

电子行业 | 行业深度研究报告

证券研究报告

2025/03/25

行业深度报告——

GTC大会召开，催生供应链新增量

证券分析师： 张世杰
分析师登记编号： S1190523020001
证券分析师： 李珏晗
分析师登记编号： S1190523080001

报告摘要

Blackwell Ultra平台发布，测试时扩展推理能力优化。英伟达于GTC发布Blackwell平台新一代产品Blackwell Ultra，该产品定位提升训练和测试时扩展推理能力。Blackwell Ultra采用台积电N4P工艺，单卡FP4浮点运算性能为15 PetaFlops，相较B200提升50%；采用8堆栈12层堆叠的HBM3e，容量提升至288GB，相较B200提升50%。Blackwell Ultra核心变化聚焦显存容量的优化，为AI推理大时代的开启作出铺垫。

Rubin Ultra NVL576性能大幅提升，“纵向扩展”推进。GTC大会预告下一代Rubin架构，基于Rubin Ultra的Rubin Ultra NVL576预计于27年下半年推出，将标配HBM4，FP4推理浮点运算能力达到15 ExaFlops，FP8训练运算能力达到5 ExaFlops，是GB300 NVL72的14倍。Rubin Ultra NVL576性能大幅提升，“纵向扩展”持续推进。

GB300&RUBIN架构变化，催生供应链新增量。变化一：回归UBB+OAM板，在GB200中，Compute Tray为集成了2颗GPU、1颗Grace CPU、LPDDR5X等器件的Bianca主板；预计英伟达在GB300中将不再采用Bianca主板，而是采用UBB+OAM模式，由英伟达提供GPU SXM Puck，客户配置更加灵活。变化二：有望引入PTFE背板架构。考虑到Rubin性能大幅提升，功耗随之增加，Rubin系列中有望迎来架构调整。正交架构或成为机柜方案，从而提升传输效率。正交架构需要高传输速率及低损耗，PTFE材料介电损耗仅为2.1MHz，有望引入PTFE背板。

风险提示：下游需求不及预期风险；行业竞争加剧风险；地缘政治风险。

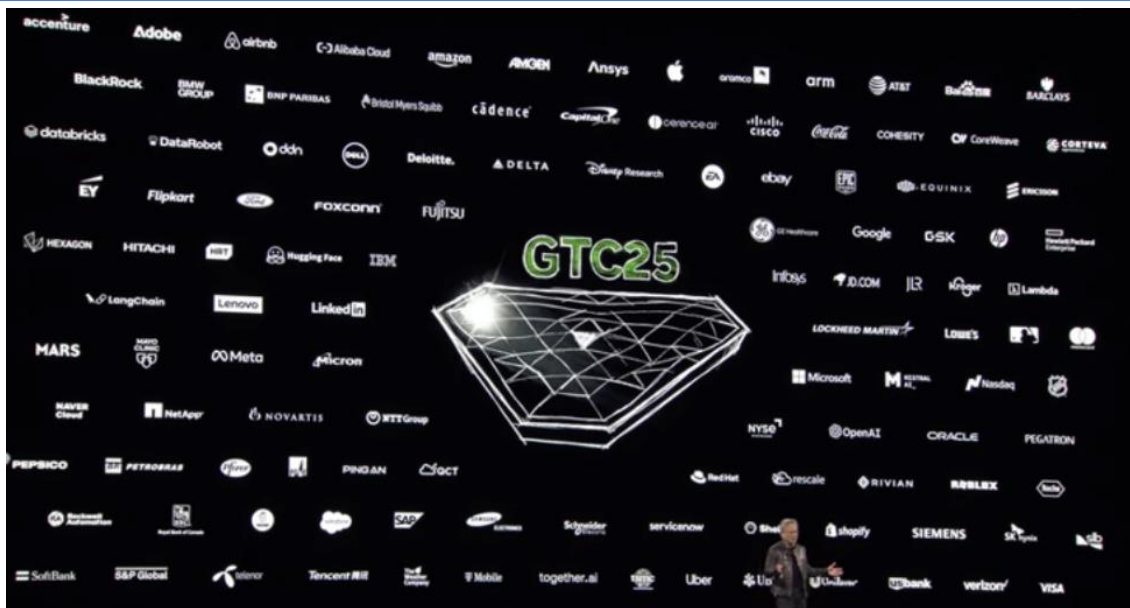
目录

- I Blackwell Ultra平台发布，测试时扩展推理能力优化
- II Rubin Ultra NVL576性能大幅提升，“纵向扩展”推进
- III GB300&Rubin平台架构变化，催生供应链新增量

