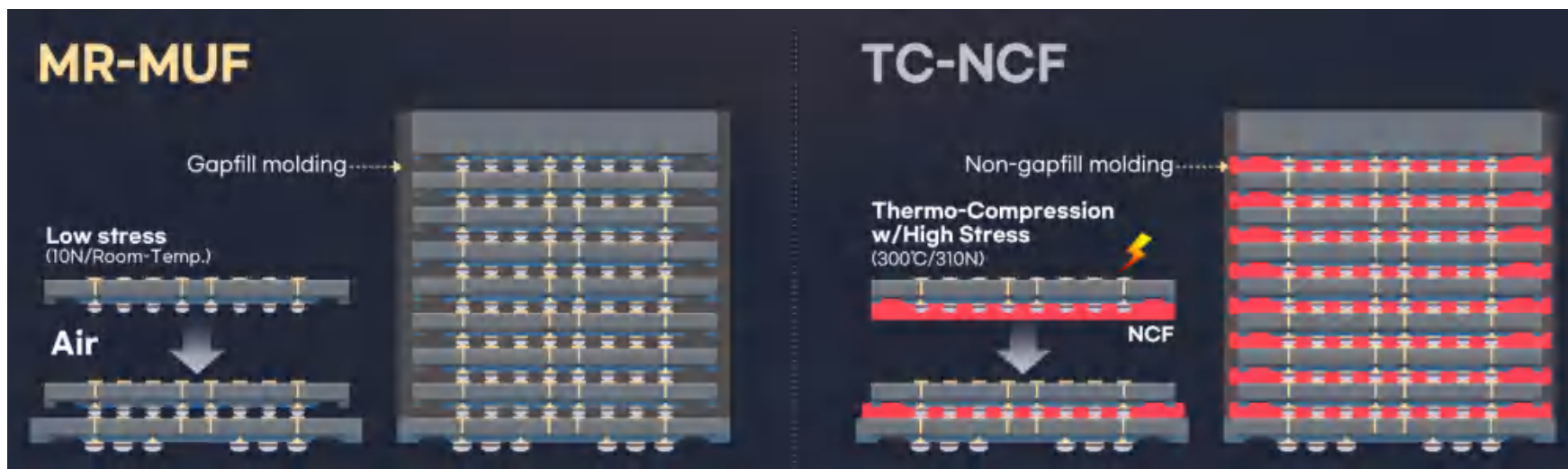
图表：SK海力士堆叠技术路线图



MR-MUF工艺

- MR-MUF全称为批量回流模制底部填充（Mass reflow molded underfill），是将多个芯片放置在下层基板上，通过回流焊一次性粘合，然后用模塑料填充芯片之间或芯片与基板之间间隙的方法，该方法主要用于FC封装和TSV芯片堆叠。
- 工艺流程：（1）在DRAM层与层之间，微凸点接触助焊剂，并进行横向引线键合层叠；（2）一次熔融全部的微凸点和助焊剂，并室温下施加压力，微凸点冷后即完成芯片和线路连接；（3）选用环氧树脂模塑料（EMC）填充芯片之间或基材之间的缝隙，同步进行绝缘和成型。该加工工艺的核心是处理层叠芯片中产生的热胀冷缩难题，及其处理芯片中间部位的空隙填充。
- 相比于TC-NCF，（1）MR-MUF不依靠高温和高压，减少了芯片键合产生的应力；（2）环氧树脂模塑料的导热系数是非导电薄膜的2~4倍，散热效率更高。

图表：海力士MR-MUF工艺和三星TC-NCF工艺比较

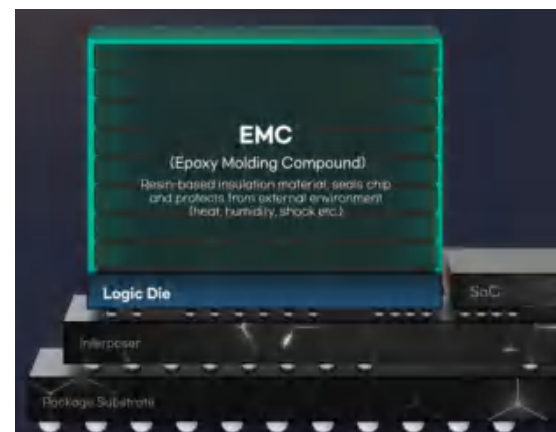
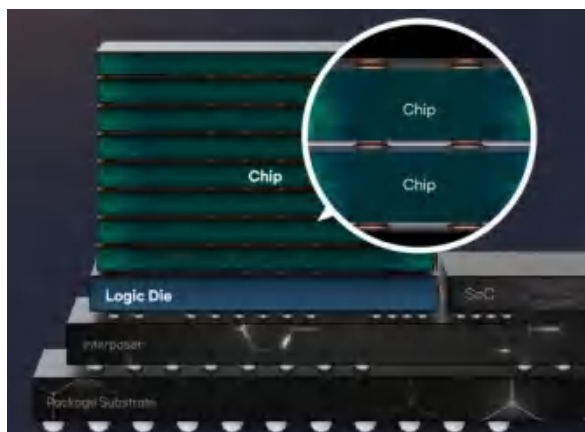


填充料EMC

- MR-MUF工艺的关键是：（1）晶圆翘曲度的控制（Chip warpage control）；（2）填充料的选取（Gapfill-MUF material）。
- 晶圆翘曲度：受加工条件（温度、压力等）影响，用量测设备检查，供应商有Camtek、ONTO。
- 填充料：EMC是先进封装常用的填充料，传统DDR或者Flash用的EMC一般是粉末状或圆柱体，而HBM中Die与Die之间的间距不足20微米（HBM2e 15微米，HBM3e 13微米），且有微凸点阻挡，粉末或者六面体流不进去。目前HBM主要使用GMC（Granular Molding Compound）或者LMC（Liquid Molding Compound）。目前海力士的GMC主要由日本Namics供应。

图表：MR-MUF工艺的微凸点键合和EMC填充流程

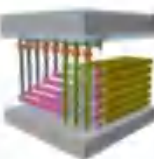
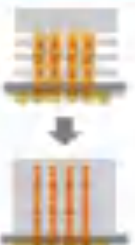
- 芯片堆叠，添加助焊剂
- 室温加压，完成芯片、线路连接
- EMC填充间隙



混合键合Hybrid Bonding

- 混合键合是将两片以上不相同的晶圆通过金属互连的工艺，不使用芯片堆叠之间的微凸点（球形），而是用铜与铜（片状，copper pad）互连，极大提升了芯片之间的信号传输速率。混合键合并非HBM专用的技术，CIS、3D NAND、逻辑芯片等均已大规模使用混合键合工艺。
- 技术优势：（1）允许无焊料键合，减少键合层厚度、缩短电气路径，并降低电阻；（2）通过直接将铜与铜接合，显著减小凸块间距。目前使用焊料很难实现10微米或更小的间距，而铜对铜直接键合可以将间距减小到不足1微米，提高芯片设计的灵活性；（3）拥有绝佳的散热性能；（4）极薄的粘合层和小间距会影响封装的形状因数，可以大大减小封装尺寸。

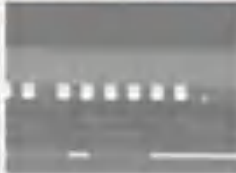
图表：混合键合工艺在各类芯片中的使用

Application	CIS	NAND	DRAM	Logic		Advanced Package	
	BSI	3D NAND	3D DRAM	Backside PDN	Sequential CFET	HBM	Disaggregation / Chiplets
Stacking Device	Sensor + Memory + Logic	Cell + Peripheral	Cell + Peripheral	Logic + Bare Si	Logic + Logic	DRAM + DRAM + Logic	
Bonding	W-W Cu Hybrid	W-W Cu Hybrid	W-W Cu Hybrid	W-W Ox Fusion	W-W Ox Fusion	D-W u Bump → Cu Hybrid	D-W / D-D Cu Hybrid
3D I/O Pitch	3um →1um	1um →0.5um	1um →0.5um	Sub um (nTSV)	Sub um (nTSV)	40um →25um	10um →1um
Structure			 Source: TEL 3D DRAM structure				 Chip partition (Chiplet) Chip Stacking Source: TEL
Status	HVM	R&D~HVM	R&D	R&D	R&D	R&D	R&D~HVM

混合键合Hybrid Bonding

- 传统DRAM使用wire bonding、Flip Chip等工艺进行电气互连，I/O数量为4/8/16个，触点密度不足200，互连长度50-2000微米，可堆叠至多8层；而TSV和Hybrid的I/O数量多达1024个，触点密度可达3.5~10k，互连长度低至5~20微米，堆叠层数可达16层甚至更高。因此，用Hybrid和TSV工艺生产的DRAM拥有高带宽、高存储密度，SK海力士预计HBM4开始使用Hybrid bonding，三星电子、美光也可能转向该技术。
- 混合键合设备的供应商主要有Besi、EV Group，其中Besi擅长Die to Wafer bonding，EV Group擅长Wafer to Wafer bonding。

