



Research and
Development Center

HBM3E 量产在即，关注国产 HBM 突破和产业链受益

2024 年 03 月 10 日

HBM3E 量产在即，关注国产 HBM 突破和产业链受益

2024 年 03 月 10 日

本期内容提要：

- **海外大厂 HBM3E 量产在即。**（1）近日，美光宣布已开始量产其 HBM3E 高带宽内存解决方案。英伟达 H200 Tensor Core GPU 将采用美光 8 层堆叠的 24GB 容量 HBM3E 内存，并于 2024 年第二季度开始出货。美光 HBM3E 引脚速率超过 9.2Gb/s，提供超过 1.2TB/s 的内存带宽。与 HBM3 相比，HBM3E 将数据传输速率和峰值内存带宽提高了 44%。（2）三星发布首款 36GB HBM3E 12H DRAM，目前为三星容量最大的 HBM。三星 HBM3E 12H 支持全天候最高带宽达 1280GB/s，产品容量也达到了 36GB。相比三星 8 层堆叠的 HBM3 8H，HBM3E 12H 在带宽和容量上大幅提升超过 50%。相比 HBM3 8H，HBM3E 12H 搭载于人工智能应用后，公司预计人工智能训练平均速度可提升 34%，同时推理服务用户数量也可增加超过 11.5 倍。（3）SK 海力士于 1 月中旬正式结束了 HBM3E 的开发工作，并顺利完成英伟达历时半年的性能评估，计划于 3 月开始大规模生产 HBM3E 产品，这批 HBM3E 将用于英伟达下一代 Blackwell 系列的 AI 芯片旗舰产品 B100 上，而英伟达则计划于 2024 年第二季度末或第三季度初推出该系列产品。
- **国内存储厂商入局 HBM 市场。**根据采招网，近日，武汉新芯发布《高带宽存储芯粒先进封装技术研发和产线建设》招标项目，利用三维集成多晶圆堆叠技术，打造更高容量、更大带宽、更小功耗和更高生产效率的国产高带宽存储器（HBM）产品，推进多晶圆堆叠工艺产业化，新增生产设备约 17 台/套，拟实现月产出能力≥3000 片（12 英寸）。我们认为，国内存储厂商在 HBM 技术上的加速突破，有望在 AI 大浪潮的需求下提升竞争实力，相关产业链也或将受益。
- **AI 浪潮驱动，HBM 先进封装、设备材料产业链环节持续受益。**HBM 制造工艺包括 TSV、Bumping 和堆叠等工艺环节。HBM 是由多个 DRAM die 堆叠而成，利用硅通孔（TSV）和微凸块（Microbump）将 die 之间相连接，多层 DRAM die 再与最下层的 Base die 连接，然后通过凸块（Bump）与硅中介层（interposer）互联。HBM 与 GPU、CPU 或 ASIC 共同铺设在硅中介层上，通过 CoWoS 等 2.5D 封装工艺相互连接，硅中介层通过 Cu Bump 连接至封装基板（Package Substrate）上，最后封装基板再通过锡球与下方的 PCB 基板相连。根据全球半导体观察，目前 HBM 存储芯片的整体良率在 65% 左右，HBM 良率的高低主要受到其堆叠架构复杂性的影响，这涉及到多层次的内存结构和作为各层连接之用的直通 TSV 技术。这些复杂技术增加了制程缺陷的风险，可能导致良率低于设计较简单的内存产品。因此存储大厂纷纷加码 HBM 先进封装，提升 HBM 良率并降低功耗。此外，根据 CFM，SK 海力士拟将新一代 HBM4 堆栈直接放置在处理器上，通过 3D 堆叠的形式进一步提高 I/O 数量。
- **投资建议：**受益于 AI 算力需求的强劲推动，HBM 或将成为存储芯片市场复苏的主要驱动力，同时国内存储厂商也加速产品布局和技术突破，相关产业链上下游的厂商有望加速受益，建议关注：（1）存储：江波龙、德明利、兆易创新、东芯股份等；（2）先进封装：通富微电、长电科技、深科技等；（3）代理：香农芯创等；（4）设备和材料：北方华创、中微公司、赛腾股份、雅克科技等。
- **风险因素：**下游需求不及预期风险；国内技术研发不及预期风险；中美贸易摩擦加剧风险。