1.1. Blackwell Ultra于GTC大会发布，预计25年下半年出货

➤英伟达于GTC大会发布Blackwell Ultra平台，开启AI推理新时代。3月17日至21日美国加州圣何塞举行GTC 2025大会，英伟达CEO黄仁勋发表主题演讲，发布Blackwell平台新一代产品Blackwell Ultra，该产品定位提升训练和测试时扩展推理能力，即通过在推理过程中增加计算量来提升准确率，开启AI推理新时代。

图表：英伟达GTC大会



资料来源：英伟达GTC2025，英伟达官网，太平洋证券

1.1. Blackwell Ultra于GTC大会发布，预计25年下半年出货

➤基于Blackwell Ultra的产品预计将于2025年下半年出货。参考GB200时间线，GB200于GTC2024正式发布后，同年四季度开始出货，25年一季度批量出货，GB300发布时间与GB200出货时间相隔仅半年左右，预测GB300或于2025年下半年左右量产出货。

图表：产品名称调整与规格预测

	B100	B200	B300	GB200	GB300	B300A	GB300A
名称调整前	B100	B200	B200 Ultra	GB200	GB200 Ultra	B200 Ultra A	GB200 Ultra
服务器	HGX	HGX	HGX	NVL72/36	NVL72/36	HGX/MGX	NVL36/MGX
CoWoS类型	CoWoS-L	CoWoS-L	CoWoS-L	CoWoS-L	CoWoS-L	CoWoS-S	CoWoS-S

资料来源：Trendforce，太平洋证券

图表：产品出货时间预测

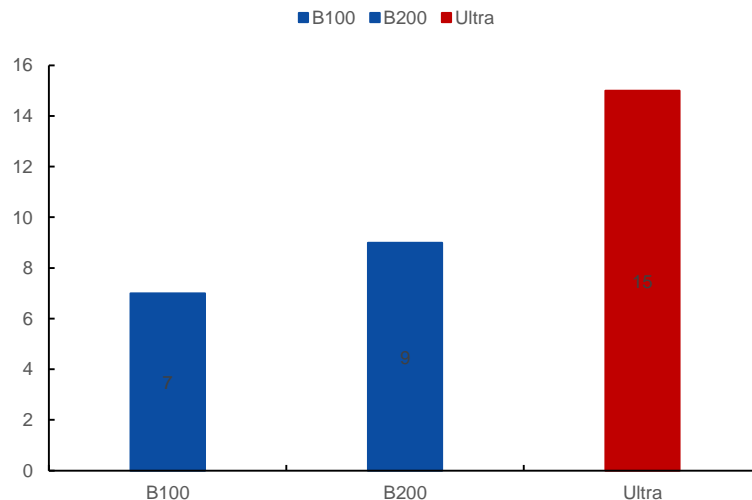
	24Q1	24Q2	24Q3	24Q4	25Q1	25Q2	25Q3	25Q4	2026
HGX		H200		B200		B300			R200
			B100		B200A				
Oberon		GH200		GB200 NVL72		GB300 NVL72			
						GB200A NVL72			

资料来源：SemiAnalysis，太平洋证券

1.1. Blackwell Ultra于GTC大会发布，预计25年下半年出货

➤**Blackwell Ultra计算性能提升，FP4 15 Petaflops。**英伟达2025年下半年将逐步过渡升级至Blackwell Ultra，Blackwell Ultra采用台积电N4P工艺，推理能力大幅优化，单卡FP4浮点运算性能为15 PetaFlops，相较B200FP4的单卡算力9 PetaFlops，Blackwell Ultra单卡FP4浮点运算性提升50%以上。