图表：DRAM的不同互连工艺

Interconnection	Wire	Flip Chip	TSV	Hybrid ³
Image				
Memory Application	DRAM (Mobile), NAND	DRAM (Computing, Graphics)	DRAM (HPC/Server, Graphics)	DRAM (HPC/Server, Graphics)
Number of I/Os	× 4, × 8, × 16	× 4, × 8, × 16	× 1,024	× 1,024
Number of Physical Interconnections	50-150 ea.	150-200 ea.	Bottom die: 5-8 K ea. Upper die: 3.5-10 K ea. (Up to 200 K ea. with dummy bumps)	Bottom die: 5-8 K ea. Upper die: 3.5-10 K ea. (Up to 200 K ea. with dummy bumps)
Interconnection Length	200 μm-2,000 μm	50 μm	20 μm	5 μm
Number of Stacks	2 / 4 / 6 / 8 (DRAM)	1 / 2 (Planar)	4 / 8 / 12	4 / 8 / 12 / 16
Max. Capacity / PKG	16 GB	4 GB	24 GB	32 GB

一、HBM：算力的内存瓶颈

二、SK海力士HBM工艺分析：TSV、EMC、混合键合成趋势

三、市场测算：未来三年CAGR超80%

四、相关标的：设备、材料迎来新成长

五、风险提示

AI刺激服务器存储容量扩充，HBM需求强劲

- AI服务器刺激更多存储器用量，大容量内存条、HBM、eSSD需求旺盛。根据Trendforce，目前服务器DRAM（模组形态为常规内存条RDIMM和LRDIMM）的普遍配置约为500~600GB，而AI服务器在单条模组上则多采64~128GB，单台服务器搭载16~36条，平均容量可达1TB以上。对于企业级SSD，由于AI服务器追求的速度更高，其要求优先满足DRAM或HBM需求，在SSD的容量提升上则呈现非必要扩大容量的态势，但配置也显著高于常规服务器。
- 随着算力卡更新迭代，HBM规格持续提升。未来在AI模型逐渐复杂化的趋势下，服务器的数据计算和存储需求将快速增长，并同步带动服务器DRAM、企业级SSD以及HBM的需求成长。相较于一般服务器而言，AI服务器多增加GPGPU的使用，以NVIDIA A100/H100 80GB配置8张计算，HBM用量约为640GB，超越常规服务器的内存条容量，H200、B100、MI300等算力卡将搭载更高容量、更高速率HBM。

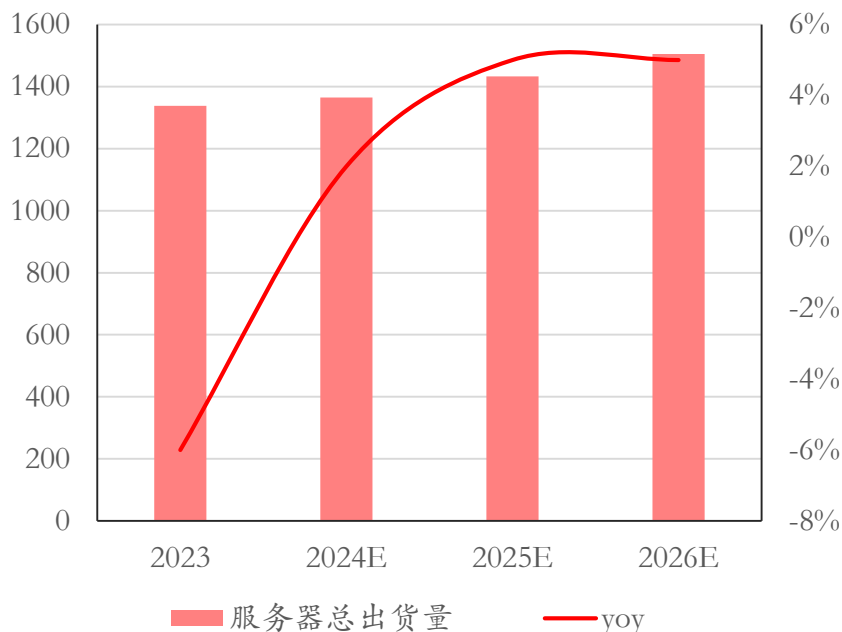
图表：常规服务器和AI服务器存储器配置对比（估算）

	服务器整体			常规服务器		AI服务器	
	2022	2023	模组形态	数量	模组价格 /美元	数量	模组价格 /美元
DRAM容量（GB）	538	603	RDIMM等	12块	1800	16~36块	2400~4800
yoy	18%	12%					
NAND容量（GB）	3351	4167	eSSD	4块	1500	8~12块	3000~4500
yoy	24%	24%					
HBM	0	0	0	0	0	8块	>10000
Raid			Riad卡	1块	100	1块	~100

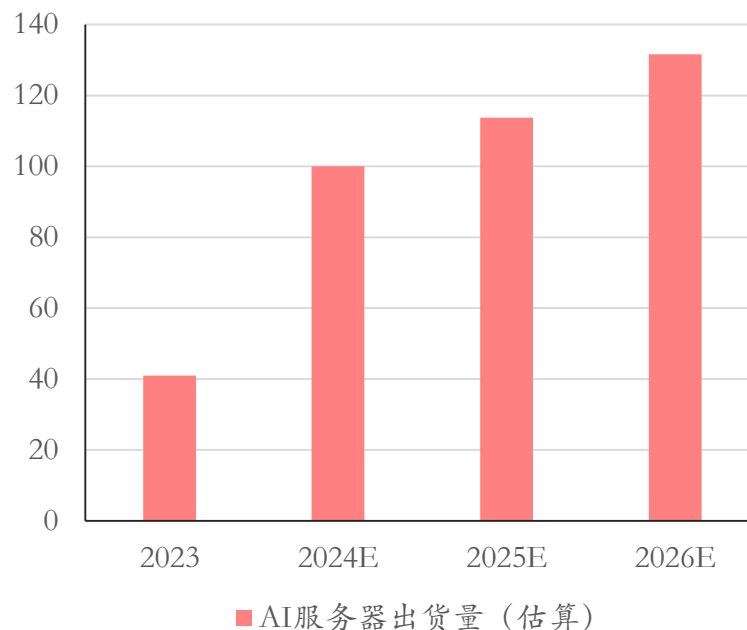
通用服务器呈现弱复苏态势，AI服务器快速增长

- 传统服务器呈现弱复苏态势，2024年出货量同比增长2%。根据Trendforce，2024年服务器出货驱动力以北美CSP为主，但受限于通货膨胀高，企业融资成本居高不下，压缩资本支出，整体需求尚未恢复至疫情前成长幅度，预计2024年全球服务器整机出货量约1365.4万台，同比增长2.05%。
- 受益于北美CSP订单带动，AI服务器ODM对2024年展望乐观。（1）广达：预计2024年AI服务器出货双位数增长，订单主要来自于Microsoft及AWS等；（2）Supermicro：预计2024年AI服务器出货量有机会翻倍成长，订单主要来自CoreWeave与Tesla，积极拓展Apple、Meta等客户AI订单；（3）Inventec：除了北美CSP需求，中国客户如ByteDance需求最强，预估2024年AI服务器出货量年成长可达双位数，占比约10~15%；（4）Foxconn：获得Oracle、AWS订单，预计2024年服务器ODM出货量增长5~7%。

图表：全球服务器出货量（万台）



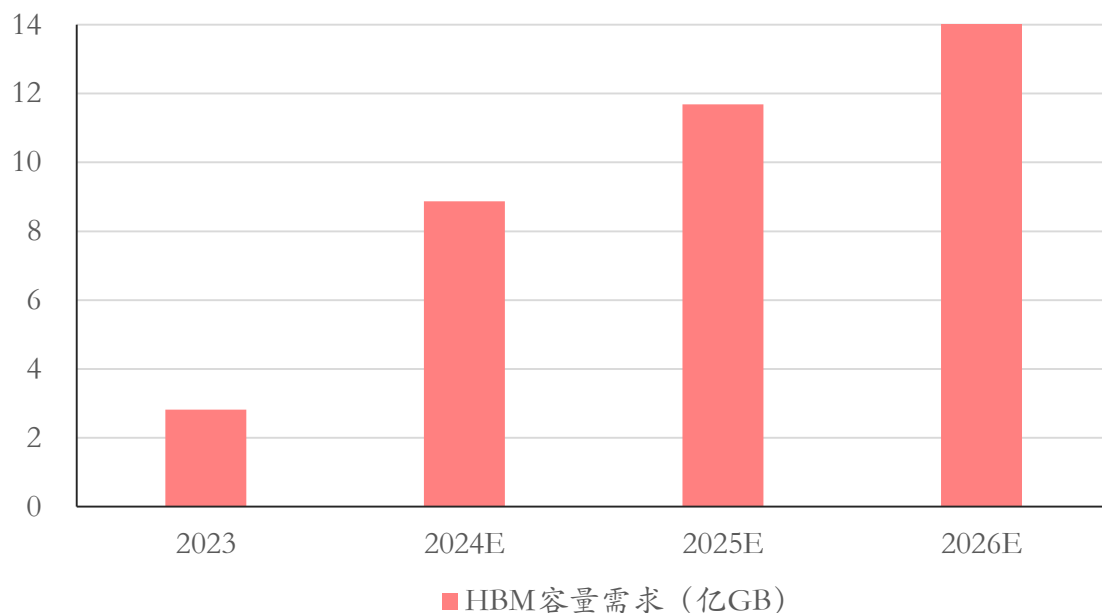
图表：AI服务器出货量（万台）



2024年HBM市场容量有望接近9亿GB

- 我们测算了全球算力卡的HBM需求，结论是：
- 1、从容量看，2023年HBM市场容量为2.8亿GB，预计2024年增长至8.9亿GB，2026年增长至14.1亿GB，2023~2026年CAGR为71%。增长驱动因素是：算力卡单卡HBM容量提升、算力卡出货量提升、新的玩家产品放量。
- AI服务器出货增速远高于传统服务器、手机、PC+NB等传统市场，且单机DRAM容量增速更快，因此HBM在全球DRAM市场的占比将逐步提升，2023年容量占比为1%（2023年大宗DRAM市场容量为275亿GB），2026年将提升至3%。