目录

美光开始量产 HBM3E，武汉新芯招标 HBM 封装产线.....	4
AI 浪潮驱动，HBM 先进封装、设备材料产业链环节持续受益.....	7
投资建议.....	9
风险因素.....	9

表目录

表 1: 美光 HBM3E 参数特征.....	4
表 2: 不同次代 HBM 规格比较.....	5

图目录

图 1: 三星 HBM3E 12H 产品.....	4
图 2: 三大原厂 HBM 产品开发进展.....	5
图 3: HBM 市场竞争格局.....	6
图 4: HBM 芯片工艺流程.....	7
图 5: 先进封装解决方案.....	7
图 6: HBM4 或将直接堆叠在逻辑芯片之上.....	8
图 7: SK 海力士独创 MR-MUF 技术.....	8

美光开始量产 HBM3E，武汉新芯招标 HBM 封装产线

近日，美光宣布已开始量产其 HBM3E 高带宽内存解决方案。英伟达 H200 Tensor Core GPU 将采用美光 8 层堆叠的 24GB 容量 HBM3E 内存，并于 2024 年第二季度开始出货。美光 HBM3E 引脚速率超过 9.2Gb/s，提供超过 1.2TB/s 的内存带宽。与 HBM3 相比，HBM3E 将数据传输速率和峰值内存带宽提高了 44%，这对于英伟达 H200 等需要大量带宽的处理器来说尤其重要。美光利用其 1 β (1-beta) 技术、先进的硅通孔 (TSV) 和其他实现差异化封装解决方案来生产 HBM3E 产品，有助于公司在数据中心级产品上提升技术竞争实力和市场占有率。

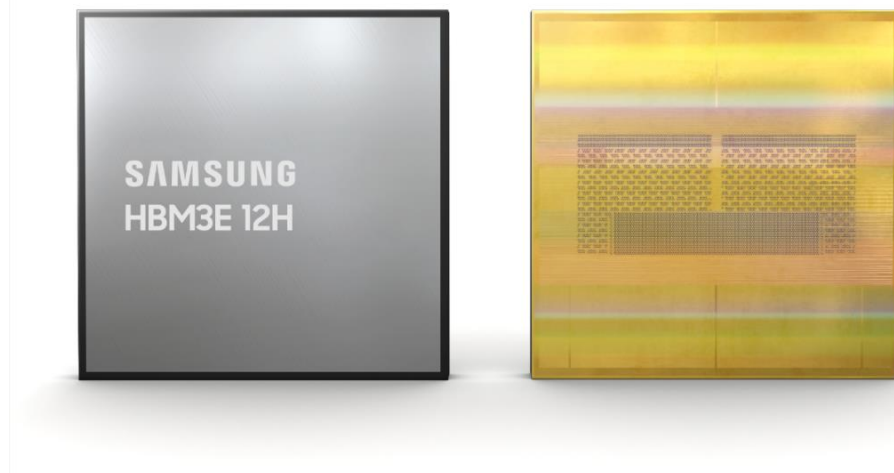
表 1: 美光 HBM3E 参数特征

特征	描述
容量	24GB, 36GB
技术	8H, 12H HBM3E
尺寸 (mm)	11mm x 11mm x 0.72mm
引脚速率	>9.2 Gb/s
突发传输长度	8
通道数	16
伪通道	2
带宽	>1.2 TB/s
工作温度	0° C ≤ T _{OPER} ≤ 105° C