图表：Blackwell Ultra FP4算力（单位：PF）

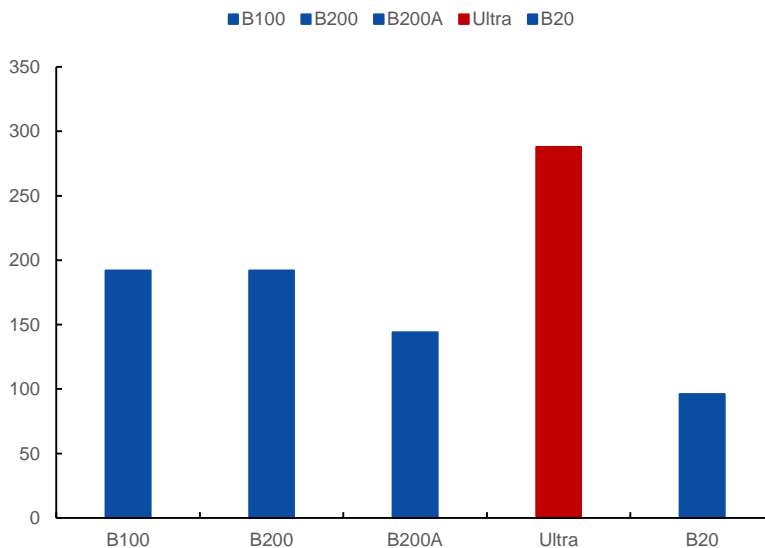


资料来源：英伟达GTC2025，英伟达官网，太平洋证券

1.1. Blackwell Ultra于GTC大会发布，预计25年下半年出货

➤**Blackwell Ultra 存力显著提升，升级至288GB 12hiHBM3e。**显存方面，Blackwell Ultra采用8堆栈12层堆叠的HBM3e，B100、B200均为8层堆叠的HBM3e，Blackwell Ultra显存容量提升至288GB，B200显存容量仅为192GB，相较提升50%。Blackwell Ultra与上一代相同，支持NVLink v5，带宽1.8TB/s。

图表：Blackwell Ultra显存容量（单位：GB）



图表：各代显存容量对比

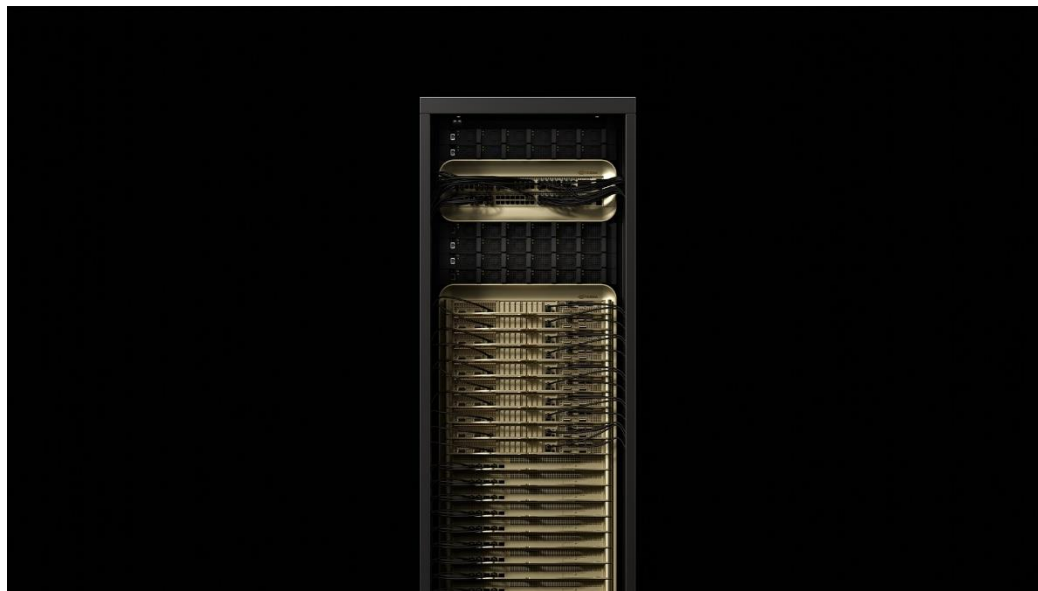
产品	HBM类型	显存容量
B100	HBM3e 8hi*8	192GB
B200	HBM3e 8hi*8	192GB
Ultra	HBM3e 12hi*8	288GB
GB200	HBM3e 8hi*8	192GB
GB300	HBM3e 12hi*8	288GB
B300A	HBM3e 12hi*4	144GB
GB300A	HBM3e 12hi*4	144GB