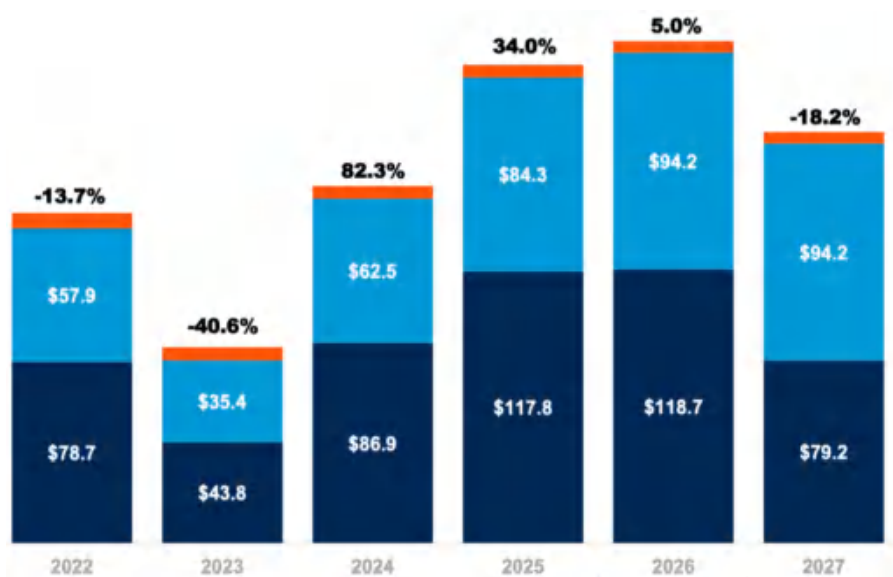
图表：全球HBM容量需求（亿GB）



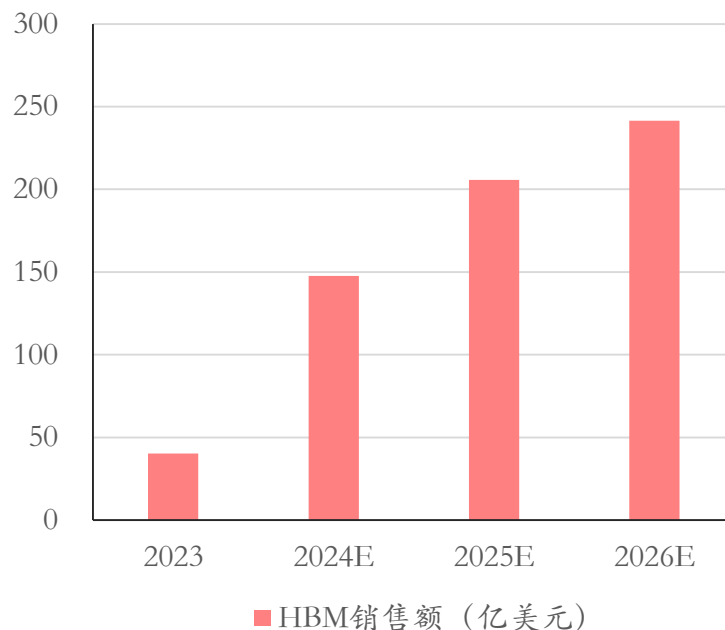
2024年HBM市场规模有望增长267%至148亿美元

- 2、从产值看：2023年HBM市场规模为40亿美元，预计2024年增长至148亿美元，2026年增长至242亿美元，2023~2026年CAGR为82%。增长驱动因素是：算力卡单卡HBM容量提升、算力卡出货量提升、新的玩家产品放量、技术迭代带来单GB HBM单价提升。
- 大宗DRAM具备强商品属性，价格周期性明显，HBM具备定制属性，技术迭代催动ASP提升，长期看HBM在DRAM市场的占比会持续提升，2023年HBM销售额在DRAM市场的占比为8%，2026年将提升至17%。

图表：全球大宗DRAM销售额（亿美元）



图表：全球HBM销售额（亿美元）



HBM市场测算明细及关键假设

- 关键假设：
 - 1、对2025年之后的算力卡增长采用保守假设：2023-2024年算力卡维持紧缺，假设2025年供需平衡，增长趋缓。
 - 2、HBM渗透率假设：2023年HBM市场的主流代际为HBM2、HBM2e和HBM3，预计2024年主流代际转变为HBM3e，由于各家原厂的HBM4方案未定，2025~2026年保守假设以HBM3e为主。
 - 3、HBM价格假设：定价上HBM3e>HBM3>HBM2e>HBM2，并且随着产品迭代，老产品价格有自然年降。其中HBM3e有24GB和36GB版本，以24GB版本测算。
 - 4、大宗DRAM的容量和销售额来自Gartner预测，包含了2024~2025年的涨价周期和2026年周期顶部的假设，数据中不含HBM。
- （当前讨论2025~2026年算力卡及HBM配置尚早，诸多方案未定，因此测算时算力卡出货、HBM容量和定价均采用了保守假设）

HBM市场测算明细及关键假设

	2023	2024E	2025E	2026E
各代际HBM出货占比假设				
HBM2/2e (8H)	60%	10%	0%	0%
HBM3 (12H)	40%	50%	20%	0%
HBM3e (12H)	0%	40%	80%	100%

单价 (美元/1GB)

HBM2/2e (8H)	13	12	11	11
HBM3 (12H)	16	14	14	13
HBM3e (12H)	22	20	19	18

HBM容量需求

海外	2.7	8.3	10.9	12.7
国内	0.2	0.5	0.8	1.4
HBM容量需求 (亿GB)	2.8	8.9	11.7	14.1
yoy		214%	32%	20%

HBM销售额

海外	381427	1415081	1967950	2272143
国内	20800	60840	88920	143606
HBM销售额 (亿美元)	40.2	147.6	205.7	241.6
yoy		267%	39%	17%

大宗DRAM市场规模 (亿美元)	438	869	1178	1187
yoy	-44%	98%	36%	1%
大宗DRAM+HBM市场规模 (亿美元)	478	1017	1384	1429
yoy	-	113%	36%	3%
HBM产值占比	8%	15%	15%	17%

大宗DRAM容量 (GB)	275	325	415	495
yoy	12%	18%	27%	19%
大宗DRAM+HBM容量 (GB)	278	334	426	509
yoy		20%	28%	19%
HBM容量占比	1%	3%	3%	3%

一、HBM：算力的内存瓶颈

二、SK海力士HBM工艺分析：TSV、EMC、混合键合成趋势

三、市场测算：未来三年CAGR超80%

四、相关标的：设备、材料迎来新成长

五、风险提示

HBM供应链国产潜在供应商

- 我们认为，目前HBM供应链以海外厂商为主，部分国内厂商打入了海外存储/HBM供应链。国产HBM正处于0到1的突破期，目前HBM供应主要为韩系、美系厂商，国内能获得的HBM资源较少。随着国产算力卡需求快速增长，对于算力卡性能至关重要的HBM也有强烈的供应保障诉求和国产化诉求。相关标的：封测、设备、材料等环节。

