

电子3月周报（3.18—3.24）

投资建议： 强于大市（维持）
 上次建议： 强于大市

GTC 2024 引领哪些硬件新方向？

► 英伟达推出 Blackwell 平台

3月19日凌晨，英伟达 CEO 黄仁勋在 GTC 大会上推出了新一代算力产品，包括 Blackwell 架构、Blackwell GPU、GB200 超级芯片组和 DGX GB200 系列服务器等。与上代产品 H100 相比，GB200 在算力、能耗和成本方面都有了很大的提升。此次 GTC 大会，黄仁勋强调了 Blackwell 平台的重要性，英伟达产品的重心也从过去提供芯片（GPU）向提供机柜（DGX 系列）、AI 数据中心转变。

► Blackwell GPU：双芯片设计，C2C 实现连接

Blackwell GPU 采用台积电定制的 4NP 制程工艺，拥有 2080 亿个晶体管，比 H100 多 1280 亿个晶体管。Blackwell GPU 采用双芯片设计，将两颗 Blackwell 架构的 GPU die 通过 NVLink-C2C 连接合封成一颗 GPU。Blackwell GPU 采用的 C2C 是多芯片封装工艺（MCM），与直接做成一颗大 die 相比，两颗小 die 的良率、设计成本及制造成本上更有优势。过去先进封装主要应用在异构芯片上，英伟达将 2 颗相同 die 合封有望引领封装进入新的应用场景。

► NVLink Switch：背板铜缆连接，成本效益显著

英伟达通过设计 NVLink Switch 芯片构建了 DGX GB200 NVL72 系统，实现了 GPU 之间的高速通信，就可以将 DGX 视为一个很大的 GPU 系统。NVLink Switch 芯片采用台积电 4NP 制造工艺，拥有 500 亿个晶体管，同时拥有 4 个带宽为 1.8TB/s 的 NVLink。NVLink Switch 采用铜连接技术，成本效益显著。使用光连接驱动 NVLink 主干需要耗电 2 万瓦，而 NVLink Switch 无需耗电就可以做到。此外，采用线缆背板架构能有效解决传统主板架构无法容纳众多 GPU 或 CPU 的问题。

► DGX SuperPOD：高效液冷机架，助力机柜散热

英伟达推出了下一代 AI 超级计算机——由 NVIDIA GB200 Grace Blackwell 超级芯片提供支持的 NVIDIA DGX SuperPOD，用于处理万亿参数模型，并具有持续的正常运行时间，以实现超大规模生成式 AI 训练和推理工作负载。新型 DGX SuperPOD 采用新型高效液冷机架规模架构，采用 NVIDIA DGX GB200 系统构建。与风冷技术相比，液冷的优势在于散热效率高、温度控制稳定且能耗低，冷板式液冷和浸没式液冷是行业目前的两条主流技术路线。

► 投资建议：新硬件新机遇

英伟达是全球 AI 龙头，其 GPU 技术将持续引领市场，Blackwell 平台带来的新硬件有望成为行业标杆，关注新硬件带来的增量市场机遇：（1）AI 芯片的制造离不开先进工艺，先进制造及封装相关标的：中芯国际、中微公司、芯源微、华海诚科等。（2）HBM 供给紧缺，相关标的：雅克科技、中微公司、华海诚科等。（3）重视铜连接、液冷产业链。

风险提示：GB200 量产、发售不及预期；HBM、CoWoS 供给不及预期。

相对大盘走势



作者

分析师：熊军
 执业证书编号：S0590522040001
 邮箱：xiongjun@glsc.com.cn

联系人：王海
 邮箱：wanghai@glsc.com.cn

相关报告

1、《电子：AI 手机有望驱动新一轮换机周期》
 2024.03.16
 2、《电子：下游需求向好，物联网终端渗透率有望提升》
 2024.03.10

正文目录

1. Blackwell 平台带来哪些新硬件?	3
1.1 双芯片设计, C2C 实现连接	4
1.2 背板铜缆连接, 成本效益显著	6
1.3 高效液冷机架, 助力机柜散热	8
2. 投资建议: 新硬件新机遇	9
2.1 关注先进封装受益标的	9
2.2 关注国产 HBM 受益标的	9
2.3 重视铜连接、液冷产业链	9
3. 风险提示	9