资料来源：GTC 2025, Trendforce, 太平洋证券

1.2. 基于Blackwell Ultra的GB300 NVL72，专注大规模AI推理场景

➤ 基于Blackwell Ultra的GB300 NVL72，专注大规模AI推理场景。Blackwell Ultra包括GB300 NVL72机架级解决方案和HGX B300 NVL16系统。GB300 NVL72在单机架连接72个Blackwell Ultra GPU+36个基于Arm Neoverse架构的Grace CPU，有18个Compute Tray+9个NVLink Switch Tray组成，专注测试时扩展的应用场景。

图表：GB300 NVL72机柜



资料来源：英伟达官网，太平洋证券

1.2. 基于Blackwell Ultra的GB300 NVL72，专注大规模AI推理场景

➤**GB300 NVL72性能优化，为上代1.5倍。**GB300 NVL72性能为GB200 NVL72的1.5倍，FP4推理浮点运算能力达到1.1ExaFlops，GB200 NVL72 FP4为1440 PetaFlops，FP8训练运算能力达到0.36 ExaFlops；配备20TB内存，上一代内存为14TB，为上一代1.5倍；总带宽为576TB/s。HGX B300 NVL16相比Hopper系列，LLM推理速度增加11倍，计算性能增加7倍，内存容量增加4倍。

图表： GB300 NVL72机柜



图表：GB300与GB200存算性能对比

类型	FP4	内存	内存带宽
GB300 NVL72	1.1 ExaFlops	20 TB	576 TB/s
GB200 NVL72	1440 PetaFlops	14 TB	576 TB/s

资料来源：英伟达GTC 2025，太平洋证券

1.3 内存模组、网卡、光模块等配置均有升级

- **内存条预计引入LPCAMM模组。**GB300的外部存储预计引入LPCAMM模块，GB200采用的方案是焊接LPDDR5X，GB300外接存储器LPCAMM支持可插拔与扩展，带来灵活性提升。
- **网卡升级为CX8，提升一倍。**GB300预计采用800G ConnectX-8，在GB200中网卡配置为400G ConnectX-7，CX8带宽较CX7提升一倍，带宽为14.4TB/s，传输速率更快。
- **单卡功耗提升，但整机柜能耗预计不变。**GB300单卡功耗提升到1.4kW，相较GB200的单卡功耗1.2kW，G300B单卡功耗提升约17%，而整体机柜能耗预计仍控制在132kW。

图表：LPCAMM2模块



资料来源：美光官网，太平洋证券

图表：英伟达网卡



资料来源：英伟达官网，太平洋证券

1.4. DGX SuperPOD赋能AI推理，助力企业快速部署

➤ **基于Blackwell Ultra的DGX SuperPOD赋能AI推理。** DGX SuperPOD AI超级计算机采用DGX GB300和DGX B300系统，集成NVQuantum-X800 InfiniBand或Spectrum-X以太网，助力企业快速部署。DGX SuperPOD提供FP4精度计算，大幅提升AI应用中Token生成速率。DGX GB300采用液冷架构，专注高级推理模型的实时代理响应，DGX B300采用风冷架构，专注生成式和代理式AI。