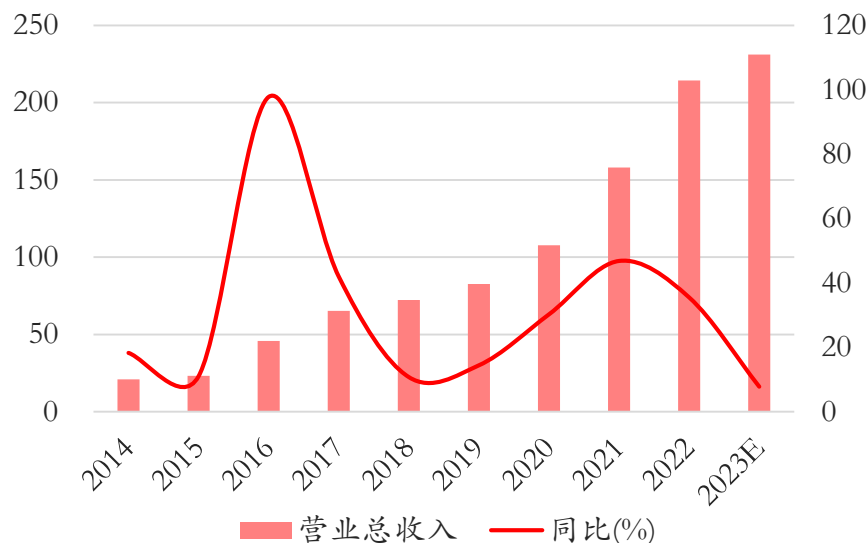
图表：HBM供应链国产潜在供应商

分类	存储/HBM相关业务	代码	公司	市值	归母净利润			PE		
					2023E	2024E	2025E	23E	24E	25E
设备	PVD、刻蚀、PECVD	002371.SZ	北方华创	1,545	38.1	52.9	69.3	40	29	22
	刻蚀、LPCVD	688012.SH	中微公司	903	17.9	19.8	25.2	51	46	36
	PECVD、ALD、混合键合	688072.SH	拓荆科技	382	6.6	8.3	11.3	57	46	34
	临时键合、解键合	688037.SH	芯源微	159	2.5	4.1	5.7	63	39	28
	减薄、CMP	688120.SH	华海清科	310	7.3	10.1	13.2	43	31	24
	量测和检测设备	603283.SH	赛腾股份	149	6.3	7.6	8.7	24	20	17
	存储检测设备	688627.SH	精智达	56	1.2	1.6	2.2	48	35	25
	固晶机	688383.SH	新益昌	69	0.6	2.6	3.4	114	26	20
材料	前驱体	002409.SZ	雅克科技	249	6.7	9.9	13.6	37	25	18
	硅微粉	688300.SH	联瑞新材	85	1.7	2.5	3.2	49	33	27
	GMC颗粒状环氧塑封料	688535.SH	华海诚科	69	0.3	0.6	0.7	208	121	96
	PSPI电镀液	300429.SZ	强力新材	52	0.2	1.5	2.0	260	35	26
	电镀液	688603.SH	天承科技	30	0.6	0.9	1.4	51	32	22
	临时键合材料	300398.SZ	飞凯材料	70	3.5	4.9	6.1	20	14	11
	Low-α球形氧化铝粉体	688733.SH	壹石通	42	0.2	1.2	1.8	172	34	23
	PCBGA 封装板	002436.SZ	兴森科技	227	2.6	4.3	6.6	87	53	34
封测	存储封测	002156.SZ	通富微电	359	1.9	9.0	12.6	187	40	29
	存储封测	600584.SH	长电科技	493	15.4	27.3	35.0	32	18	14
	DRAM封测	000021.SZ	深科技	226	7.6	9.3	11.2	30	24	20
代理	海力士HBM代理	300475.SZ	香农芯创	172	3.4	4.4	5.3	50	39	32

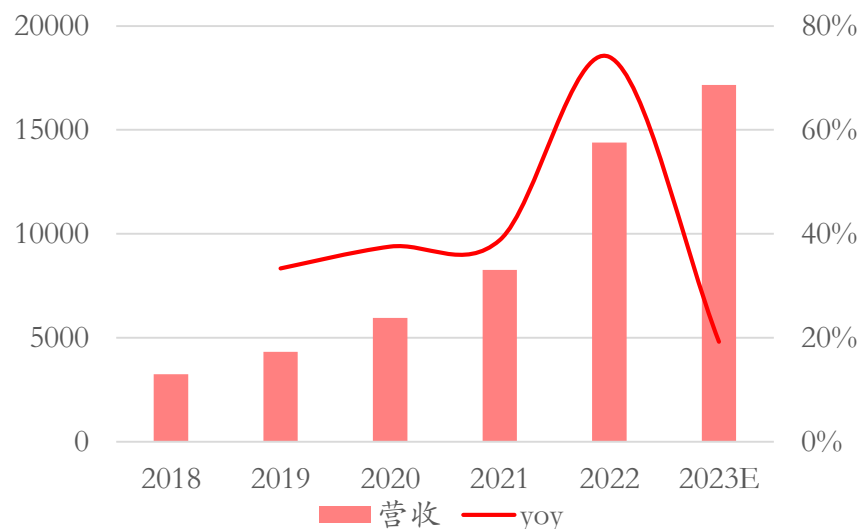
通富微电：绑定AMD，突破Chiplet、存储等高端封测技术

- 公司为全球第四大、中国大陆第二大第三方封测厂商，主要客户有AMD、恩智浦、联发科、德州仪器、意法半导体、英飞凌、兆易创新、长鑫存储、长江存储等。
- 并购AMD苏州&槟城工厂，跃居全球第四，突破Chiplet、存储等高端封测技术。2016年公司收购AMD苏州及槟城各85%的股权，通过并购，公司与AMD形成了“合资+合作”的模式，成为AMD最大的封测供应商，占AMD封测订单的80%以上。得益于此，公司2022年营收突破200亿元，全球市占率达到6.51%。AMD完成对全球FPGA龙头赛灵思的收购，实现了CPU+GPU+FPGA+AI的全方位布局，双方在客户资源、IP和技术组合上具有高度互补性。此外，公司已经建成Chiplet、2.5D/3D等封装技术平台，也实现了NAND Flash和LPDDR的多层堆叠封装。基于高端处理器和AI芯片封测需求的不断增长，公司的AMD槟城工厂正在大力扩产。

图表：通富微电历年营收及增速（亿元，%）



图表：超威苏州&槟城工厂历年营收及增速（百万元，%）



中微公司：超高深宽比刻蚀设备取得突破

- 等离子体刻蚀设备在先进存储已量产。在存储芯片制造环节，公司的等离子体刻蚀设备已大量用于先进三维闪存和动态随机存储器件的量产。公司可提供超高深宽比掩膜 ($\geq 40:1$) 和超高深宽比介质刻蚀 ($\geq 60:1$) 的全套解决方案，开发了配备超低频偏压射频的ICP刻蚀机用于超高深宽比掩膜的刻蚀，并且开发了配备超低频高功率偏压射频的CCP刻蚀机用于超高深宽比介质刻蚀。
- 薄膜沉积设备获得突破。公司已实现多种LPCVD设备的研发交付以及ALD设备的重大突破，多款新产品研发项目在快速推进。公司开发的新型号CVD钨设备用创新的工艺解决方案，满足了器件中超高深宽比的接触孔填充要求。公司开发的ALD钨设备，能够满足3D NAND等三维器件结构中金属钨的填充需求。同时公司在开发的另一ALD产品系列，具有业界领先的阶梯覆盖率、均一性和输出效率，能够满足先进逻辑和存储器件中金属阻挡层和金属栅极的应用需求。

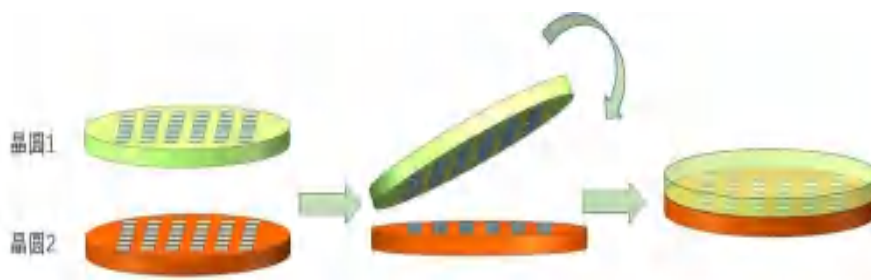
图表：中微公司CCP、ICP刻蚀设备



拓荆科技：已布局晶圆级3D集成的混合键合设备

- 公司目前已形成 PECVD、ALD、SACVD、HDPCVD 等薄膜设备产品系列，广泛应用于国内集成电路逻辑芯片、存储芯片等制造产线。
- 针对先进封装领域需求，公司已布局应用于晶圆级三维集成的混合键合设备。主要包括两种设备类别：
(1) 晶圆对晶圆键合 (Wafer to Wafer Bonding)：在常温下实现复杂的12英寸晶圆对晶圆多材料表面的键合工艺，同时还可实现熔融键合工艺；(2) 芯片到晶圆键合 (Die to Wafer Bonding)：分为预处理和键合两道工序，芯片对晶圆键合表面预处理产品可以实现芯片对晶圆键合前表面预处理工序。