图表目录

图表 1: 英伟达基于 Blackwell 平台推出 GB200 系列产品	3
图表 2: AI 算力 8 年提升 1000 倍	3
图表 3: Blackwell GPU 性能远优于 Hopper GPU	3
图表 4: GB200 算力是 H100 的 30 倍	4
图表 5: 1 颗 Blackwell 架构 die 通过 C2C 工艺合封成 1 颗 Blackwell GPU	5
图表 6: Blackwell	5
图表 7: CoWoS 工艺	5
图表 8: 23-24 年全球 HBM 与 DRAM 产业产值	6
图表 9: 23-24 年各供货商 HBM/TSV 产能 (万片/月)	6
图表 10: NVLink Switch 芯片	6
图表 11: DGX GB200 NVL72	6
图表 12: 机柜背面采用铜连接技术	7
图表 13: Blackwell 实现 AI 大模型更有成本、能耗优势	7
图表 14: GB200 NVL72 采用高效液冷架构	8
图表 15: 中国液冷服务器市场规模 (亿美元)	9

1. Blackwell 平台带来哪些新硬件？

3月19日凌晨，英伟达 CEO 黄仁勋在 GTC 大会上推出了新一代算力产品，包括 Blackwell 架构、Blackwell GPU、GB200 超级芯片组和 DGX GB200 系列服务器等。此次 GTC 大会，黄仁勋强调了 Blackwell 平台的重要性，英伟达产品的重心也从过去提供芯片（GPU）向提供机柜（DGX 系列）、AI 数据中心转变。

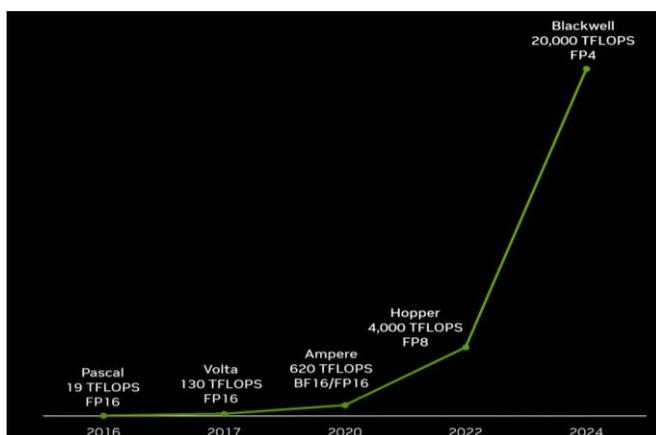
图表1：英伟达基于 Blackwell 平台推出 GB200 系列产品



资料来源：NVIDIA 官网，国联证券研究所

与上代产品 H100 相比，GB200 在算力、能耗和成本方面都有了很大的提升。过去 8 年，GPU 计算速度提升了近 1000 倍。Blackwell 架构的 token 生成能力是 Hopper 的 5 倍，推理能力是其 5 倍。

