仅供内部参考，请勿外传

国海证券 SEALAND SECURITIES

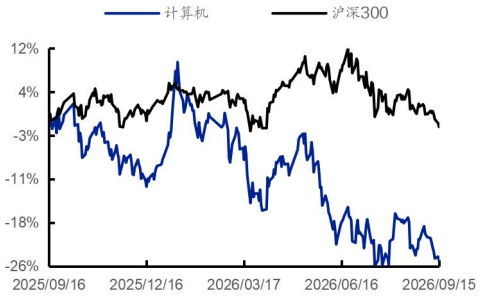
2026 年 09 月 16 日 | 行业研究 | 评级：推荐(维持)

研究所：
证券分析师：刘熹 S0350523040001
liux10@ghzq.com.cn
联系人：谢婧茹 S0350125070015
xiejr01@ghzq.com.cn

# 国产算力的 HBM 供应何解？

## ——计算机行业专题研究

最近一年走势




| 行业相对表现 | | | 2026/09/15 |
|---|---|---|---|
| 表现 | 1M | 3M | 12M |
| 计算机 | -8.1% | -7.5% | -23.1% |
| 沪深 300 | -4.6% | -9.0% | -1.8% |

相关报告

《AI ASIC：推理场景优势显著，与 GPU 共筑算力新格局——计算机国产算力专题系列（推荐）*计算机*刘熹》——2026-08-11

《计算机行业专题研究：北美 CSP 继续上调资本开支指引，从 AI ROI 角度分析现金流中长期影响（推荐）*计算机*刘熹》——2026-08-05

《计算机行业专题研究：从规模增长到价值兑现，头部模型厂进入新盈利阶段（推荐）*计算机*刘熹》——2026-06-17

《计算机事件点评：华为韬(τ)定律：国产算力迎新解决方案（推荐）*计算机*刘熹》——2026-05-27

《计算机“算力租赁+国产算力”系列（5）：中国 AI 核心矛盾：算力交付（推荐）*计算机*刘熹》——2026-05-18

## 投资要点：

本篇报告解决了以下核心问题：国产 AI 算力扩张下的 HBM 供需缺口现状及未来缓解趋势分析。

■ **HBM 是 AI 芯片关键内存组件，具备高带宽/能效比/低延时特性**

**HBM 是基于 2.5/3D 先进封装技术实现的大容量、高带宽 DDR 组合阵列，由 C-die 核心芯片与 B-die 基础芯片两部分构成。**其内部结构与传统内存不同，并非将内存芯片并排布置，而是利用 TSV 硅通孔和微凸块将多层 DRAM 裸片垂直堆叠键合，再与 Base Die 连接，最后通过凸块与硅中阶层互联。**凭借这一架构，HBM 具备超高带宽、高能效比和低延时三大核心优势：**HBM3E 单堆栈带宽已达 1.2TB/s，HBM4 预计跃升至 2.0TB/s；TSV 技术使能耗较 GDDR5 降低约 30-50%；3D 堆叠设计使 PCB 占用面积仅为传统方案的 1/3，超短互连距离则显著降低了信号传输延时。

**HBM 在 AI 芯片中扮演核心配套内存组件角色，其带宽与容量直接影响大模型的推理与训练效率。**推理场景中若带宽不足，GPU 将长时间处于等待状态。训练场景中，前向与反向传播所需数据均通过 HBM 流转，HBM 的带宽与容量直接决定大规模模型的训练效率与可扩展性。以英伟达为例，从 H200 到最新的 Rubin GPU，单卡 HBM 配置在容量与型号上持续升级，Rubin GPU 已配备 288GB HBM4；**国产 GPU 厂商华为、阿里、寒武纪等当前单卡 HBM 配置以 HBM3 等为主，容量约为 144GB 至 196GB。**

■ **国内 CSP+主权 AI 投资景气度走高，对应较大 HBM 需求量**

**国内 CSP 算力采购扩张明显。**2026Q2，腾讯资本开支达 527.84 亿元，同比增长 176%，上半年累计 847.20 亿元，已超 2025 年全年 792 亿元规模；阿里资本支出达 676.78 亿元，同比增长 75%，三年 3800 亿元 AI 基建计划截至 6 月季度末已累计投入 1900 亿元；字节正评估将 2026 年资本支出推至最高 700 亿美元（约 4700 亿元人民币），较 2025 年翻倍有余，此前 2026 年 AI 资本开支亦上修至逾 2000 亿元，新增投资集中投向数据中心扩容与 AI 基础设施。

仅供内部参考，请勿外传



**国内主权 AI 投资景气度正处于投资规模量级跃升、国产化导向明确的阶段。**政策端，工信部印发《信息通信行业发展“十五五”规划》，明确到 2030 年智能算力规模达 9800 EFLOPS，信息基础设施累计投资 3.8 万亿元。基于此，我们预测 2027 年寒武纪、海光信息、华为及其他出货量分别为 55、45、290、260 万颗，对应 2027 年等效 24GB HBM3/3E 需求量为 330、270、1740、1560 万颗，合计需求量达 3900 万颗。