资料来源: Micron, 信达证券研发中心

三星发布首款 36GB HBM3E 12H DRAM，目前为三星容量最大的 HBM。三星 HBM3E 12H 支持全天候最高带宽达 1280GB/s，产品容量也达到了 36GB。相比三星 8 层堆叠的 HBM3 8H，HBM3E 12H 在带宽和容量上大幅提升超过 50%。HBM3E 12H 采用了先进的热压非导电薄膜 (TC NCF) 技术，使得 12 层和 8 层堆叠产品的高度保持一致，以满足当前 HBM 封装的要求。相比 HBM3 8H，HBM3E 12H 搭载于人工智能应用后，公司预计人工智能训练平均速度可提升 34%，同时推理服务用户数量也可增加超过 11.5 倍。

图 1: 三星 HBM3E 12H 产品



资料来源: 三星半导体, 信达证券研发中心

HBM (High Bandwidth Memory) 是三星、AMD 和 SK 海力士发起的一种基于 3D 堆栈工艺的高性能 DRAM，通过增加带宽，扩展内存容量，让更大的模型，更多的参数留在离核心计算更近的地方，从而减少内存和存储解决方案带来的延迟，适用于高存储器带宽需求的应用场合。从技术角度看，HBM 使 DRAM 从传统 2D 转变为立体 3D，充分利用空间、缩小面积，契合半导体行业小型化、集成化的发展趋势。

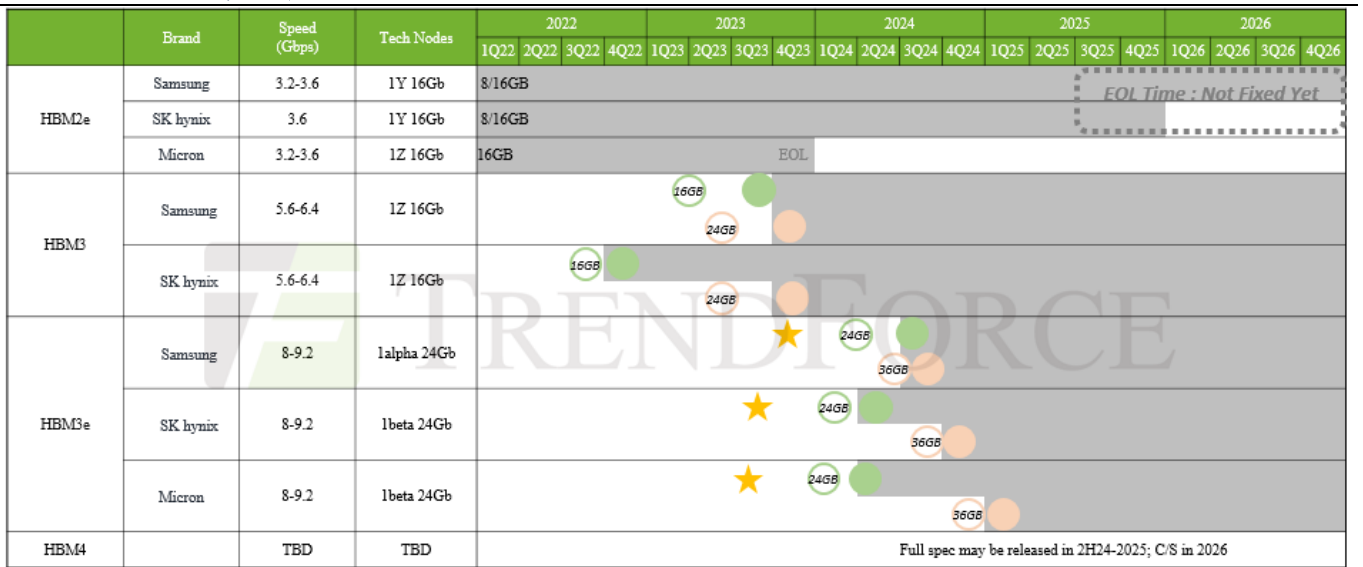
表 2: 不同次代 HBM 规格比较

次代	速率 (Gb/s)	带宽 (GB/s)	堆叠层数	Die 容量 (Gb)	HBM 容量 (GB)
HBM	1	128	8	16	16
HBM2	2	256	8	16	16
HBM2E	3.6	461	12	24	36
HBM3	6.4	819	16	32	64
HBM3E	9.6	1229	16	32	64