图表2：AI 算力 8 年提升 1000 倍



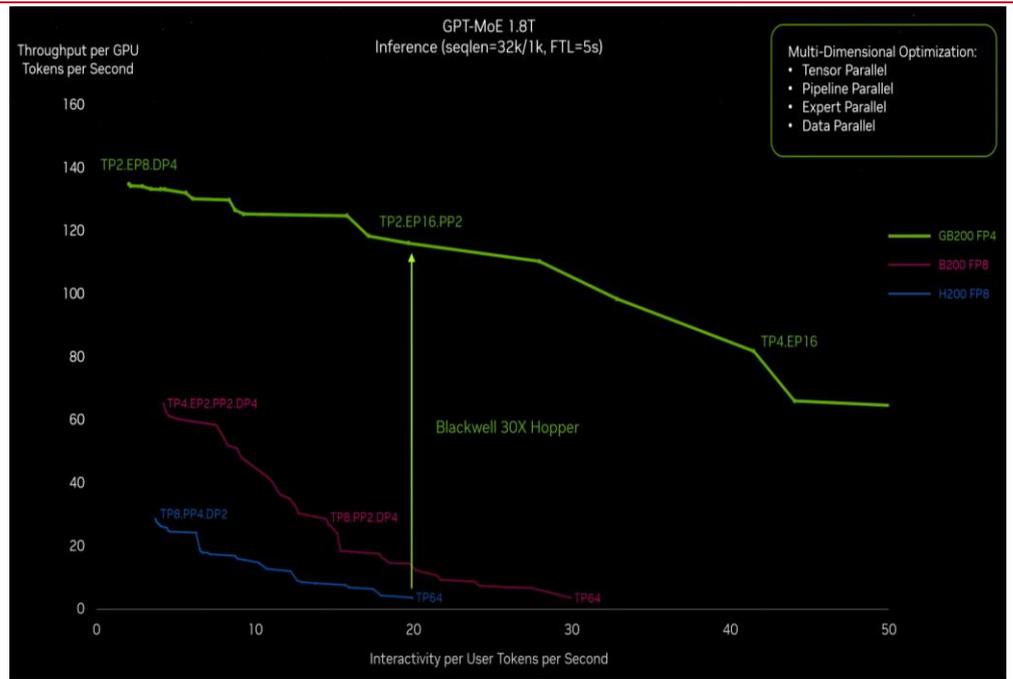
资料来源：NVIDIA 官网，国联证券研究所

图表3：Blackwell GPU 性能远优于 Hopper GPU

Blackwell GPU		
FP8	20 PFLOPS	2.5X Hopper
NEW FP6	20 PFLOPS	2.5X
NEW FP4	40 PFLOPS	5X
HBM Model Size	740B param	6X
HBM Bandwidth	34T param/sec	5X
NVLINK All-Reduce with SHARP	7.2 TB/s	4X