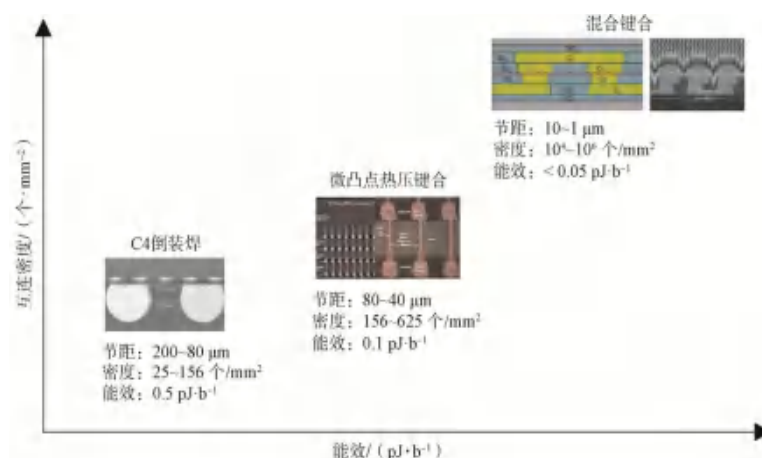
图表：晶圆键合设备应用示意图



图表：公司混合键合产品两类情况

主要产品型号	主要应用
Dinone 300	晶圆对晶圆常温混合键合(Hybrid Bonding)和熔融键合(Fusion Bonding)
Pollux	晶圆及切割后芯片的表面活化及清洗

图表：键合技术演进



- 混合键合设备是晶圆级三维集成应用中最前沿的核心设备之一，可以在常温下通过永久键合工艺技术实现芯片或晶圆的堆叠，提升芯片间的通信速度及整体芯片性能，相比于传统工艺还有 有效缩短芯片开发周期、有效保证键合面气密性等特点。

芯源微：临时键合与解键合技术进展顺利

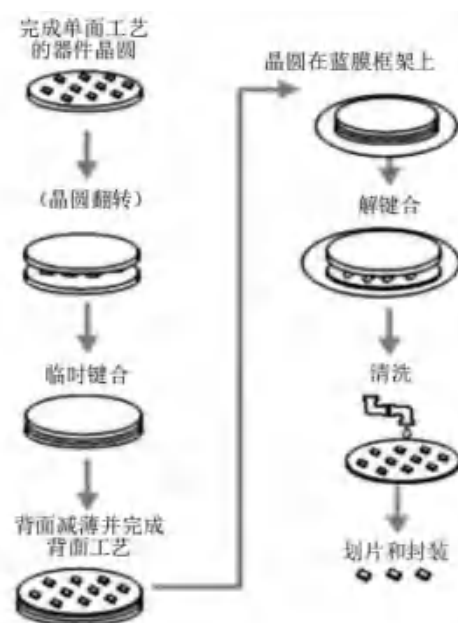
- 晶圆与键合载片键合后，键合载片可以为晶圆提供机械支撑；在减薄抛光工艺完成后，通过解键合将晶圆与键合载片分离，该过程称为临时键合及解键合技术。（1）临时键合主要包括3个步骤：键合介质涂覆、前烘固化、晶圆与键合载片对准，以及真空热压。（2）解键合是在减薄抛光完成后，将晶圆与载片分开的技术。解键合按照翘曲矫正、热解滑移和晶圆清洗3个过程进行。
- 芯源微产品包括光刻工序涂胶显影设备和单片式湿法设备，是国内唯一可以提供量产型前道涂胶显影机的厂商，目前已完成在前道晶圆加工环节28nm以上工艺节点的全覆盖，并可持续向更高工艺等级迭代。公司临时键合机、解键合键合机进展顺利，于2023年度申请并生效了《晶圆解键合系统及其控制方法》和《一种键合片处理装置》的发明专利。

图表：光刻工艺流程



- TSV和三维堆叠型3D集成制造中，减薄后晶圆厚度越来越薄；
- 超薄器件晶圆机械强度降低，容易翘曲和起伏，易造成器件性能降低、产品均一性变差，生产过程中碎片率增加。

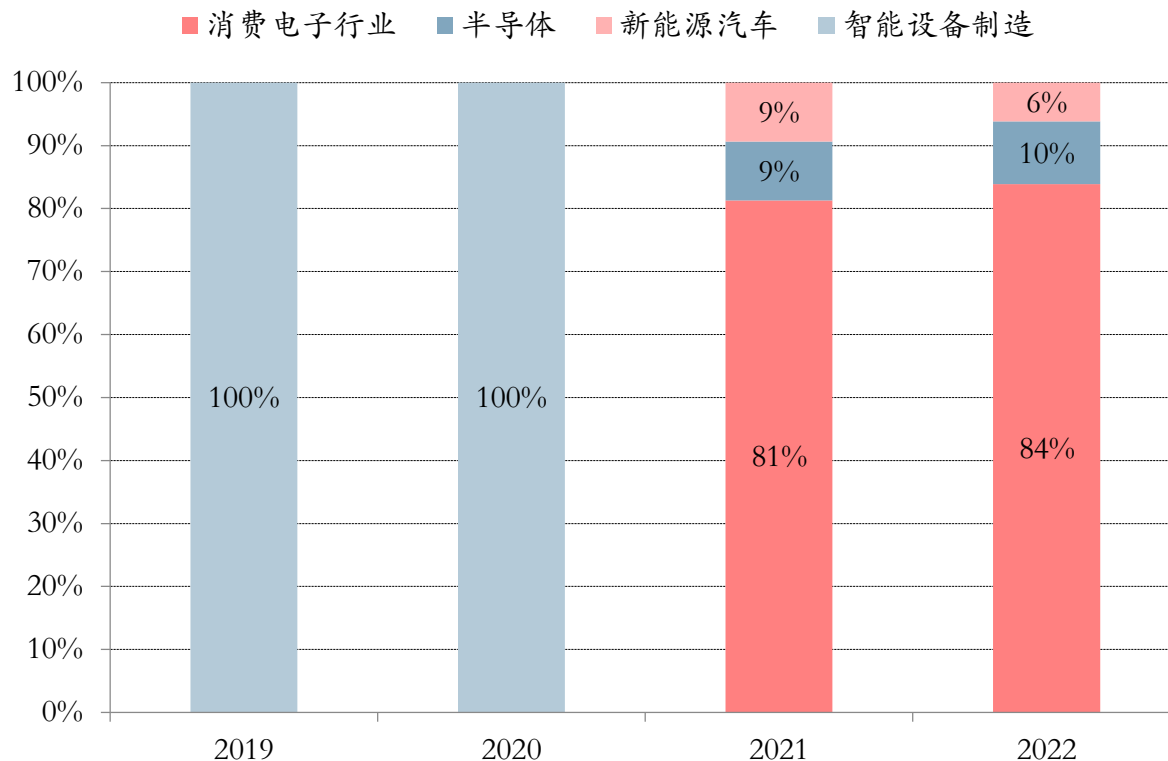
图表：临时键合和解键合步骤



赛腾股份：收购Optima进入晶圆检测市场

- 赛腾股份2011年切入苹果供应链进入发展快车道，2019年通过收购全球领先的晶圆检测设备供应商日本Optima涉足晶圆检测装备领域。在半导体方面，2022年收入达到2.91亿元，占营收比例为10%，目前公司拥有Sumco、Sksiltron、Samsung、奕斯伟、中环半导体、金瑞泓等优质半导体客户。

图表：2019年-2022年赛腾股份营业收入构成



数据来源：赛腾股份、中信建投

赛腾股份：收购Optima进入晶圆检测市场

- 收购Optima之后，公司实现了在国内高端半导体设备市场的进一步突破。Optima主要从事半导体硅片检查设备和测量设备的开发、制造和销售，产品线主要包括硅片边缘缺陷自动检测设备RXW-1200、晶圆片用背面检测设备BMW1200、边缘/表背面复合检测设备RXM-1200、晶圆制造外观检测AXM-1200四类产品，应用于半导体制造过程中键合、修边、刻蚀等的质量控制环节，有效保证芯片生产良品率。