资料来源: Rambus, 信达证券研发中心

SK 海力士率先量产 HBM3E, HBM4 或于 2026 年推出。根据半导体行业观察, SK 海力士于 1 月中旬正式结束了 HBM3E 的开发工作, 并顺利完成英伟达历时半年的性能评估, 计划于 3 月开始大规模生产 HBM3E 产品, 这批 HBM3E 将用于英伟达下一代 Blackwell 系列的 AI 芯片旗舰产品 B100 上, 而英伟达则计划于 2024 年第二季度末或第三季度初推出该系列产品。根据 TrendForce, HBM4 预计规划于 2026 年推出, 或采取 HBM 堆栈在 SoC 主芯片之上的封装方式。

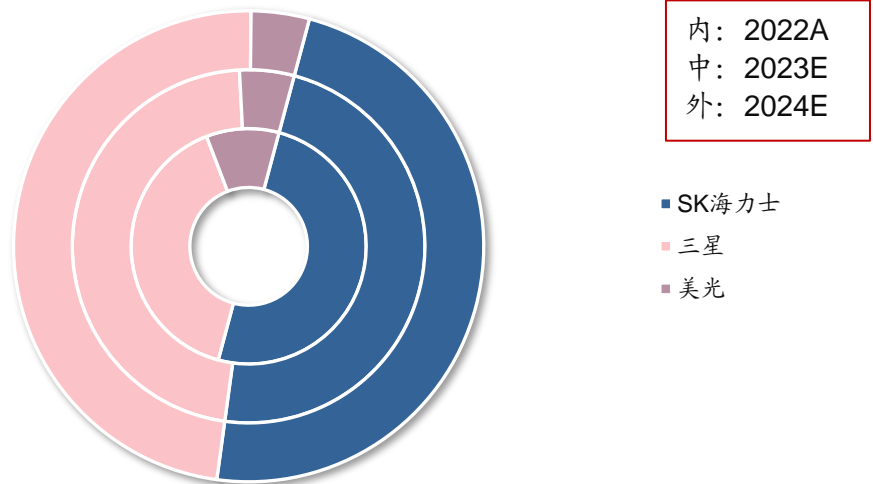
图 2: 三大原厂 HBM 产品开发进展



资料来源: TrendForce, 信达证券研发中心

SK海力士、三星和美光三分天下。根据 TrendForce 预测，2023 年 SK 海力士 HBM3 产品领先其他原厂，是 NVIDIA Server GPU 的主要供应商；三星则着重满足其他云端服务业者的订单，在客户加单下，2023 年与 SK 海力士的市占率差距大幅缩小，2023~2024 年两家业者 HBM 市占率预估相当，合计拥有 HBM 市场约 95% 的市占率，不过因客户组成略有不同，在不同季度的位元出货表现上或有先后。美光 2023 年专注开发 HBM3E 产品，相较两家韩厂大幅扩产的规划，TrendForce 预期 2023~2024 年美光的市占率会受排挤效应而略为下滑。