资料来源：NVIDIA 官网，国联证券研究所

图表4：GB200 算力是 H100 的 30 倍

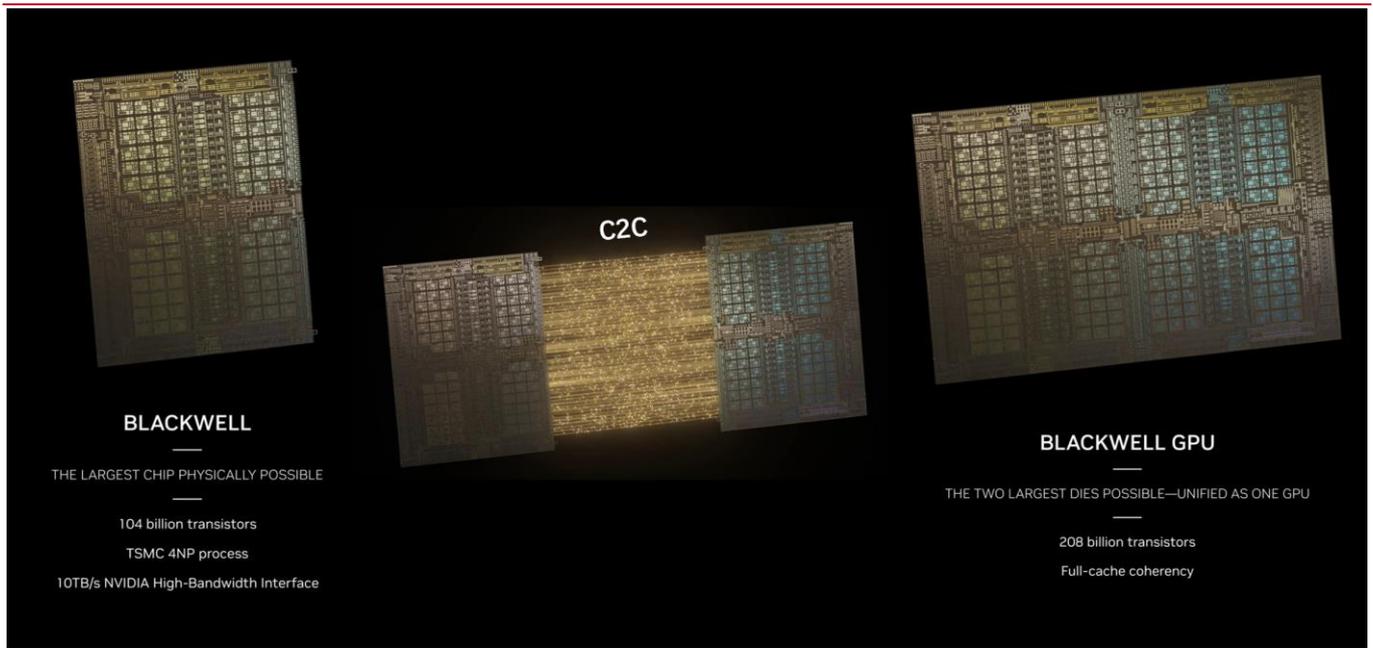


资料来源：NVIDIA 官网，国联证券研究所

1.1 双芯片设计，C2C 实现连接

Blackwell GPU 采用台积电定制的 4NP 制程工艺，拥有 2080 亿个晶体管，比 H100 多 1280 亿个晶体管。Blackwell GPU 采用双芯片设计，将两颗 Blackwell 架构的 GPU die 通过 NVLink-C2C 连接合封成一颗 GPU。两颗 GPU die 之间的数据传输速率可以达到 10TB/s，从而不存在内存局部性和缓存一致性问题，因此可将其视为单个 CUDA GPU。Blackwell GPU 采用的 C2C 是多芯片封装工艺 (MCM)，与直接做成一颗大 die 相比，两颗小 die 的良率、设计成本及制造成本上更有优势。

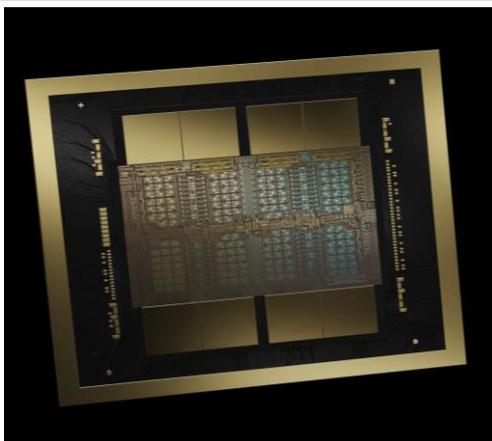
图表5: 1颗 Blackwell 架构 die 通过 C2C 工艺合封成 1颗 Blackwell GPU



资料来源: NVIDIA 官网, 国联证券研究所

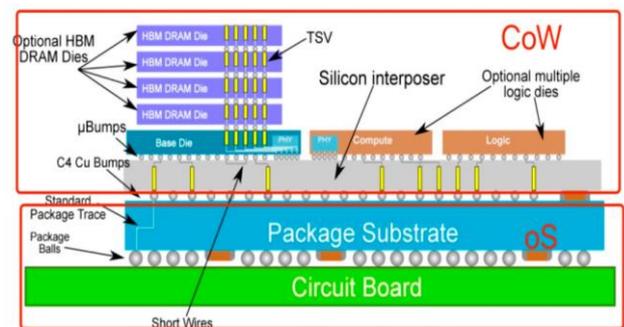
算力快速提升下, HBM、CoWoS 或供不应求。Blackwell GPU 搭配了 8 颗 HBM3e 内存, 容量可以达到 192GB, 相比 H200 增加 36%。GPU 和 HBM 通过采用 CoWoS 工艺合封在一块载板上。HBM 是垂直堆叠 DRAM 芯片, 通过硅通孔 (TSV) 连接, 并使用 TCB 键合。随着算力需求的快速提升, 对 HBM 的需求也将快速提升, 对 CoWoS 的产能也提出更多需求。此外, HBM 需求的快速提升, 对于高端 DRAM 颗粒的需求也有望提升, 存储颗粒市场库存有望实现较快消耗。