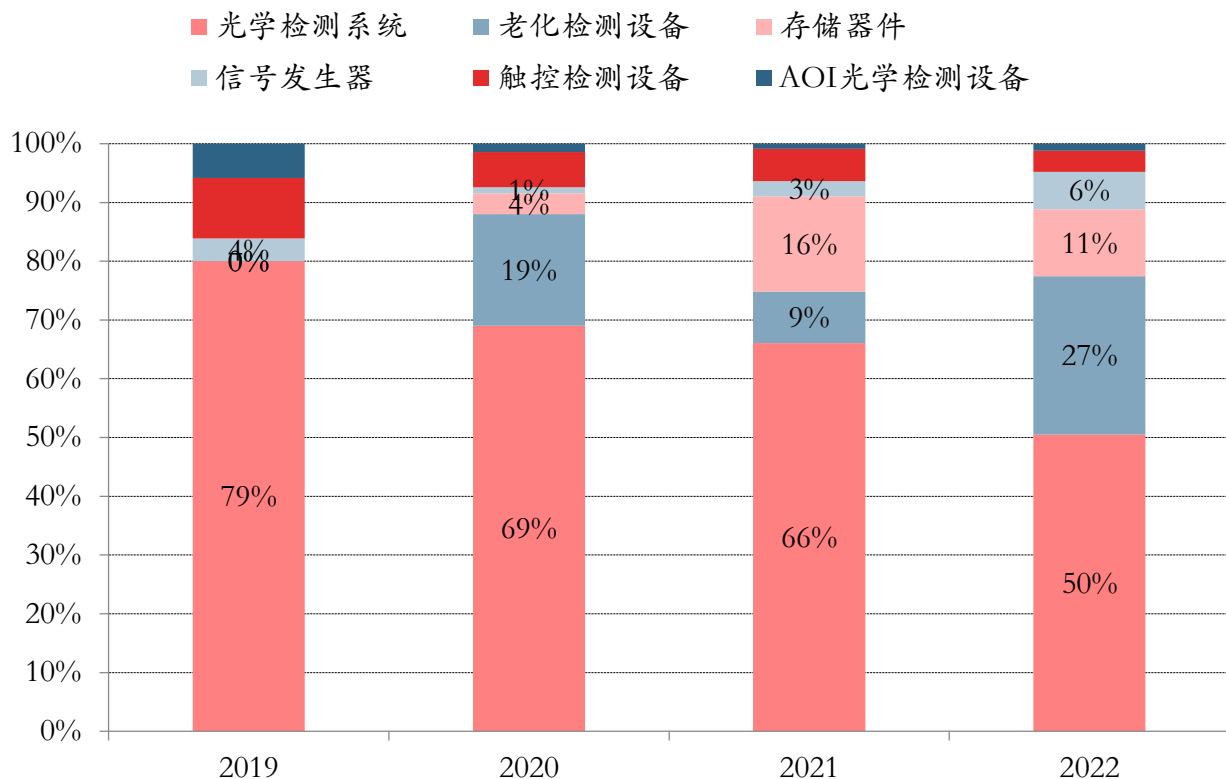
图表：Optima主要检测设备

设备名称	用途	制造工艺	图例
硅片边缘缺陷自动检测设备RXW-1200	在硅片制造，晶圆片制造过程，对硅片边缘部的缺陷检出和分类，要求部位的尺寸进行量测的在线检测设备。	倒角/表面研磨/表面抛光/清洗/干燥/PW终检/外延、SOI/成膜/光刻/CMP/修边/键合/减薄	
晶圆片用背面检测设备BMW-1200	能够以高灵敏度检测，晶圆片工艺中在晶片背面产生的缺陷和异物并测量提取缺陷的微小三维形状的检测设备。	成膜/光刻/蚀刻/CMP	
边缘/表背面复合检测设备RXM-1200	检测硅片制造（抛光，外延）过程中，检测出边缘/表背面出现的各种缺陷的复合设备。	倒角/表面研磨/表面抛光/清洗/干燥/PW终检/Epi、SOI等	
AXM-1200	AXM 系列是最先进的设备，从Notch检查新方法开始，进一步提高了边缘/正面/背面的检查能力，通过新开发的原始算法自动进行缺陷分类，并且配备了各种新功能。	硅片制造工艺中的边缘/正面/背面/Notch检查	

精智达：卡位DRAM晶圆测试机与FT测试机

- 精智达业务涵盖新型显示器件的光学检测和校正修复、老化、信号发生等领域，在半导体领域与韩国UniTest合作开展DRAM晶圆老化测试设备的研发和DRAM老化修复设备的本地化生产，同时独立开展MEMS探针卡、DRAM晶圆测试机、DRAM FT测试机等研发，已开发客户包括长鑫存储、兆易创新、沛顿科技、晋华集成、通富微电、太极实业等。

图表：2019年-2022年精智达营业收入构成（%）情况



精智达：卡位DRAM晶圆测试机与FT测试机

- 投资高铂科技布局高速测试机SOC。公司IPO募投1.62亿元，投入半导体存储器件测试设备研发，包括动态存储技术测试系统研发，MEMS探针卡研发试制等。另外，公司2023年以人民币2900万元投资深圳高铂科技，将持有标的公司10.03%股权。标的公司提供存储芯片测试开发服务，设计完成的SC9800高性能数字混合信号测试芯片，用于存储测试，覆盖DDR4、DDR5、LPDDR4、LPDDR5等，最高支持9Gbps速率传输协议，该芯片已于2023年10月正式投产。

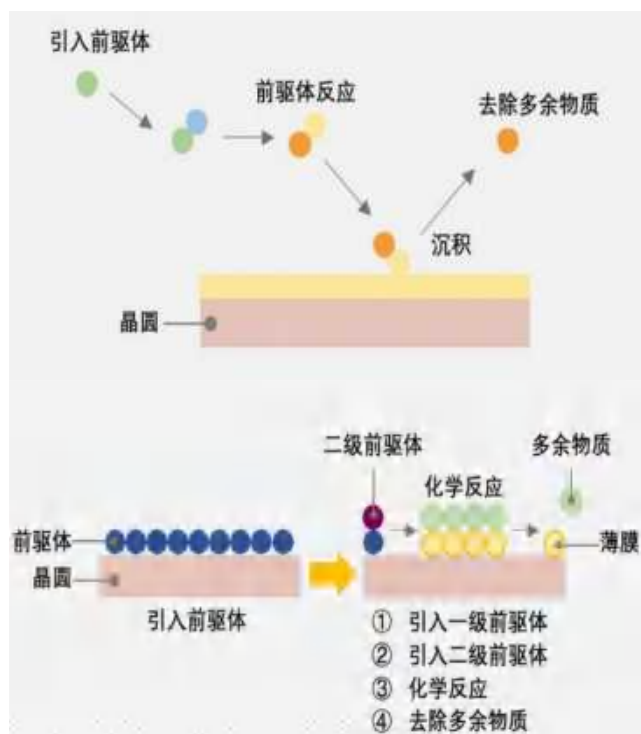
图表：精智达半导体存储器件测试设备

设备名称	功能	图例
MEMS探针卡	主要用晶圆测试时实现测试机与被测裸片的电气联接，通过传输信号对芯片参数进行测试。	
DRAM老化修复设备	对封装后的芯片颗粒进行高低温与大电流环境下的老化测试，在测试中对颗粒内部缺陷进行修复。融合高低温、老化冲击、功能测试等各项测试工艺，并对检测出的不良进行软件算法修复，可以取代多道化统的品圆及封装老化流程，实现高吞吐容量的电学性能与可靠性检验要求。	
DRAM晶圆测试机	设计面向新一代DRAM存储器晶圆的测试机，包括探针卡接口装置;已与睿力集成(长鑫存储)就样机测试验证签订合作协议。	
DRAM FT 测试机	对封装后的芯片颗粒进行实际应用条件下的功能指标测试，对芯片施加输入信号、采集输出信号、并判断芯片在不同工作条件下功能与性能的有效性，通过通信接口将测试结果传送给分选机，分选机据此对被测试芯片进行标记、分选等。	

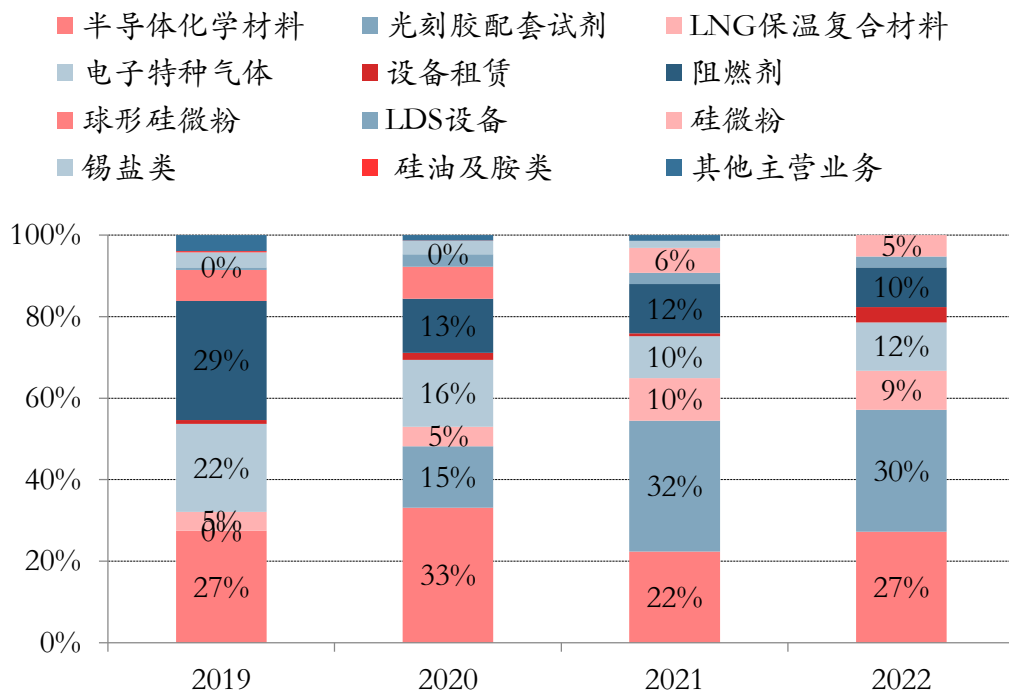
雅克科技：前驱体为海力士和三星核心供应商， High-K材料已获突破

- 雅克科技作为前驱体核心供应商，受益于存储器、HBM扩产。公司是前驱体材料唯一国产供应商，2022年全球市占率约为10%。全球80%的HBM市场份额由SK海力士和三星占据，而雅克科技是三星和海力士的核心供应商，将受益HBM需求持续增加。
- High-K材料已突破。公司进一步布局半导体先进制程配套的前驱体材料，推出新High-K材料及超高/低温硅类产品，部分产品已经进入客户端测试。2022年半导体化学材料和光刻胶配套试剂为雅克科技前两大业务，分别占比27%和30%。