图表：DGX SuperPod



资料来源：英伟达官网，太平洋证券

目录

- I Blackwell Ultra平台发布，测试时扩展推理能力优化
- II Rubin Ultra NVL576性能大幅提升，“纵向扩展”推进
- III GB300&Rubin平台架构变化，催生供应链新增量

2.1. 新继代Vera Rubin预计26年出货，配置全面更新

➤ **新一继代Rubin架构将于2026年发布。**英伟达继Hopper、Blackwell之后的下一代Rubin发布，相较Hopper、Blackwell的每两年发布新一代，Rubin Ultra将于预计2027年发布。Rubin架构基于台积电3nm工艺，采用CoWoS-L封装。

图表：Rubin平台时间线

	24Q1	24Q2	24Q3	24Q4	25Q1	25Q2	25Q3	25Q4	26Q1	26Q2	26Q3	26Q4	2027
HGX		H200		B200		Blackwell Ultra				Rubin			Rubin Ultra
			B100		B200A								
Oberon		GH200		GB200 NVL72		GB300 NVL72					Vera Rubin NVL144		
							GB200A NVL72						

资料来源：英伟达GTC2025，太平洋证券

2.1. 新继代Vera Rubin预计26年出货，配置全面更新

➤**Vera Rubin架构正式问世，配置全面更新。**Rubin GPU FP4精度下达到50 petaflops，是Blackwell Ultra 15 petaflops的3.3倍以上，推理能力大幅提升；配置8颗16层堆叠HBM4，容量288GB。Vera CPU是基于英伟达自主架构Olympus的处理器，集成88个定制Arm核心，总线程数达到176个，支持多线程技术，相较上一代Grace CPU为72核，仅支持单线程。Vera CPU通过NVLink-C2C与Rubin GPU链接，带宽达1.8TB/s。

图表：Rubin GPU



图表：Vera CPU



资料来源：英伟达GTC2025，英伟达官网，太平洋证券

图表：Vera Rubin与Grace Blackwell Ultra对比

	Rubin	Blackwell Ultra
工艺	N4P	3nm
FP4推理算力	15 PetaFlops	50 PetaFlops
HBM类型	HBM3e	HBM4
内存容量	288GB	288GB
功耗	1400W	1800W

	Vera GPU	Grace CPU
架构	Arm Olympus	Arm Neoverse
核心	88核	72核
线程	多线程	单线程
互联带宽	1.8 TB/s	900GB/s