图表6: Blackwell



资料来源: NVIDIA 官网, 国联证券研究所

图表7: CoWoS 工艺

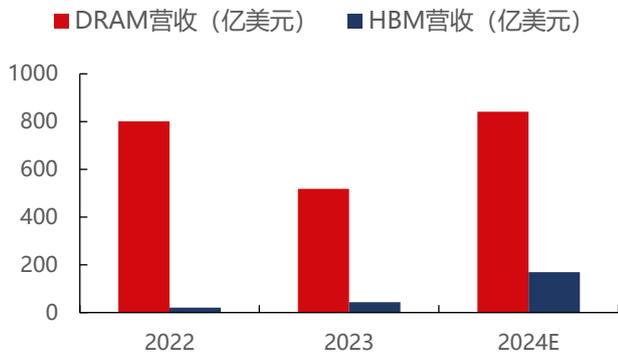


资料来源: 与非网, 国联证券研究所

HBM 市场快速增长, 供给持续紧缺。根据 TrendForce, 2023 年 HBM 产值约为 43.56 亿元, 占 DRAM 产业比重约 8.4%, 至 2024 年底扩大至 20.1%。以现阶段主流产品 HBM3 市占率来看, 海力士占比超过 90%, 三星将随着后续数个季度 AMD MI300 放

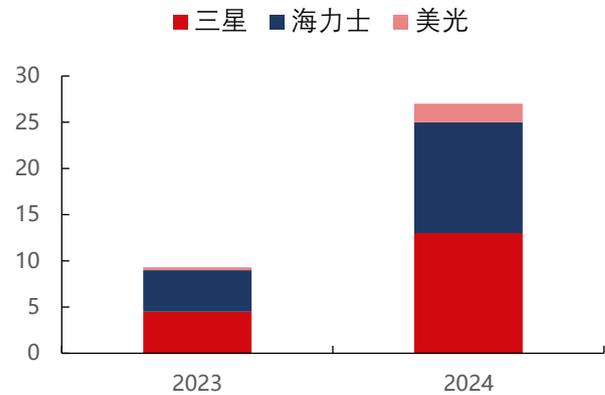
量持续紧追。2024年三星、海力士对于HBM产能的布局最积极，三星有望扩至13万片/月，海力士有望扩至12万片。考虑到HBM良率比DDR5低约20-30%、生产周期较DDR5多1.5-2个月，产能释放需要时间，HBM市场供应或将持续紧俏。

图表8：23-24年全球HBM与DRAM产业产值



资料来源：TrendForce，国联证券研究所

图表9：23-24年各供货商HBM/TSV产能（万片/月）



资料来源：TrendForce，国联证券研究所

1.2 背板铜缆连接，成本效益显著

通过NVLink Switch构建巨大的GPU系统，以实现GPU之间的全速通信。英伟达通过设计NVLink Switch芯片构建了DGX GB200 NVL72系统，实现了GPU之间的高速通信，就可以将DGX作为一个巨大的GPU系统。NVLink Switch芯片采用台积电4NP制造工艺，拥有500亿个晶体管，同时拥有4个带宽为1.8TB/s的NVLink。第五代NVLink发布后，最多可连接576个GPU。

图表10：NVLink Switch芯片



资料来源：NVIDIA官网，国联证券研究所

图表11：DGX GB200 NVL72



资料来源：NVIDIA官网，国联证券研究所

NVLink Switch采用铜连接技术，成本效益显著。在AI服务器中，采用线缆背

板架构的原因在于它能有效解决传统主板架构无法容纳众多 GPU 或 CPU 的问题。在单个 DXG 机柜里，每秒钟有 130TB 的数据通过机箱的背面，几乎可以在一秒内触及每个互联网客户。每个机柜里面有 5000 根 NVLink 电缆，总长度达 2 英里。如果使用光传输的话，就必须使用光模块和 retimer，使用这两个器件驱动 NVLink 主干就需要耗电 2 千瓦，而 NVLink Switch 无需耗电就可以做到。