## ■ 2027 年 HBM 供给紧张态势延续，国产 HBM 有望予以支撑

**HBM 核心供应商 SK 海力士、三星与美光已分配完 60-70%的 2027 年 HBM 产能，且主要转向 HBM4。**扩产方面，三星计划到 2026 年末将 HBM 月产能提升至 25 万片晶圆（以 12 英寸计），较 2025 年末约 17 万片增长约 47%，扩产主要为 HBM4 订单做准备；SK 海力士 HBM 产能 2025 年底约为每月 15 万片，至 2026 年底增长至 20 万片。产品与客户结构上，三星三大客户为英伟达、AMD、博通，计划于 2026 年下半年对上述客户放量供货，且 2026 年下半年 HBM4 将占其 HBM 总收入的 60%以上；SK 海力士 2026 年 HBM 出货量中约三分之二分配给英伟达，计划上半年主力供应 HBM3E，下半年起扩大 HBM4 供应量，并已与十家大客户签订长期供货协议。

国内方面，长鑫在合肥运营两座 12 英寸 DRAM 晶圆厂、在北京运营一座 12 英寸 DRAM 晶圆厂，合计月产能约 30 万片；上海新工厂计划 2027 年起投产，达产后整体 DRAM 产能将为现有水平的两倍以上。**其 DRAM 产品覆盖 DDR 与 LPDDR 等，并已在技术上具备 HBM3 量产能力。**如果长鑫 2027 年可出货 300 万颗 24GB 等效 HBM3，按照单卡 96-144G 规格测算，共可支撑 50-75 万颗国产 GPU 需求，有望对包括寒武纪、海光等在内的国产 GPU 生产提供保障。

- **行业评级及投资策略：**当前 HBM 供给紧张，**但从中期看，随着国产 HBM 供应逐步落地，这一供需缺口有望得到缓解。与此同时，AI 算力需求景气度仍在持续走高，行业需求端具备较强支撑。**在供给瓶颈边际改善、需求持续向好的背景下，我们维持对计算机行业的“推荐”评级。
- **相关公司：**1）CPU：海光信息、中国长城（飞腾信息）；2）AI 芯片：寒武纪、沐曦股份、壁仞科技、天数智芯、燧原科技、芯原股份；3）连接：澜起科技、盛科通信、万通发展、曦智科技；4）超节点：浪潮信息、华勤技术、联想集团、中科曙光、紫光股份、锐捷网络、工业富联；5）液冷：鼎通科技、飞荣达、申菱环境；6）电源：中国长城、欧陆通、麦格米特、中恒电气；7）算力租赁：协创数据、利通电子、宏景科技、盛视科技、智微智能、润建股份、赛意信息；8）垂类 AI 应用：中控技术、科大讯飞、鼎捷数智、万兴科技等。

仅供内部参考，请勿外传



- **风险提示**：宏观经济影响下游需求、AI 模型发展不及预期、市场竞争加剧、中美博弈加剧、供应链风险、扩产不及预期、AI 商业化兑现不及预期、国产 HBM 量产进度及良率不及预期、GPU 出货量假设偏差、算力资本开支不及预期、政策规划落地不及预期。

仅供内部参考，请勿外传



# 内容目录

1、 HBM 是 AI 芯片的关键内存组件，具备高带宽/能效比/低延时特性 ..... 6
2、 国内 CSP+主权 AI 投资景气度走高，对应较大 HBM 需求量 ..... 9
3、 2027 年 HBM 供给紧张态势延续，国产 HBM 有望予以支撑 ..... 11
4、 投资建议 ..... 13
5、 风险提示 ..... 13

仅供内部参考，请勿外传



# 图表目录

图 1: HBM 由 C-die 和 B-die 两部分构成 ..... 6
图 2: HBM 多层堆叠结构 ..... 7
图 3: HBM 对比传统内存的超高带宽、高能效比以及低延时特性 ..... 7
图 4: 英伟达及国产 GPU HBM 配置 ..... 8
图 5: 腾讯/阿里季度资本支出（亿元） ..... 9
图 6: 《信息通信行业发展“十五五”规划》指出到 2030 年智能算力规模达到 9800 EFLOPS ..... 10
图 7: 国产 GPU 出货量预测 ..... 11
图 8: HBM 迭代路径 ..... 12
图 9: 长鑫 DRAM 产能爬坡预测曲线 ..... 12