图 3：HBM 市场竞争格局



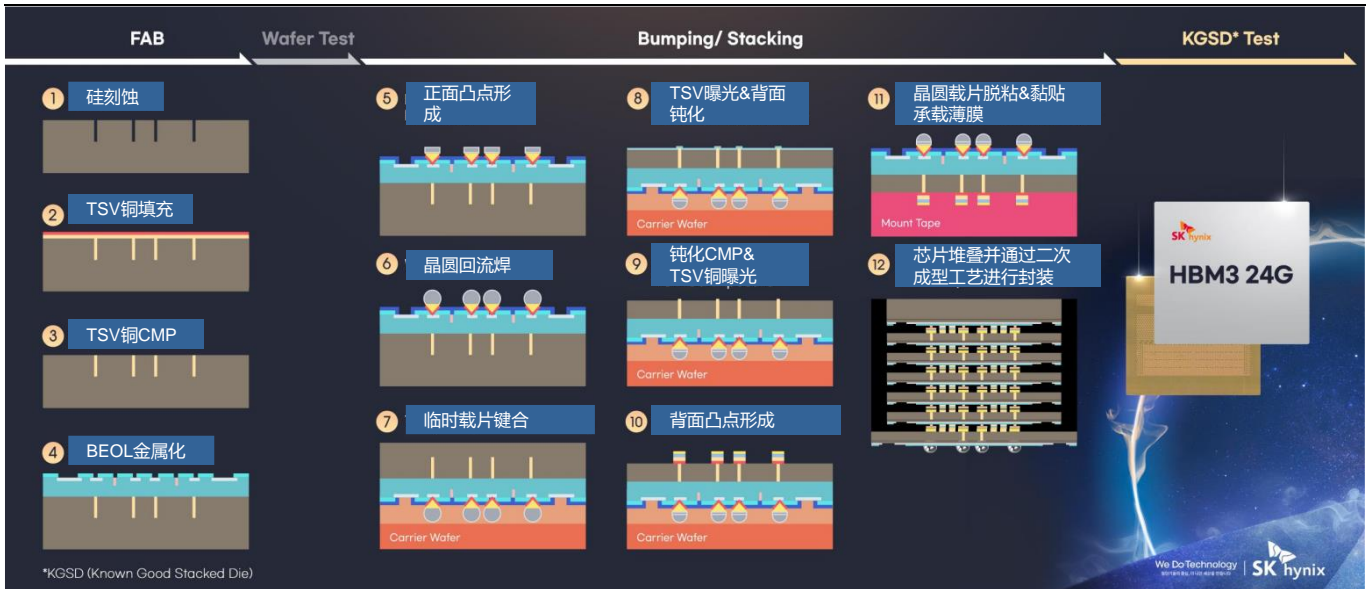
资料来源：TrendForce，信达证券研发中心

国内存储厂商入局 HBM 市场。根据采招网，近日，武汉新芯发布《高带宽存储芯粒先进封装技术研发和产线建设》招标项目，利用三维集成多晶圆堆叠技术，打造更高容量、更大带宽、更小功耗和更高生产效率的国产高带宽存储器（HBM）产品，推进多晶圆堆叠工艺产业化，新增生产设备约 17 台/套，拟实现月产出能力≥3000 片（12 英寸）。我们认为，国内存储厂商在 HBM 技术上的加速突破，有望在 AI 大浪潮的需求下提升竞争实力，相关产业链也或将受益。

AI 浪潮驱动，HBM 先进封装、设备材料产业链环节持续受益

HBM 制造工艺包括 TSV、Bumping 和堆叠等工艺环节。HBM 是由多个 DRAM die 堆叠而成，利用硅通孔（TSV）和微凸块（Microbump）将 die 之间相连接，多层 DRAM die 再与最下层的 Base die 连接，然后通过凸块（Bump）与硅中阶层（interposer）互联。HBM 与 GPU、CPU 或 ASIC 共同铺设在硅中阶层上，通过 CoWoS 等 2.5D 封装工艺相互连接，硅中阶层通过 Cu Bump 连接至封装基板（Package Substrate）上，最后封装基板再通过锡球与下方的 PCB 基板相连。