图表：前驱体在CVD、ALD中原理解图



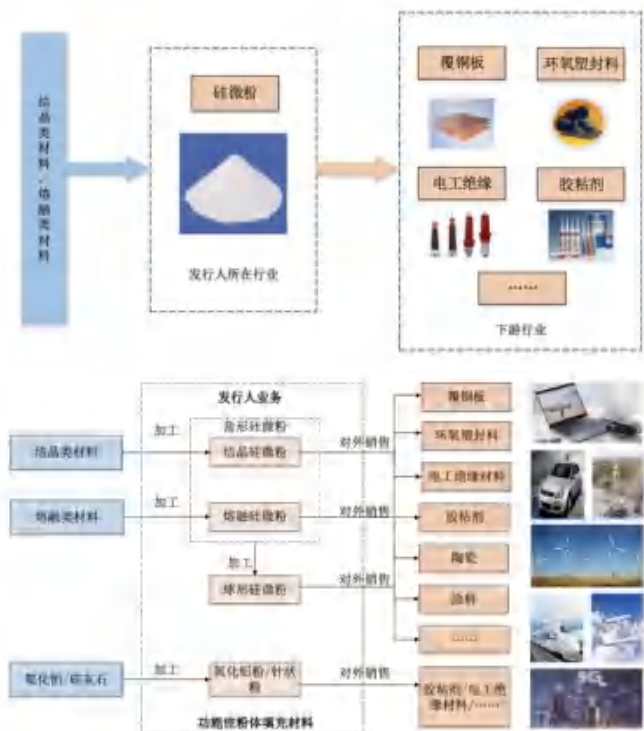
图表：2019年-2022年雅克科技营业收入结构



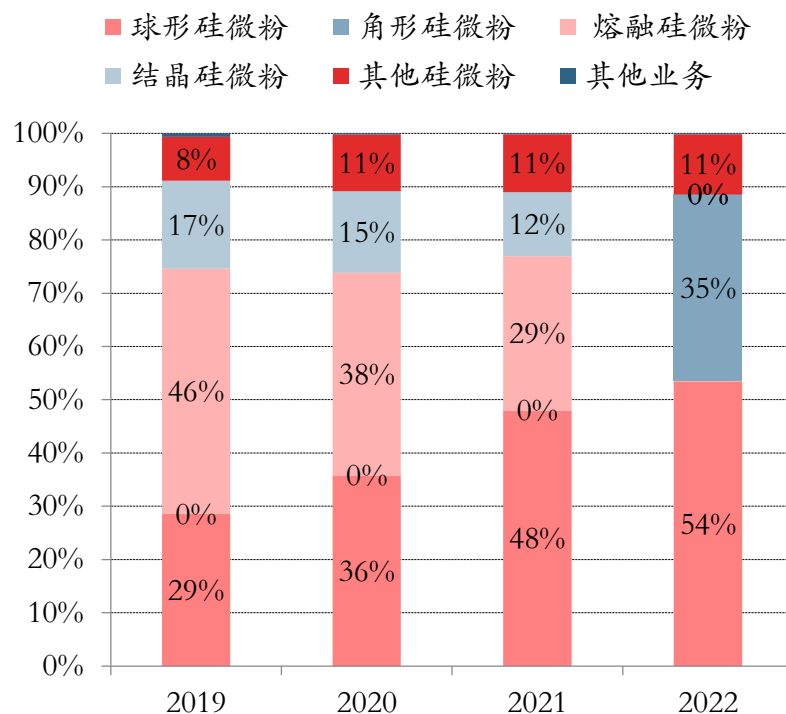
联瑞新材：配套供应GMC所用球硅和Low-α球铝

- 联瑞新材是国内领先的电子级硅微粉企业，产品包括结晶硅微粉、熔融硅微粉和球形硅微粉，产品可用于电子电路用覆铜板、芯片封装用环氧塑封料以及电工绝缘材料、胶粘剂、陶瓷、涂料等。
- 公司Low-α球硅和球铝已批量供应GMC客户。公司球形硅微粉2022年收入3.54亿元，占营收比例54%。HBM所需的颗粒封装材料（GMC）中需要添加TOPCUT 20um以下球硅和Low-α球铝，公司部分客户是全球知名的GMC供应商，公司配套供应HBM所用球硅和Low-α球铝。

图表：公司所处行业与其上下游行业关系



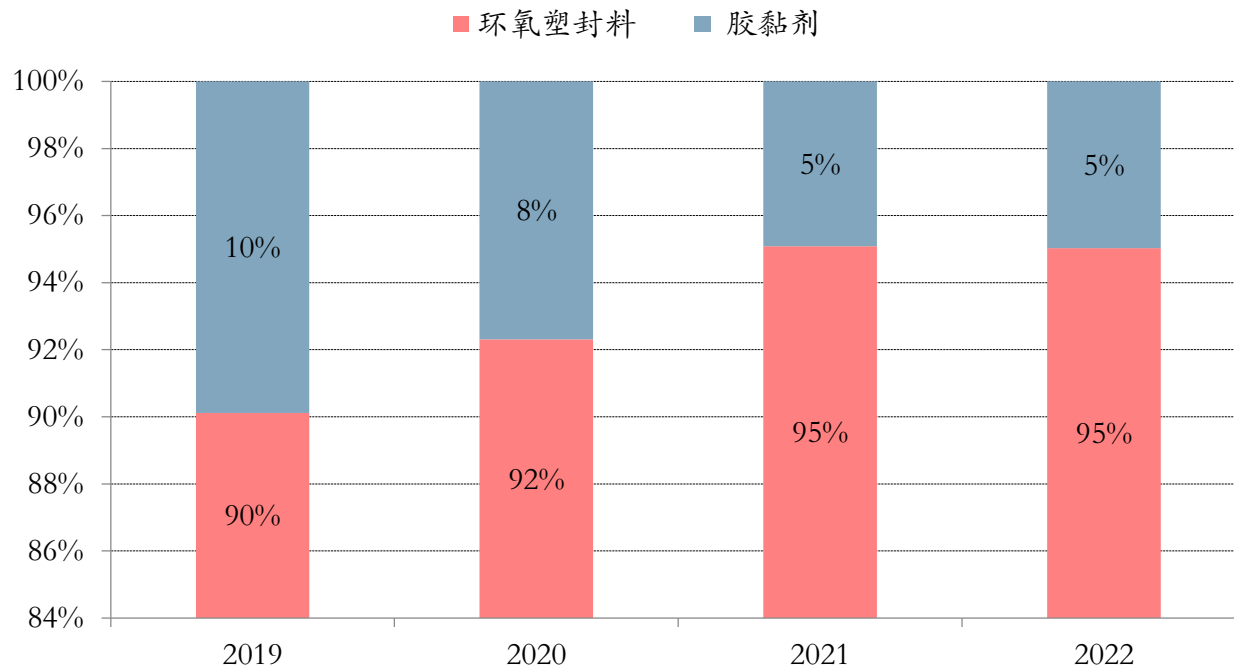
图表：2019年-2022年联瑞新材分业务营业收入结构(%)



华海诚科：成功开发LMC与GMC材料，逐步打破外资垄断

- 华海诚科主要产品为环氧塑封料和电子胶黏剂，构建了可应用于传统封装（包括DIP、TO、SOT、SOP等）与先进封装（QFN/BGA、SiP、FC、FOWLP/FOPLP等）的产品体系。
- 得益于扬杰科技、华天科技等功率、封装类客户需求的增长，公司近几年收入规模快速扩张，2014年~2021年收入从0.78亿元增长至3.47亿元，CAGR=23%，2022年~2023年因行业景气度下行，客户端稼动率不满，收入规模有所承压。2024年随着下游客户稼动率回升+高端材料突破，公司业绩有望出低回升。