仅供内部参考，请勿外传



**本篇报告解决了以下核心问题：**

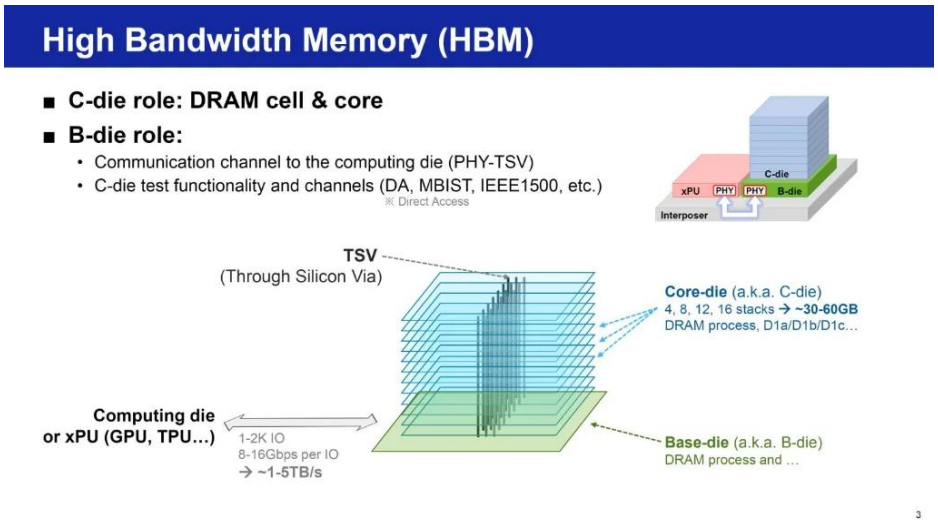
国产 AI 算力扩张下的 HBM 供需缺口现状及未来缓解趋势分析。

# 1、HBM 是 AI 芯片的关键内存组件，具备高带宽/能效比/低延时特性

**HBM 是 AI 芯片中的关键内存组件，由 C-die 和 B-die 两部分构成，同时具备多层堆叠结构。**

HBM，即高带宽内存，本质上是基于 2.5/3D 先进封装技术来实现的大容量、高带宽的 DDR 组合阵列；主要由两部分芯片构成：1）C-die 核心芯片：包含 DRAM 存储单元，可垂直堆叠最多 16 层。2）B-die 基础芯片：位于整个堆叠结构底部，负责处理各类 DRAM 控制功能，并通过 PHY 物理接口层与 GPU 等计算芯片通信。

**图 1：HBM 由 C-die 和 B-die 两部分构成**



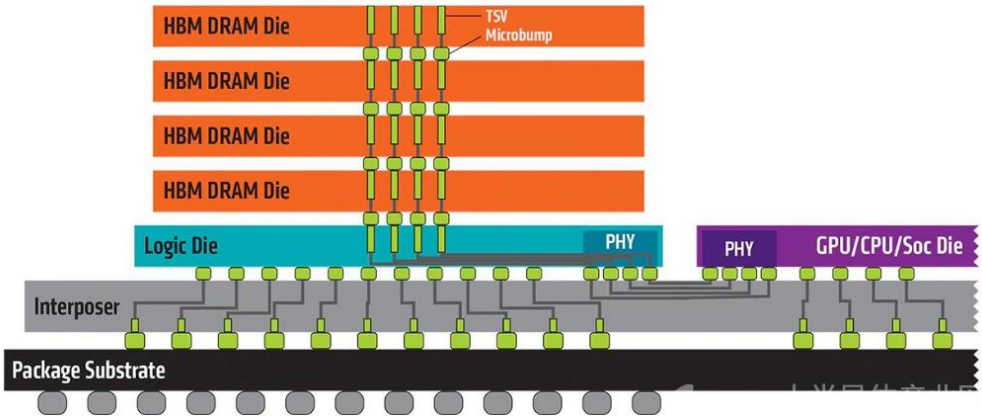

资料来源：华尔街见闻

**其内部结构与传统内存不同，**HBM 不是将所有内存芯片并排布置在平板上，**而是将多个 DRAM 像多层建筑一样垂直堆叠，主要利用先进工艺 TSV（硅通孔）和微凸块（Micro bump）将裸片进行 TCB 堆叠键合而相连接；**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。

仅供内部参考，请勿外传



**图 2：HBM 多层堆叠结构**




资料来源：SEMI

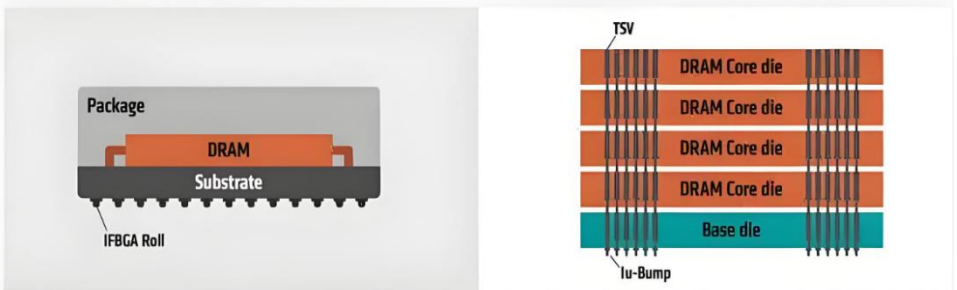
**核心优势方面，HBM 与传统内存相比具备超高带宽、高能效比以及低延时特性：**