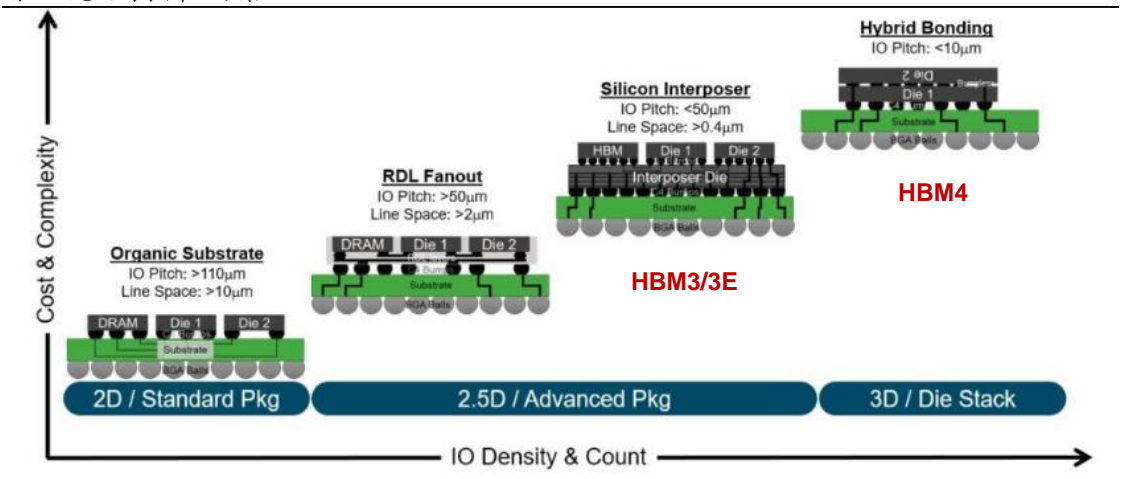
图 4: HBM 芯片工艺流程



资料来源: SK 海力士, 信达证券研发中心

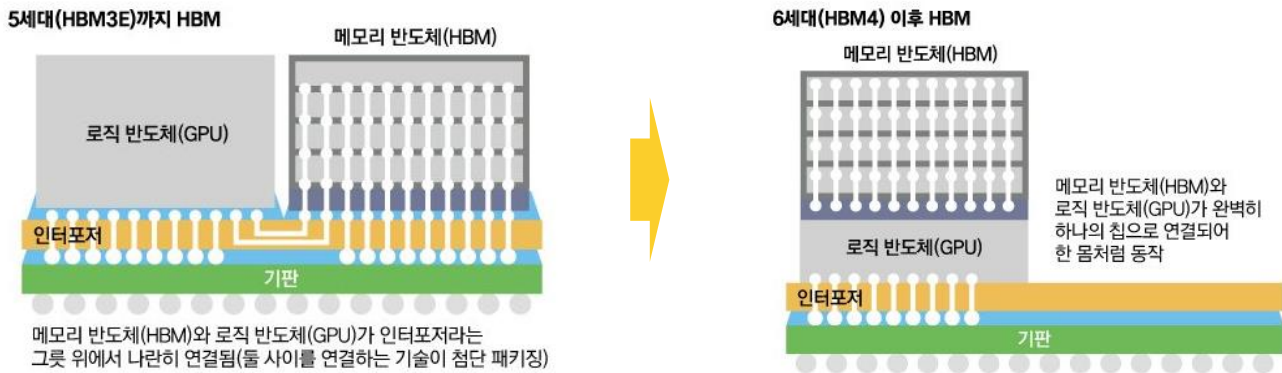
存储大厂高度聚焦先进封装，HBM4 或将直接放置在处理器上。根据全球半导体观察，目前 HBM 存储芯片的整体良率在 65% 左右，HBM 良率的高低主要受到其堆叠架构复杂性的影响，这涉及到多层次的内存结构和作为各层连接之用的直通 TSV 技术。这些复杂技术增加了制程缺陷的风险，可能导致良率低于设计较简单的内存产品。因此存储大厂纷纷加码 HBM 先进封装，提升 HBM 良率并降低功耗。此外，根据 CFM，SK 海力士拟将新一代 HBM4 堆栈直接放置在处理器上，通过 3D 堆叠的形式进一步提高 I/O 数量。