图表：2019年-2022年华海诚科分业务营业收入结构

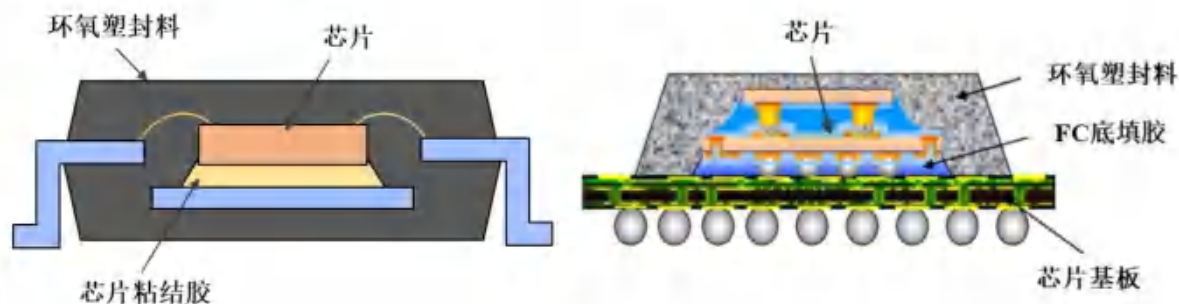


华海诚科：成功开发LMC与GMC材料，逐步打破外资垄断

- 公司针对系统级封装与晶圆级封装成功研发了液态塑封材料（LMC）、颗粒状环氧塑封料（GMC）、FC底填胶等，有望打破外资厂商在先进封装用高端材料的垄断。与传统封装采用固态饼状环氧塑封料不同，应用于FOWLP封装的GMC与LMC以颗粒状与液态为主，因而对塑封料厂商的生产工艺提出了更高要求。以GMC为例，目前制备颗粒状环氧塑封料的主流技术为离心法和热切割法，对塑封料厂商的配方技术、生产工艺、生产设备、产品测试方法等要求较高，故市场基本由外资厂商垄断。

图表：华海诚科核心技术，及环氧塑封料应用场景

材料类型	基本情况
颗粒状环氧塑封料（GMC）	颗粒状环氧塑封料在塑封过程采用均匀撒粉的方式，在预热后变为液态，将带有芯片的承载板浸入到树脂中而成型，凭借操作简单、工时较短、成本较低等优势，GMC有望发展成为主要的晶圆级封装塑封材料之一，市场发展前景良好
液态塑封料（LMC）	LMC是通过将液态树脂挤压到产品中央，在塑封机温度和压力的作用下增强液态树脂的流动性，从而填充满整个晶圆。LMC具备可中低温固化、低翘曲、模塑过程无粉尘、低吸水率以及高可靠



一、HBM：算力的内存瓶颈

二、SK海力士HBM工艺分析：TSV、EMC、混合键合成趋势

三、市场测算：未来三年CAGR超80%

四、相关标的：设备、材料迎来新成长

五、风险提示

风险提示

- AI技术落地不及预期。AI属于创新技术，需投入较大前期研发成本，并且AI技术的落地应用是一个复杂的过程，受到许多因素的影响，包括技术成熟度、数据质量、人才短缺、法规限制等。因此，AI技术的落地可能会出现一些挑战和困难，导致实际应用效果不如预期；
- 国际环境变化影响。美联储货币政策不确定，影响科技行业估值；美国不断对中国科技施压，衍生出供应链安全、知识产权争议、政策和监管环境等问题。
- 宏观环境的不利因素。宏观不利因素可能使得全球经济增速放缓，AI服务器需求不及预期。如果宏观经济波动较大或长期处于低谷，政府、企业资本开支将受到抑制，AI服务器市场的波动会影响产业链相关公司的经营业绩。
- 市场竞争加剧风险。HBM市场寡头垄断，且迭代快速，国产厂商与海外差距巨大，若研发进展不及预期，可能导致技术差距扩大，导致未来可获得的市场份额低于预期。

分析师介绍

刘双锋：电子行业分析师，执业证书编号：S1440520070002。电子&TMT海外牵头人及港深研究组长。3年深南电路，5年华为工作经验，从事市场洞察、战略规划工作，涉及通信服务、云计算及终端领域，专注于通信服务领域，2018年加入中信建投通信团队。2018年IAMAC最受欢迎卖方分析师通信行业第一名团队成员，2018《水晶球》最佳分析师通信行业第一名团队成员。覆盖消费电子与半导体领域。

孙芳芳：电子行业分析师，SAC执证编号：S1440520060001。同济大学材料学硕士，2015年8月加入浙商证券，任电子行业首席，2020年5月加入中信建投电子团队，覆盖半导体设备与材料、安防、面板、芯片设计等。

章合坤：电子行业分析师，SAC执证编号：S1440522050001。上海交通大学材料科学与工程硕士，2020年加入中信建投电子团队，覆盖芯片设计、光学等领域。

研究助理介绍

何显灵：电子行业研究助理。复旦大学硕士，2022年加入中信建投电子团队。

评级说明

投资评级标准		评级	说明
报告中投资建议涉及的评级标准为报告发布日后6个月内的相对市场表现，也即报告发布日后的6个月内公司股价（或行业指数）相对同期相关证券市场代表性指数的涨跌幅作为基准。A股市场以沪深300指数作为基准；新三板市场以三板成指为基准；香港市场以恒生指数作为基准；美国市场以标普500指数为基准。	股票评级	买入	相对涨幅15%以上
		增持	相对涨幅5%—15%
		中性	相对涨幅-5%—5%之间
		减持	相对跌幅5%—15%
		卖出	相对跌幅15%以上
	行业评级	强于大市	相对涨幅10%以上
		中性	相对涨幅-10-10%之间
		弱于大市	相对跌幅10%以上

分析师声明

本报告署名分析师在此声明：(i) 以勤勉的职业态度、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，结论不受任何第三方的授意或影响。(ii) 本人不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

法律主体说明

本报告由中信建投证券股份有限公司及/或其附属机构（以下合称“中信建投”）制作，由中信建投证券股份有限公司在中华人民共和国（仅为本报告目的，不包括香港、澳门、台湾）提供。中信建投证券股份有限公司具有中国证监会许可的投资咨询业务资格，本报告署名分析师所持中国证券业协会授予的证券投资咨询执业资格证书编号已披露在报告首页。

在遵守适用的法律法规情况下，本报告亦可能由中信建投（国际）证券有限公司在香港提供。本报告作者所持香港证监会牌照的中央编号已披露在报告首页。

一般性声明

本报告由中信建投制作。发送本报告不构成任何合同或承诺的基础，不因接收者收到本报告而视其为中信建投客户。

本报告的信息均来源于中信建投认为可靠的公开资料，但中信建投对这些信息的准确性及完整性不作任何保证。本报告所载观点、评估和预测仅反映本报告出具日该分析师的判断，该等观点、评估和预测可能在不发出通知的情况下有所变更，亦有可能因使用不同假设和标准或者采用不同分析方法而与中信建投其他部门、人员口头或书面表达的意见不同或相反。本报告所引证券或其他金融工具的过往业绩不代表其未来表现。报告中所含任何具有预测性质的内容皆基于相应的假设条件，而任何假设条件都可能随时发生变化并影响实际投资收益。中信建投不承诺、不保证本报告所含具有预测性质的内容必然得以实现。

本报告内容的全部或部分均不构成投资建议。本报告所包含的观点、建议并未考虑报告接收人在财务状况、投资目的、风险偏好等方面的具体情况，报告接收者应当独立评估本报告所含信息，基于自身投资目标、需求、市场机会、风险及其他因素自主做出决策并自行承担投资风险。中信建投建议所有投资者应就任何潜在投资向其税务、会计或法律顾问咨询。不论报告接收者是否根据本报告做出投资决策，中信建投都不对该等投资决策提供任何形式的担保，亦不以任何形式分享投资收益或者分担投资损失。中信建投不对使用本报告所产生的任何直接或间接损失承担责任。

在法律法规及监管规定允许的范围内，中信建投可能持有并交易本报告中所提公司的股份或其他财产权益，也可能在过去12个月、目前或者将来为本报中所提公司提供或者争取为其提供投资银行、做市交易、财务顾问或其他金融服务。本报告内容真实、准确、完整地反映了署名分析师的观点，分析师的薪酬无论过去、现在或未来都不会直接或间接与其所撰写报告中的具体观点相联系，分析师亦不会因撰写本报告而获取不当利益。

本报告为中信建投所有。未经中信建投事先书面许可，任何机构和/或个人不得以任何形式转发、翻版、复制、发布或引用本报告全部或部分内容，亦不得从未经中信建投书面授权的任何机构、个人或其运营的媒体平台接收、翻版、复制或引用本报告全部或部分内容。版权所有，违者必究。

中信建投证券研究发展部

北京
东城区朝内大街2号凯恒中心B
座12层
电话：(8610) 8513-0588
联系人：李祉瑶
邮箱：lizhiyao@csc.com.cn

上海
浦东新区浦东南路528号南塔2103室
电话：(8621) 6882-1612
联系人：翁起帆
邮箱：wengqifan@csc.com.cn

深圳
福田区福中三路与鹏程一路交汇处
广电金融中心35楼
电话：(86755) 8252-1369
联系人：曹莹
邮箱：caoying@csc.com.cn

中信建投（国际）

香港
中环交易广场2期18楼
电话：(852) 3465-5600
联系人：刘泓麟
邮箱：charleneliu@csci.hk