- HBM 的带宽能力远超传统 DRAM，目前 HBM3E 的单堆栈带宽已达 1.2TB/s，预计 HBM4 将进一步跃升至 2.0TB/s。**高带宽能力使其能解决 AI 训练中的内存瓶颈问题，让 GPU 高效处理数十亿甚至数万亿参数的模型。**

- 能效方面，TSV 技术缩短了数据传输距离，使 HBM 的能耗比较 GDDR5 降低约 30-50%。同时垂直堆叠结构减少信号传输路径，降低了驱动互连所需的功率，大幅提升能效。而且在空间尺寸上，3D 堆叠设计使内存模块尺寸大幅减小，HBM 的 PCB 占用面积仅为传统方案的 1/3。

- 在低延时特性上，HBM 的超短互连距离和直接垂直连接降低了信号传输延时，使处理器能够更快访问数据，**在 AI 推理等实时性要求高的场景中，这种低延时特性尤为关键。**

**图 3：HBM 对比传统内存的超高带宽、高能效比以及低延时特性**




| GDDR5 | Per Package | HBM |
|---|---|---|
| 32-bit | Bus Width | 1024-bit |
| Up to 1750MHz (7GBps) | Clock Speed | Up to 500MHz (1GBps) |
| Up to 28GB/s per chip | Bandwidth | >100GB/s per stack |
| 1.5V | Voltage | 1.3V |

资料来源：SEMI

仅供内部参考，请勿外传



**因此 HBM 在 AI 芯片中扮演着核心配套内存组件的角色，尤其在高性能 GPU 等 AI 芯片中，其带宽与容量直接影响大模型的推理与训练效率：**

**在推理场景中，**模型权重需要驻留于 HBM 中，以便 GPU 即时读取。同时除权重外，HBM 还存储 KV 缓存。每次生成下一个 token 时，GPU 首先从 HBM 读取权重，同时检索整个 KV 缓存，以便在自注意力阶段将新 token 与对话历史进行比对。计算完成后，GPU 将新生成的键和值追加回 HBM，从而持续扩大缓存。**这一过程对带宽要求极高，若内存带宽无法达到每秒数百 GB 级别，GPU 将长时间处于等待内存状态，而非执行计算。**

**在训练场景中，**GPU 进行前向与反向传播所需的所有数据均通过 HBM 流转。**前向传播时，**模型权重存储于 HBM，各层在处理一批数据时可快速读取；每一层处理完毕后，会将中间激活值写入 HBM 以供后续使用。**反向传播时，**GPU 重新访问存储的激活值与权重，从 HBM 读取并计算梯度。HBM 的带宽与容量直接影响大规模模型的训练效率与可扩展性。

**以英伟达为例，从 H200 到最新的 Rubin GPU，其单卡 HBM 配置在容量与型号上持续升级，Rubin GPU 已配备 288GB HBM4。**国产 GPU 厂商如华为、阿里、寒武纪等当前单卡 HBM 配置规格也较高，以 HBM3 等为主，配置容量约为 144GB 至 196GB。

**图 4：英伟达及国产 GPU HBM 配置**

| 厂商 | 产品型号 | FP16 算力 | HBM 配置 |
|---|---|---|---|
| 英伟达 | H200 | 1000/2000T | 141GB HBM3E |
| | B100 | 1800/3500T | 192GB HBM3E |
| | B200 | 2250/4500T | 192GB HBM3E |
| | B300 | 2250/4500T | 288GB HBM3E |
| | Rubin | 50 P（FP4 算力） | 288GB HBM4 |
| 华为 | 昇腾 950DT | / | 144GB HBM（HiZQ 2.0） |
| 阿里 | 真武 M890 | / | 144GB HBM |
| 寒武纪 | MLU690 | 700T+ | 196GB HBM3 |