图 5: 先进封装解决方案



资料来源: Synopsys, 信达证券研发中心

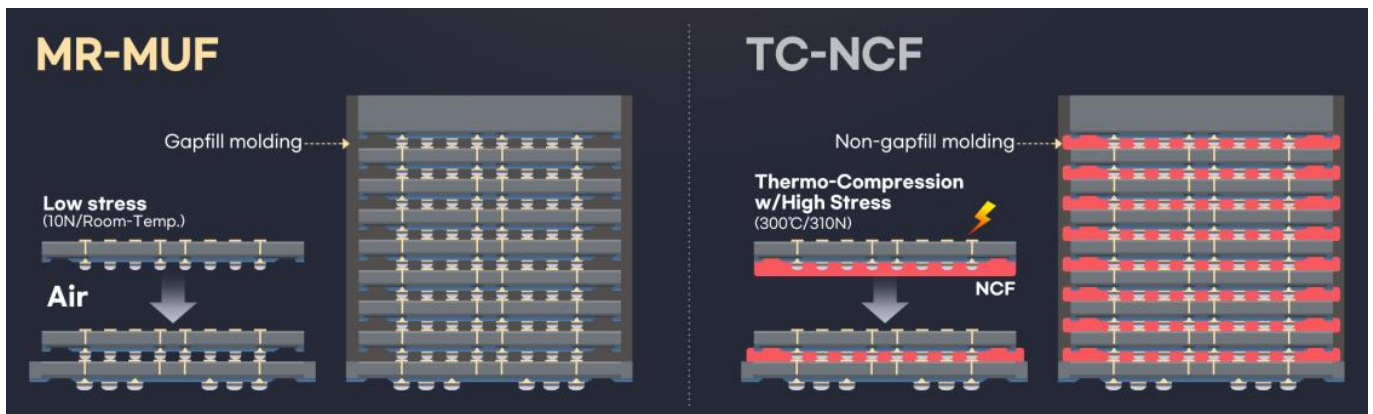
图 6: HBM4 或将直接堆叠在逻辑芯片之上



资料来源: 闪存市场 (CFM), 信达证券研发中心

SK 海力士凭借 MR-MUF 技术占据 HBM 市场份额, 三星考虑将 MUF 用于 3DS RDIMM。 SK 海力士在 HBM2E 和 HBM3 的生产中引入了大规模回流成型底部填充 (MR-MUF) 工艺, 提高了 HBM 超过 10 万个微凸块互连的质量, 此外, 该工艺充分增加了热虚拟凸块的数量, 同时由于采用了高导热率的模制底部填充 (MUF) 材料, 因此散热性能更加出色。根据全球半导体观察, 三星认为 MUF 非常适合 3DS RDIMM, 而目前 3DS RDIMM 使用 TSV 技术制造, 主要应用于服务器上, 三星正在考虑在其下一代 DRAM 中 MUF 技术。若三星也导入 MUF, 相关制造材料或将持续受益。

图 7: SK 海力士独创 MR-MUF 技术



资料来源: SK 海力士, 信达证券研发中心

投资建议

受益于 AI 算力需求的强劲推动，HBM 或将成为存储芯片市场复苏的主要驱动力，同时国内存储厂商也加速产品布局和技术突破，相关产业链上下游的厂商有望加速受益，建议关注：

- 存储：江波龙、德明利、兆易创新、东芯股份等；
- 先进封装：通富微电、长电科技、深科技等；
- 代理：香农芯创等；
- 设备和材料：北方华创、中微公司、赛腾股份、雅克科技等。

风险因素

下游需求不及预期风险；

国内技术研发不及预期风险；

中美贸易摩擦加剧风险。

研究团队简介

莫文宇，电子行业分析师，S1500522090001。毕业于美国佛罗里达大学，电子工程硕士，2012-2022年就职于长江证券研究所，2022年入职信达证券研发中心，任副所长、电子行业首席分析师。

韩宇杰，电子行业研究员。华中科技大学计算机科学与技术学士、香港中文大学硕士。研究方向为半导体设备、半导体材料、集成电路设计。

郭一江，电子行业研究员。本科兰州大学，研究生就读于北京大学化学专业。2020年8月入职华创证券电子组，后于2022年11月加入信达证券电子组，研究方向为光学、消费电子、汽车电子等。