资料来源：英伟达GTC2025，太平洋证券

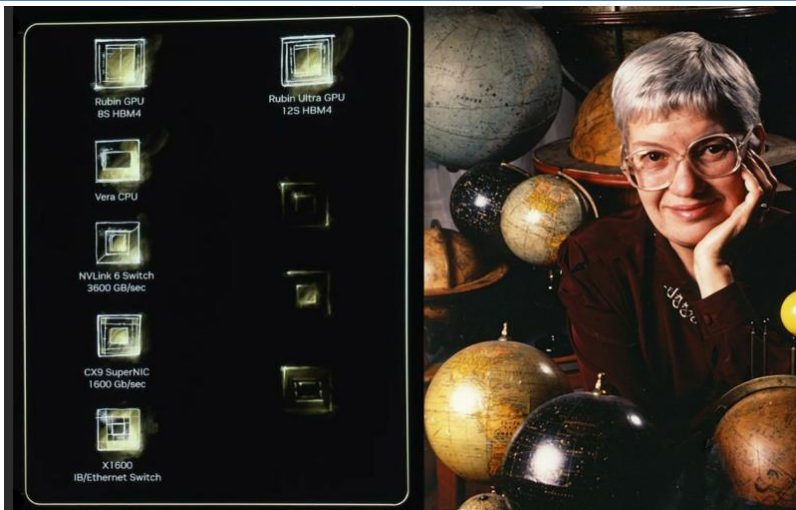
2.2. Rubin Ultra NVL576性能突破性提升

➤**Vera Rubin NVL144提供3.3倍NVL72性能。**在Blackwell架构为双GPU架构，即1个Blackwell芯片上集成2个GPU，并以Blackwell芯片数量计数，而从Rubin系列开始，英伟达将使用GPU die数量作为计数，每个GPU die为一个GPU，每个NVL均连接到GPU，集成为Vera Rubin NVL144。Vera Rubin NVL144将标配HBM4，带宽为13TB/s，FP4推理浮点运算能力达到3.6 ExaFlops，FP8训练运算能力达到1.2 ExaFlops，是GB300 NVL72的3.3倍；配置Connect-X9，带宽为28.8TB/s，为上一代Connect-X8的2倍；支持NVLink6，带宽为260TB/s，带宽提高一倍。

图表：Vera Rubin NVL144机柜



图表：Vera Rubin全面升级



资料来源：英伟达GTC2025，英伟达官网，太平洋证券

图表：Vera Rubin NVL144与GB300 NVL72对比

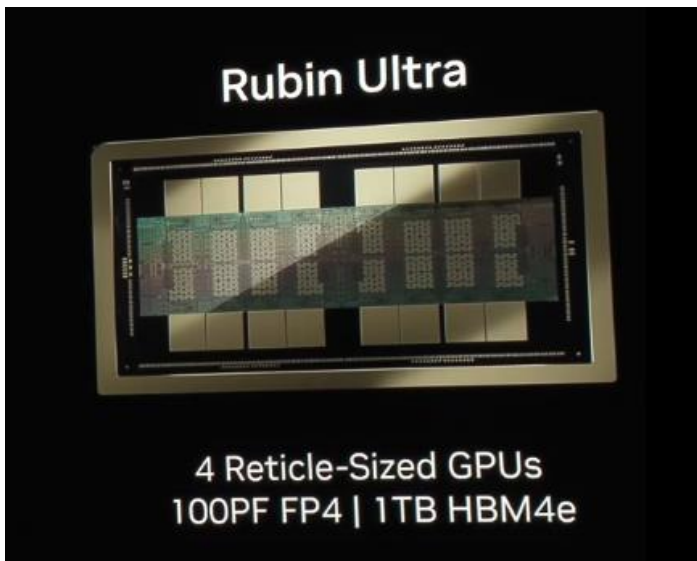
参数	Vera Rubin NVL144	GB300 NVL72	提升
FP4推理算力	3.6 Eflops	1.1 Eflops	+230%
FP8训练算力	1.2 Eflops	0.36 Eflops	+230%
内存带宽	13 TB/s HBM4	8 TB/s HBM3e	+62%
快速存储	75TB	40TB	+88%
互联带宽	260 TB/s 第六代NVLink	1.8 TB/s 第五代NVLink	+100%
网络	28.8TB/s ConnectX-8	14.4TB/s ConnectX-8	+100%

资料来源：英伟达GTC2025，太平洋证券

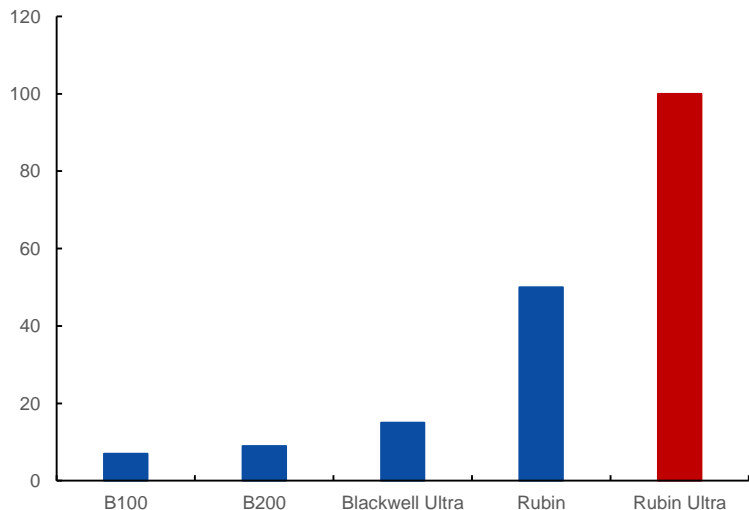
2.2. Rubin Ultra NVL576性能突破性提升

➤下一代Rubin Ultra由4个GPU die封装，预计2027年正式发布。Rubin架构平台的新一子代Rubin Ultra将于2027年发布，Rubin Ultra芯片上将集成4个GPU die。Rubin Ultra的FP4推理算力高达100 Petaflops，是Rubin的50 Petaflops FP4算力的2倍，同时将配置HBM4e，内存容量升级至1TB，是Rubin的3.5倍。