资料来源：九章智算云、卓普云 AI Droplet 公众号、新智元、电子工程专辑、平头哥官网、壹壹捌芯城、国海证券研究所

仅供内部参考，请勿外传



## 2、国内 CSP+主权 AI 投资景气度走高，对应较大 HBM 需求量

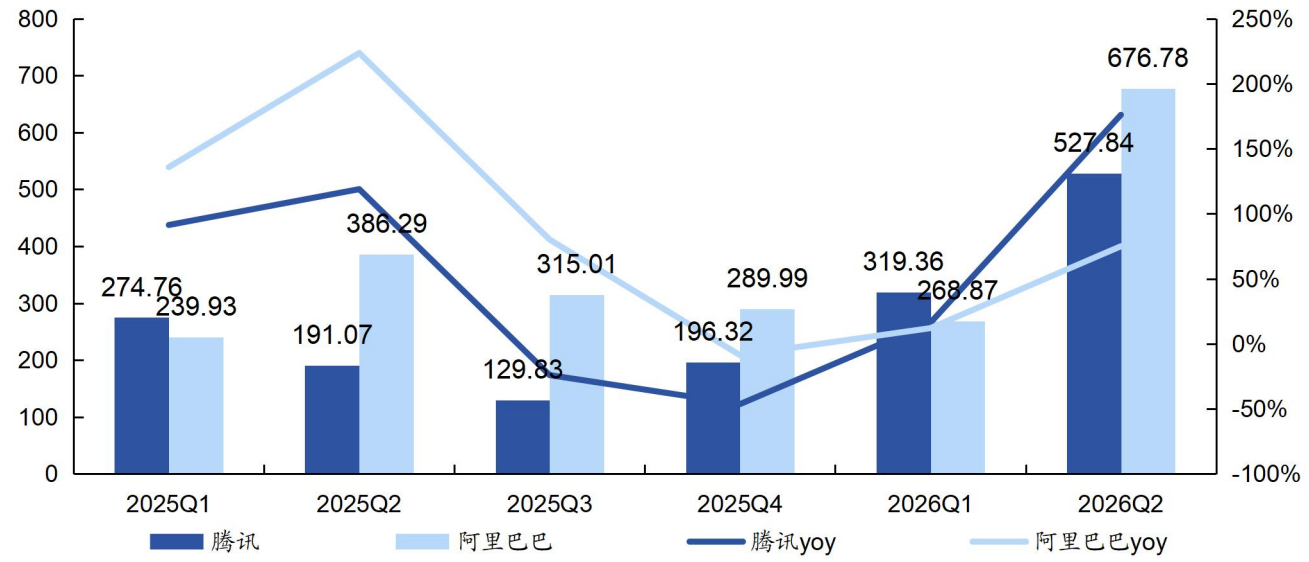
**需求侧，当前国内 CSP+主权 AI 需求景气度持续走高：**

**国内 CSP 算力采购扩张明显。**2026Q2，腾讯资本开支达到 527.84 亿元，同比增长 176%、环比增长 65%，**上半年资本开支累计达到 847.20 亿元，已经超过 2025 年全年 792 亿元的规模。**腾讯表示，在基础设施层，公司大幅增加了算力采购，将助力公司把应用及模型的使用转化为收入。

**阿里方面，2026Q2 资本支出达 676.78 亿元，同比增长 75%。**阿里 CEO 吴泳铭指出此前公布的三年 3800 亿元 AI 基建投资计划，到 6 月季度末已累计投入 1900 亿元。同时二季度，阿里云外部商业化收入同比增长了 45%，创 22 个季度以来增速新高。而且二季度，AI 相关产品收入 123.76 亿元，连续第 12 个季度实现三位数同比增长。

同时据新智核公众号，**字节正评估将 2026 年资本支出推至最高 700 亿美元（约 4700 亿元人民币），较 2025 年翻倍有余；**此前其 2026 年 AI 资本开支也已上修至逾 2000 亿元人民币，较早前计划上调至少 25%。**新增投资集中投向数据中心扩容与 AI 基础设施。**

**图 5：腾讯/阿里季度资本支出（亿元）**




资料来源：Bloomberg、每日经济新闻公众号、21 世纪经济报道公众号、财联社公众号、说财猫公众号、国海证券研究所

**国内主权 AI 投资的景气度正处于投资规模量级跃升、国产化导向明确的阶段。**工信部印发的《信息通信业“十五五”规划》指出：**到 2030 年智能算力规模达到 9800 EFLOPS，信息基础设施累计投资 3.8 万亿元，并明确“有序部署万卡、十万卡及以上智算集群”和“加大力度适配国产算力芯片”。**

仅供内部参考，请勿外传



同时，2026 年 6 月，发改委指出算力是继热力、电力之后的新型基础资源，算力网是数字经济的核心底座。目前，全国一体化算力网建设正在加速推进，**依托“东数西算”工程布局了 8 大国家算力枢纽和 10 大数据中心集群。据界面新闻，截至 2026 年 3 月底，我国已建成智能算力规模 188.2 万 P**，是 2025 年同期的 2.5 倍，预计还将保持高速增长态势。