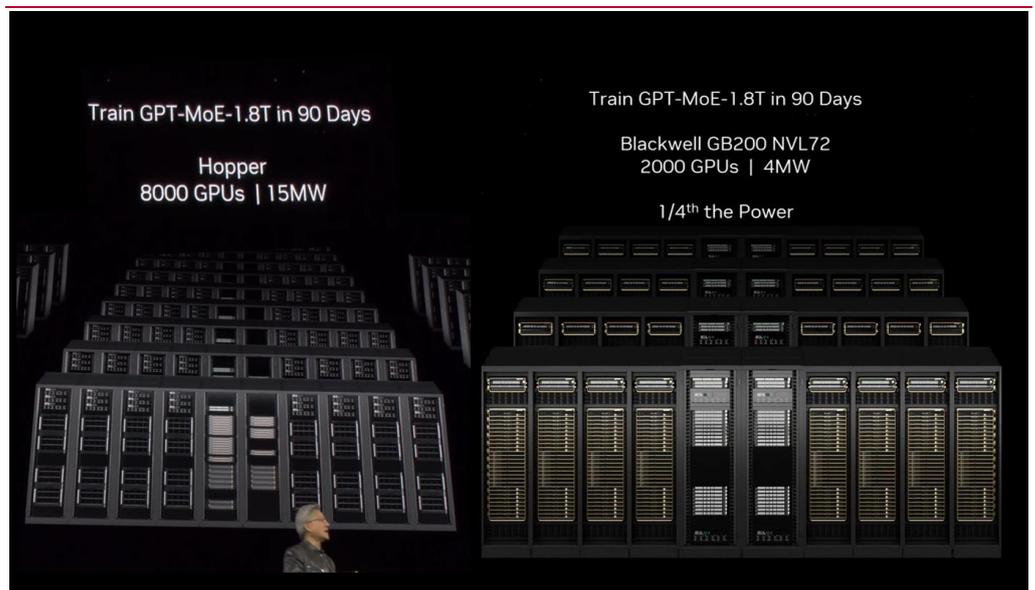
图表12: 机柜背面采用铜连接技术



资料来源: NVIDIA 官网, 国联证券研究所

相同 AI 大模型下, Blackwell 成本、能耗更有优势。如果要在 90 天内训练 1 个 1.8 万亿参数的模型, Hopper GPU 需要使用 8000 片、耗电约 15 兆瓦, Blackwell GPU 则只需要 2000 片、耗电仅 4 兆瓦。