图表：Rubin Ultra芯片



图表：各产品FP4（单位：PetaFlops）

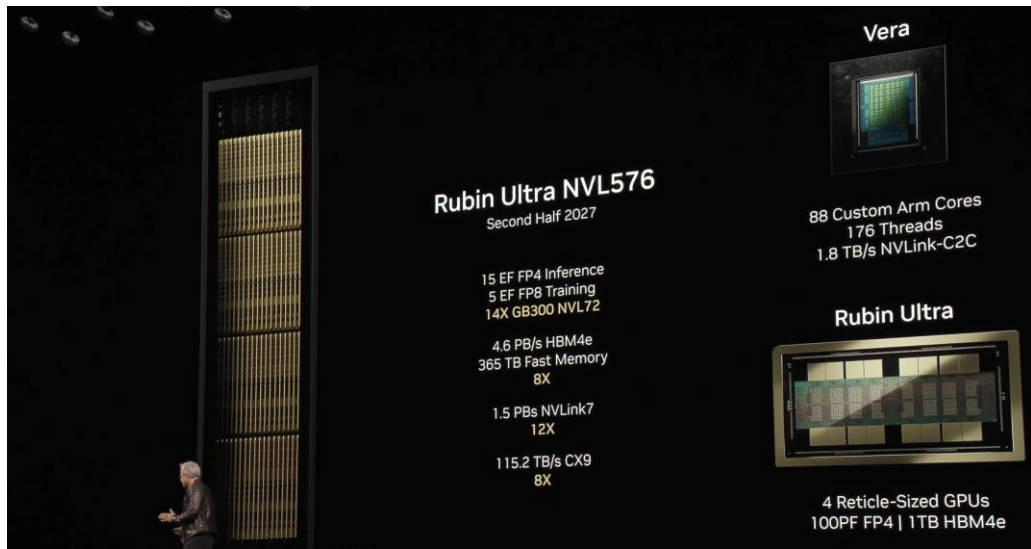


资料来源：英伟达GTC2025，英伟达官网，太平洋证券

2.2. Rubin Ultra NVL576性能突破性提升

➤**Rubin Ultra NVL576提供14倍GB300 NVL72性能。**基于Rubin Ultra的Rubin Ultra NVL576预计于27年下半年推出，将标配HBM4，带宽为4.6PB/s，FP4推理浮点运算能力达到15 ExaFlops，FP8训练运算能力达到5 ExaFlops，是GB300 NVL72的14倍；配置Connect-X9，带宽为28.8TB/s，为上一代Connect-X9的2倍；支持NVLink7，带宽为1.5PB/s，带宽为上一代的12倍。