**图 6：《信息通信行业发展“十五五”规划》指出到 2030 年智能算力规模达到 9800 EFLOPS**

专栏 1 “十五五”信息通信行业发展主要指标

| 序号 | 指标名称 | 2025 年 | 2030 年 | 年均/累计 |
|---|---|---|---|---|
| 1 | 信息通信行业收入（万亿元） | 3.8 | 4.1 | 1.5% |
| 2 | 信息基础设施累计投资（万亿元） | 3.7* | 3.8* | 〔0.1〕 |
| 3 | 电信业务总量增速（%） | 9 | — | 7 |
| 4 | 研发投入强度（%） | 4.7 | 4.8 | 〔0.1〕 |
| 5 | 单位电信业务总量二氧化碳排放量下降幅度（%） | 6 | 17* | — |
| 6 | 新建大型超大型算力设施运行电能利用效率（PUE） | 1.25 | <1.2 | 〔>0.05〕 |
| 7 | 每万人拥有 5G（含 5G-A）基站数（个） | 34 | 50 | 〔16〕 |
| 8 | 50G-PON 端口数（万个） | 2 | 100 | 〔98〕 |
| 9 | 智能算力规模（每秒百亿亿次浮点运算） | 1590 | 9800 | 〔8210〕 |
| 10 | 先进存力规模（EB） | 540 | 1700 | 〔1160〕 |
| 11 | 5G（含 5G-A）用户普及率（%） | 85.7 | 95 | 〔9.3〕 |
| 12 | 千兆以上宽带用户数（亿户） | 2.4 | 3.2 | 〔0.8〕 |
| 13 | 通信网络终端连接数（亿个） | 54 | 65 | 〔11〕 |

注：①〔〕内为 5 年累计变化数。②带*的为连续 5 年累计值。③信息通信行业收入、信息基础设施累计投资、研发投入强度包括基础电信企业及规模以上互联网企业数据。

资料来源：21 世纪经济报道

基于此，我们预测，寒武纪、海光信息、华为以及其他在 2027 年的出货量分别为 55、45、290、260 万颗，对应 2027 年等效 24GB HBM3/3E 需求量为 330、270、1740、1560 万颗，合计需求量达 3900 万颗。



仅供内部参考，请勿外传



图 7：国产 GPU 出货量预测

| 国产 GPU 厂商 | 2027 年出货量预测（万颗） | 2027 年 HBM 需求量预测（万颗/每颗 24GB 等效容量） |
|---|---|---|
| 寒武纪 | 55 | 330 |
| 海光信息 | 45 | 270 |
| 华为 | 290 | 1740 |
| 其他 | 260 | 1560 |
| 合计 | 650 | 3900 |

资料来源：国海证券研究所

## 3、2027 年 HBM 供给紧张态势延续，国产 HBM 有望予以支撑

**HBM 核心供应商 SK 海力士、三星与美光已分配完 60-70%的 2027 年 HBM 产能，且主要转向 HBM4：**

据电子工程专辑，三星计划到 2026 年末，将 HBM 的月产能提升至 25 万片晶圆（以 12 英寸晶圆计算），相较于 2025 年末约每月 17 万片的产能，增幅高达约 47%，扩产主要为 HBM4 订单做准备。

对于 SK 海力士，其 HBM 产能在 2025 年底约为每月 15 万片，占其 DRAM 总产能 54.5 万片的约 28%。DRAM 总产能预计到 2026 年底增至每月 60 万片，同比增加 5.5 万片，其中 90%以上的增量将用于 HBM。即 SK 海力士 HBM 月产能至 2026 年底将增长至 20 万片。

**产品及客户结构方面，**三星三大客户为英伟达、AMD、博通，其计划于 2026 年下半年对上述三大客户放量供货，且 2026 年下半年 HBM4 将占其 HBM 总收入的 60%以上。SK 海力士 2026 年 HBM 出货量中约三分之二分配给英伟达，且计划上半年主力供应 HBM3E，下半年起扩大 HBM4 供应量，并与十家大客户签订长期供货协议。