图表13: Blackwell 实现 AI 大模型更有成本、能耗优势

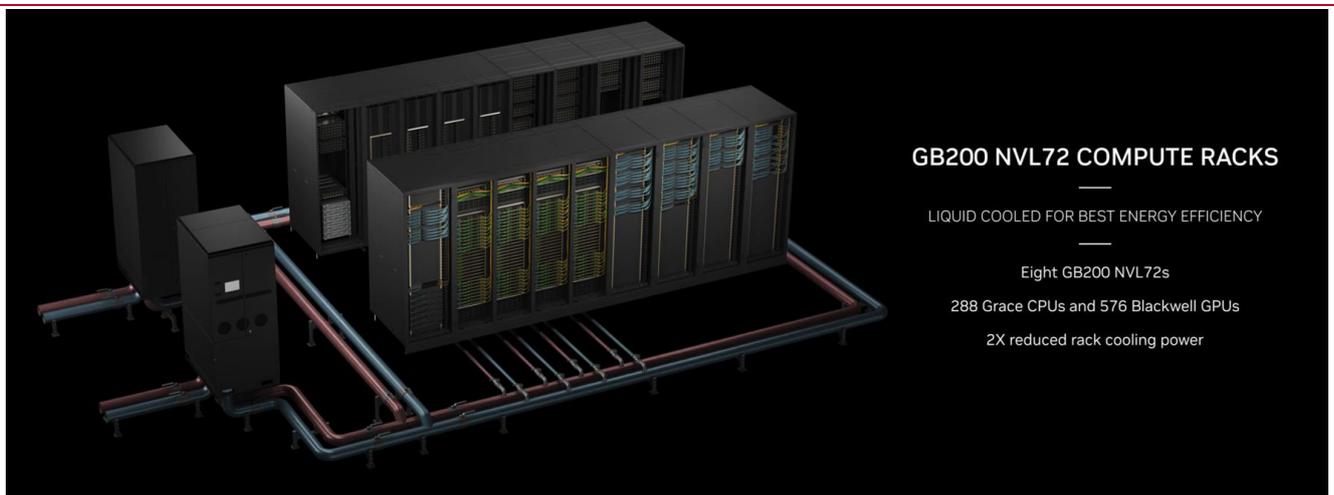


资料来源: NVIDIA 官网, 国联证券研究所

1.3 高效液冷机架，助力机柜散热

采用新型高效液冷机架规模架构。英伟达推出了下一代 AI 超级计算机——由 NVIDIA GB200 Grace Blackwell 超级芯片提供支持的 NVIDIA DGX SuperPOD，用于处理万亿参数模型，并具有持续的正常运行时间，以实现超大规模生成式 AI 训练和推理工作负载。新型 DGX SuperPOD 采用新型高效液冷机架规模架构，采用 NVIDIA DG GB200 系统构建。与风冷技术相比，液冷的优势在于散热效率高、温度控制稳定且能耗低，冷板式液冷和浸没式液冷是行业目前的两条主流技术路线。

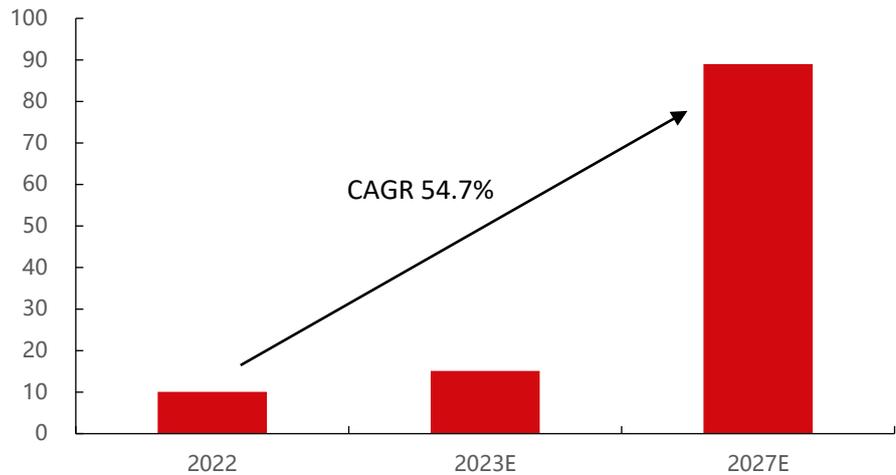
图表14：GB200 NVL72 采用高效液冷架构



资料来源：NVIDIA 官网，国联证券研究所

中国液冷服务器市场或将保持快速增长。2023 上半年中国液冷服务器市场规模达到 6.6 亿美元，同比增长 283.3%，预计 2023 年全年将达到 15.1 亿美元。IDC 预计，2022-2027 年，中国液冷服务器市场年复合增长率将达到 54.7%，2027 年市场规模将达到 89 亿美元。背板铜连接技术可带来成本效益，同时对散热提出更高要求，英伟达背板铜连接技术的使用有望加速液冷服务器市场规模快速增长。

图表15: 中国液冷服务器市场规模 (亿美元)



资料来源: IDC, 国联证券研究所

2. 投资建议: 新硬件新机遇

英伟达是全球 AI 龙头, 其 GPU 技术将持续引领市场, Blackwell 平台带来的新硬件有望成为行业标杆, 建议关注新硬件带来的增量市场机遇。

2.1 关注先进封装受益标的

先进算力离不开先进技术, 英伟达新一代 GPU 采用 4NP 制程和 CoWoS 封装工艺, 相关产业链集中在中国台湾、韩国、欧美等地, 国产发展空间广阔。随着国内加大对半导体产业的支持, 国内企业有望实现技术突破, 相关标的: 中芯国际、中微公司、芯源微、华海诚科等。

2.2 关注国产 HBM 受益标的

HBM 技术在高性能计算和人工智能领域展现了显著的高速、高带宽和低功耗特性, 在 AI 市场热潮下具有广泛的应用前景。当前全球 HBM 市场主要被海外垄断, 国内 HBM 产业尚处于发展初期, 部分企业开始布局, 相关标的: 雅克科技、中微公司、华海诚科等。