吴加正，电子行业研究员。复旦大学工学学士、理学博士，德国慕尼黑工业大学2年访问研究经验。2020年9月入职上海微电子装备(集团)股份有限公司，任光刻机系统工程师，于2022年12月加入信达证券电子组，研究方向为精密电子仪器、半导体设备及零部件、半导体工艺等。

王义夫，电子行业研究员。西南财经大学金融学士，复旦大学金融硕士，2023年加入信达证券电子组，研究方向为存储芯片、模拟芯片等。

李星全，电子行业研究员。哈尔滨工业大学学士，北京大学硕士。2023年加入信达证券电子组，研究方向为服务器、PCB、消费电子等。

分析师声明

负责本报告全部或部分内容的每一位分析师在此申明，本人具有证券投资咨询执业资格，并在中国证券业协会注册登记为证券分析师，以勤勉的职业态度，独立、客观地出具本报告；本报告所表述的所有观点准确反映了分析师本人的研究观点；本人薪酬的任何组成部分不曾与，不与，也将不会与本报告中的具体分析意见或观点直接或间接相关。

免责声明

信达证券股份有限公司（以下简称“信达证券”）具有中国证监会批复的证券投资咨询业务资格。本报告由信达证券制作并发布。

本报告是针对与信达证券签署服务协议的签约客户的专属研究产品，为该类客户进行投资决策时提供辅助和参考，双方对权利与义务均有严格约定。本报告仅提供给上述特定客户，并不面向公众发布。信达证券不会因接收人收到本报告而视其为本公司的当然客户。客户应当认识到有关本报告的电话、短信、邮件提示仅为研究观点的简要沟通，对本报告的参考使用须以本报告的完整版本为准。

本报告是基于信达证券认为可靠的已公开信息编制，但信达证券不保证所载信息的准确性和完整性。本报告所载的意见、评估及预测仅为本报告最初出具日的观点和判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会出现不同程度的波动，涉及证券或投资标的的历史表现不应作为日后表现的保证。在不同时期，或因使用不同假设和标准，采用不同观点和分析方法，致使信达证券发出与本报告所载意见、评估及预测不一致的研究报告，对此信达证券可不发出特别通知。

在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议，也没有考虑到客户特殊的投资目标、财务状况或需求。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，若有必要应寻求专家意见。本报告所载的资料、工具、意见及推测仅供参考，并非作为或被视为出售或购买证券或其他投资标的的邀请或向人做出邀请。

在法律允许的情况下，信达证券或其关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并可能会为这些公司正在提供或争取提供投资银行业务服务。

本报告版权仅为信达证券所有。未经信达证券书面同意，任何机构和个人不得以任何形式翻版、复制、发布、转发或引用本报告的任何部分。若信达证券以外的机构向其客户发放本报告，则由该机构独自为此发送行为负责，信达证券对此等行为不承担任何责任。本报告同时不构成信达证券向发送本报告的机构之客户提供的投资建议。

如未经信达证券授权，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。信达证券将保留随时追究其法律责任的权利。

评级说明

投资建议的比较标准	股票投资评级	行业投资评级
本报告采用的基准指数：沪深300指数（以下简称基准）； 时间段：报告发布之日起6个月内。	买入 ：股价相对强于基准20%以上；	看好 ：行业指数超越基准；
	增持 ：股价相对强于基准5%~20%；	中性 ：行业指数与基准基本持平；
	持有 ：股价相对基准波动在±5%之间；	看淡 ：行业指数弱于基准。
	卖出 ：股价相对弱于基准5%以下。	

风险提示

证券市场是一个风险无时不在的市场。投资者在进行证券交易时存在赢利的可能，也存在亏损的风险。建议投资者应当充分深入地了解证券市场蕴含的各项风险并谨慎行事。

本报告中所述证券不一定能在所有的国家和地区向所有类型的投资者销售，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顾问的意见。在任何情况下，信达证券不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任，投资者需自行承担风险。