**内存芯片供应紧张态势预计将持续至 2027 年，**据科创板日报援引财联社，三星、SK 海力士和美光已对 2027 年 DRAM 和 HBM 产能进行分配。客户最终获得的配货量为其初始请求的 60%至 70%。

仅供内部参考，请勿外传



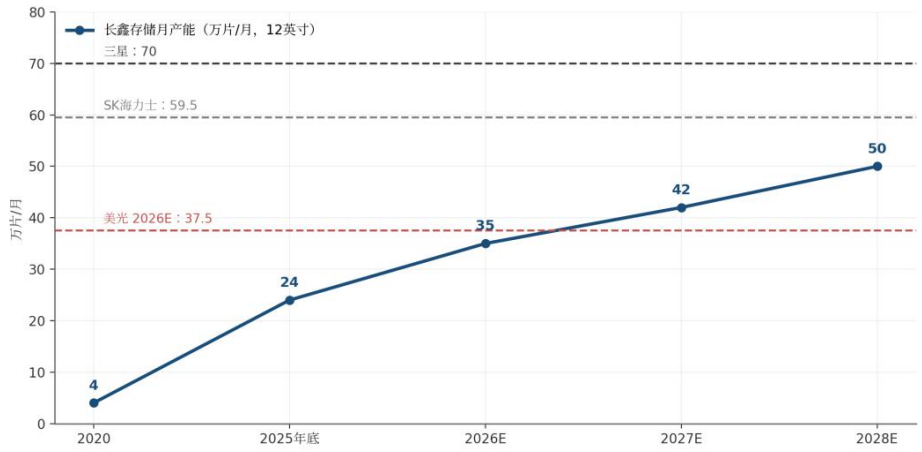
**图 8：HBM 迭代路径**

| | HBM | HBM2 | HBM2E | HBM3 | HBM3E | HBM4 |
|---|---|---|---|---|---|---|
| 核心供应商 | SK海力士 | SK海力士、三星 | SK海力士、三星、美光 | SK海力士、三星 | SK海力士、三星、美光 | SK海力士、三星、美光 |
| 首款产品发布年份 | 2014年 | 2018年 | 2020年 | 2022-2023年 | 2024年 | 2025-2026年 |
| 芯片密度 | 2Gb | 8-16Gb | 16Gb | 16Gb | 24Gb | 24-32Gb |
| 工艺制程 | 2x | 2y/2z | 1y/1z | 1z | 1a/1b/1β | 1b/1β/1c/1γ |
| 内存带宽 | 128GB/s | 307GB/s | 460GB/s | 819GB/s | 1.2TB/s | ≥2TB/s |
| 堆叠高度 | 4层 | 4-8层 | 4-8层 | 8-12层 | 8-12层 | 12-16层 |
| 主要封装技术 | TSV&Microbumps | TSV&Microbumps | TSV&Microbumps | TSV&Microbumps | TSV&Microbumps | Cu-Cu Hybrid Bonding for 16Hi |
| I/O速率 | 1Gbps | 2-2.4Gbps | 3.2-3.6Gbps | 5.6-6.4Gbps | 8.0-9.2Gbps | ≥9Gbps |
| 容量 | 1GB | 4-8GB | 8-16GB | 16-24GB | 24-36GB | 36-64GB |

资料来源：SEMI

**国内长鑫在合肥运营两座 12 英寸 DRAM 晶圆厂、在北京运营一座 12 英寸 DRAM 晶圆厂，合计月产能约 30 万片；**上海新工厂计划 2027 年起投产，达产后整体 DRAM 产能将为现有水平的两倍以上。其 DRAM 产品覆盖 DDR 与 LPDDR 等，广泛应用于移动终端、PC、服务器、虚拟现实及物联网等领域，**并已在技术上具备 HBM3 量产能力。**如果长鑫 2027 年可出货 300 万颗 24GB 等效 HBM3，按照单卡 96-144G 规格测算，共可支撑 50-75 万颗国产 GPU 需求，有望对包括寒武纪、海光等在内的国产 GPU 生产提供保障。

**图 9：长鑫 DRAM 产能爬坡预测曲线**




资料来源：AI-Foresight 公众号

仅供内部参考，请勿外传



# 4、投资建议

当前 HBM 供给紧张，**但从中期看，随着国产 HBM 供应逐步落地，这一供需缺口有望得到缓解。与此同时，AI 算力需求景气度仍在持续走高，行业需求端具备较强支撑。**在供给瓶颈边际改善、需求持续向好的背景下，我们维持对计算机行业的“推荐”评级。

**相关公司：**

1）CPU：海光信息、中国长城（飞腾信息）；2）AI 芯片：寒武纪、沐曦股份、壁仞科技、天数智芯、燧原科技、芯原股份；3）连接：澜起科技、盛科通信、万通发展、曦智科技；4）超节点：浪潮信息、华勤技术、联想集团、中科曙光、紫光股份、锐捷网络、工业富联；5）液冷：鼎通科技、飞荣达、申菱环境；6）电源：中国长城、欧陆通、麦格米特、中恒电气；7）算力租赁：协创数据、利通电子、宏景科技、盛视科技、智微智能、润建股份、赛意信息；8）垂类 AI 应用：中控技术、科大讯飞、鼎捷数智、万兴科技等。

# 5、风险提示

宏观经济影响下游需求、AI 模型发展不及预期、市场竞争加剧、中美博弈加剧、供应链风险、扩产不及预期、AI 商业化兑现不及预期、国产 HBM 量产进度及良率不及预期、GPU 出货量假设偏差、算力资本开支不及预期、政策规划落地不及预期。

1）宏观经济影响下游需求：AI 落地与宏观经济高度相关，若需求复苏不及预期，可能导致周期波动下 AI 落地发展或不及预期。

2）AI 发展不及预期：AI 大模型的发展或推动 AI+硬件发展，若 AI 大模型发展不及预期，或导致行业发展不及预期。

3）市场竞争加剧：更多企业加入进行相关研发，或导致行业竞争加剧。

4）中美博弈加剧：中美关系反复震荡，全球局势不确定性大幅增加。

5）供应链风险：AI 芯片制程工艺与先进封装技术持续演进，若新一代芯片量产进度不及预期或良率爬坡受阻，可能影响 AI 产业供给节奏。

6）AI 商业化兑现不及预期：若 AI 大模型在行业场景中的渗透速度低于预期，可能导致资本开支无法及时转化为相应收入与利润。

7）扩产不及预期：若相关企业因资金筹措困难、核心设备交付延迟或厂房建设进度缓慢等因素，能导致扩产进度低于预期。

仅供内部参考，请勿外传



8）国产 HBM 量产进度及良率不及预期：若国内 HBM 厂商在技术突破、量产时间节点或产品良率爬坡方面不及预期，可能导致国产 AI 芯片面临核心存储器件供应瓶颈。