2.3 重视铜连接、液冷产业链

基于 Blackwell 平台, 英伟达发布了一系列超级计算产品, 新增铜连接、液冷技术, 而新技术的引入对 AI 计算的成本效益、算力性能均有所提升。我们认为, 英伟达新技术的引入有望加速这些新技术在 AI 领域的渗透, 重视铜连接、液冷产业链。

3. 风险提示

1) **GB200 量产、发售不及预期。** 英伟达发布的新品若量产、售卖不及预期, 对

相关产业链产生不利影响。

- 2) **HBM、CoWoS 供给不及预期。**高算力 GPU 芯片必须配套 HBM 等产品或工艺，若相关产能供给不及预期，对产业链造成不利影响。

分析师声明

本报告署名分析师在此声明：我们具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，本报告所表述的所有观点均准确地反映了我们对标的证券和发行人的个人看法。我们所得报酬的任何部分不曾与，不与，也将不会与本报告中的具体投资建议或观点有直接或间接联系。

评级说明

投资建议的评级标准		评级	说明
报告中投资建议所涉及的评级分为股票评级和行业评级（另有说明的除外）。评级标准为报告发布日后6到12个月内的相对市场表现，也即：以报告发布日后的6到12个月内的公司股价（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。其中：A股市场以沪深300指数为基准，新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准；香港市场以摩根士丹利中国指数为基准；美国市场以纳斯达克综合指数或标普500指数为基准；韩国市场以柯斯达克指数或韩国综合股价指数为基准。	股票评级	买入	相对同期相关证券市场代表指数涨幅20%以上
		增持	相对同期相关证券市场代表指数涨幅介于5%~20%之间
		持有	相对同期相关证券市场代表指数涨幅介于-10%~5%之间
		卖出	相对同期相关证券市场代表指数跌幅10%以上
	行业评级	强于大市	相对同期相关证券市场代表指数涨幅10%以上
		中性	相对同期相关证券市场代表指数涨幅介于-10%~10%之间
		弱于大市	相对同期相关证券市场代表指数跌幅10%以上

一般声明

除非另有规定，本报告中的所有材料版权均属国联证券股份有限公司（已获中国证监会许可的证券投资咨询业务资格）及其附属机构（以下统称“国联证券”）。未经国联证券事先书面授权，不得以任何方式修改、发送或者复制本报告及其所包含的材料、内容。所有本报告中使用的商标、服务标识及标记均为国联证券的商标、服务标识及标记。

本报告是机密的，仅供我们的客户使用，国联证券不因收件人收到本报告而视其为国联证券的客户。本报告中的信息均来源于我们认为可靠的已公开资料，但国联证券对这些信息的准确性及完整性不作任何保证。本报告中的信息、意见等均仅供客户参考，不构成所述证券买卖的出价或征价邀请或要约。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。客户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、税收等方面咨询专家的意见。对依据或者使用本报告所造成的一切后果，国联证券及/或其关联人员均不承担任何法律责任。

本报告所载的意见、评估及预测仅为本报告出具日的观点和判断。该等意见、评估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，国联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。

国联证券的销售人员、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。国联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。国联证券的资产管理部门、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。

特别声明

在法律许可的情况下，国联证券可能会持有本报告中提及公司所发行的证券并进行交易，也可能为这些公司提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。因此，投资者应当考虑到国联证券及/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一参考依据。

版权声明

未经国联证券事先书面许可，任何机构或个人不得以任何形式翻版、复制、转载、刊登和引用。否则由此造成的一切不良后果及法律责任有私自翻版、复制、转载、刊登和引用者承担。

联系我们

北京：北京市东城区安定门外大街208号中粮置地广场A塔4楼

无锡：江苏省无锡市金融一街8号国联金融大厦12楼

电话：0510-85187583

上海：上海浦东新区世纪大道1198号世纪汇一座37楼

深圳：广东省深圳市福田区益田路4068号卓越时代广场1期13楼