9）GPU 出货量假设偏差：实际 GPU 出货量可能与假设有偏差。

10）算力资本开支不及预期：若宏观经济承压或 AI 商业模式闭环受阻，将导致互联网大厂、云计算厂商及运营商等主体削减算力基础设施的资本开支或投资节奏放缓。

11）政策规划落地不及预期：若国家及地方针对 AI 产业的扶持政策、规划目标及补贴措施在落地执行过程中受阻，可能影响 AI 产业的长期发展速度与确定性。

仅供内部参考，请勿外传


## 【计算机小组介绍】

刘熹，计算机行业首席分析师，上海交通大学硕士，多年计算机行业研究经验，善于把握由技术、政策驱动的科技产业新趋势，致力于进行前瞻重磅推荐。2024 年 Wind 金牌分析师，2024 年同花顺最受欢迎分析师，2025 年 Wind 金牌分析师。
唐锦珂，计算机行业研究助理，中山大学岭南学院硕士，主要覆盖服务器、液冷、算力租赁、IDC 等方向。
谢婧茹，计算机行业研究助理，上海财经大学经济学院硕士，主要覆盖半导体、云计算、教育 IT 等方向。
朱谓晟，计算机行业研究助理，复旦大学金融硕士，主要覆盖存储、网络等方向。

## 【分析师承诺】

刘熹，本报告中的分析师均具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，以勤勉的职业态度，独立，客观的出具本报告。本报告清晰准确的反映了分析师本人的研究观点。分析师本人不曾因，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取到任何形式的补偿。

## 【国海证券投资评级标准】

行业投资评级

推荐：行业基本面向好，行业指数领先沪深 300 指数；
中性：行业基本面稳定，行业指数跟随沪深 300 指数；
回避：行业基本面向淡，行业指数落后沪深 300 指数。

股票投资评级

买入：相对沪深 300 指数涨幅 20%以上；
增持：相对沪深 300 指数涨幅介于 10%～20%之间；
中性：相对沪深 300 指数涨幅介于-10%～10%之间；
卖出：相对沪深 300 指数跌幅 10%以上。

## 【免责声明】

本报告的风险等级定级为 R3，仅供符合国海证券股份有限公司（简称“本公司”）投资者适当性管理要求的客户（简称“客户”）使用。本公司不会因接收人收到本报告而视其为客户。客户及/或投资者应当认识到有关本报告的短信提示、电话推荐等只是研究观点的简要沟通，需以本公司的完整报告为准，本公司接受客户的后续问询。

本公司具有中国证监会许可的证券投资咨询业务资格。本报告中的信息均来源于公开资料及合法获得的相关内部外部报告资料，本公司对这些信息的准确性及完整性不作任何保证，不保证其中的信息已做最新变更，也不保证相关的建议不会发生任何变更。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，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波动。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。报告中的内容和意见仅供参考，在任何情况下，本报告中所表达的意见并不构成对所述证券买卖的出价和征价。本公司及其本公司员工对使用本报告及其内容所引发的任何直接或间接损失概不负责。本公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问或者金融产品等服务。本公司在知晓范围内依法合规地履行披露义务。

## 【风险提示】

仅供内部参考，请勿外传



市场有风险，投资需谨慎。投资者不应将本报告视为作出投资决策的唯一参考因素，亦不应认为本报告可以取代自己的判断。在决定投资前，如有需要，投资者务必向本公司或其他专业人士咨询并谨慎决策。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议，请公众投资者慎重使用未经授权刊载或者转发的本公司的证券研究报告。投资者务必注意，其据此做出的任何投资决策与本公司、本公司员工或者关联机构无关。

若本公司以外的其他机构（以下简称“该机构”）发送本报告，则由该机构独自为此发送行为负责。通过此途径获得本报告的投资者应自行联系该机构以要求获悉更详细信息。本报告不构成本公司向该机构之客户提供的投资建议。

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。本公司、本公司员工或者关联机构亦不为该机构之客户因使用本报告或报告所载内容引起的任何损失承担任何责任。

**【郑重声明】**

本报告版权归国海证券所有。未经本公司的明确书面特别授权或协议约定，除法律规定的情况外，任何人不得对本报告的任何内容进行发布、复制、编辑、改编、转载、播放、展示或以其他任何方式非法使用本报告的部分或者全部内容，否则均构成对本公司版权的侵害，本公司有权依法追究其法律责任。