图表：Rubin Ultra NVL576机柜



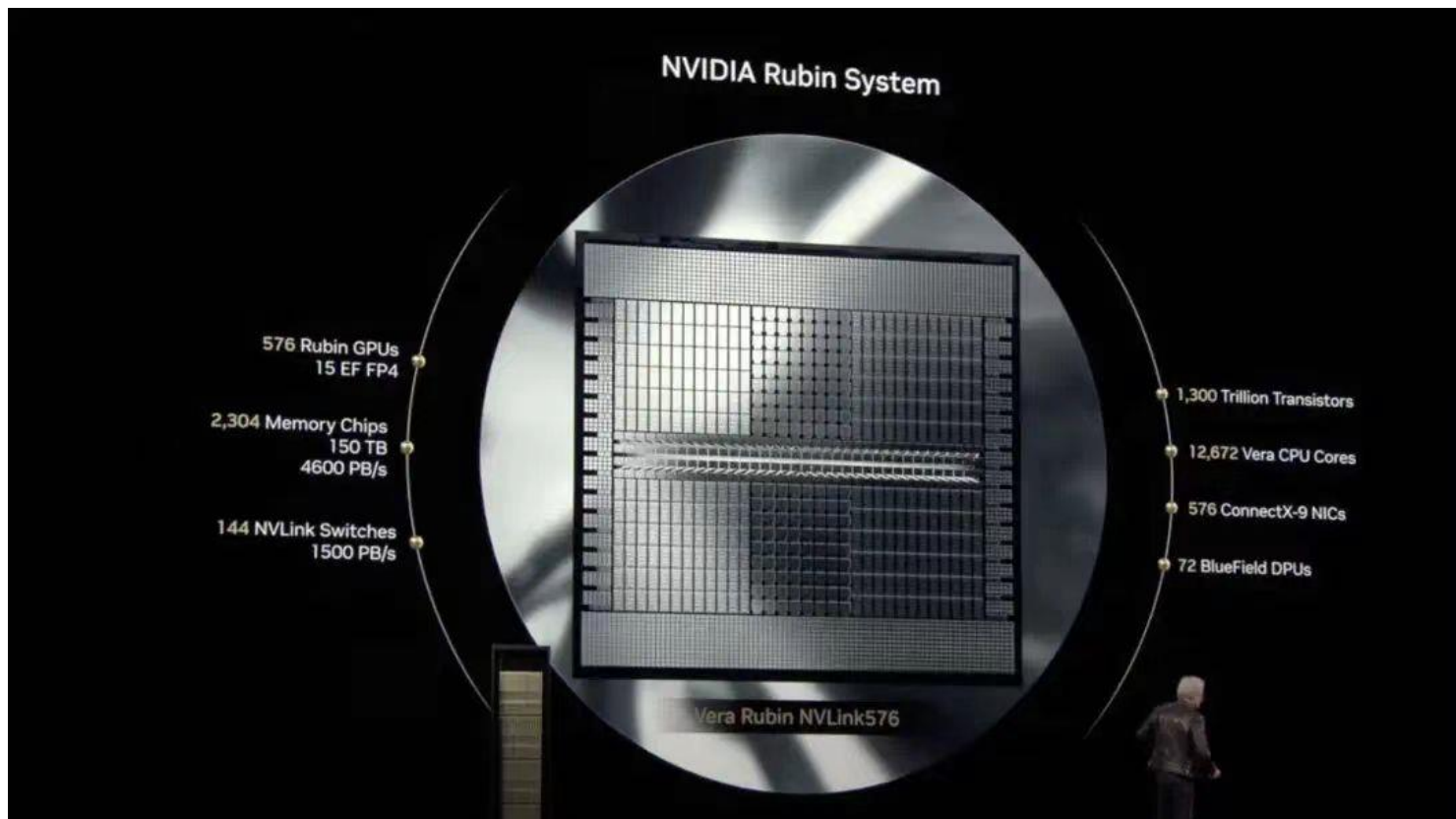
资料来源：英伟达GTC2025，英伟达官网，太平洋证券

图表：Rubin Ultra NVL576与GB300 NVL72对比

参数	Rubin Ultra NVL576	GB300 NVL72	提升
FP4推理算力	15 EFlops	1.1 EFlops	+1300%
FP8训练算力	5 EFlops	0.36 EFlops	+1300%
内存带宽	4.6 PB/s HBM4e	8 TB/s HBM3e	+57400%
快速存储	365TB	40TB	+700%
互联带宽	1.5 PB/s 第七代NVLink	1.8 TB/s 第五代NVLink	+1100%
网络	115.2 TB/s ConnectX-9	14.4TB/s ConnectX-8	+700%

资料来源：英伟达GTC2025，太平洋证券

图表：Rubin System



资料来源：英伟达GTC2025，英伟达官网，太平洋证券

目录

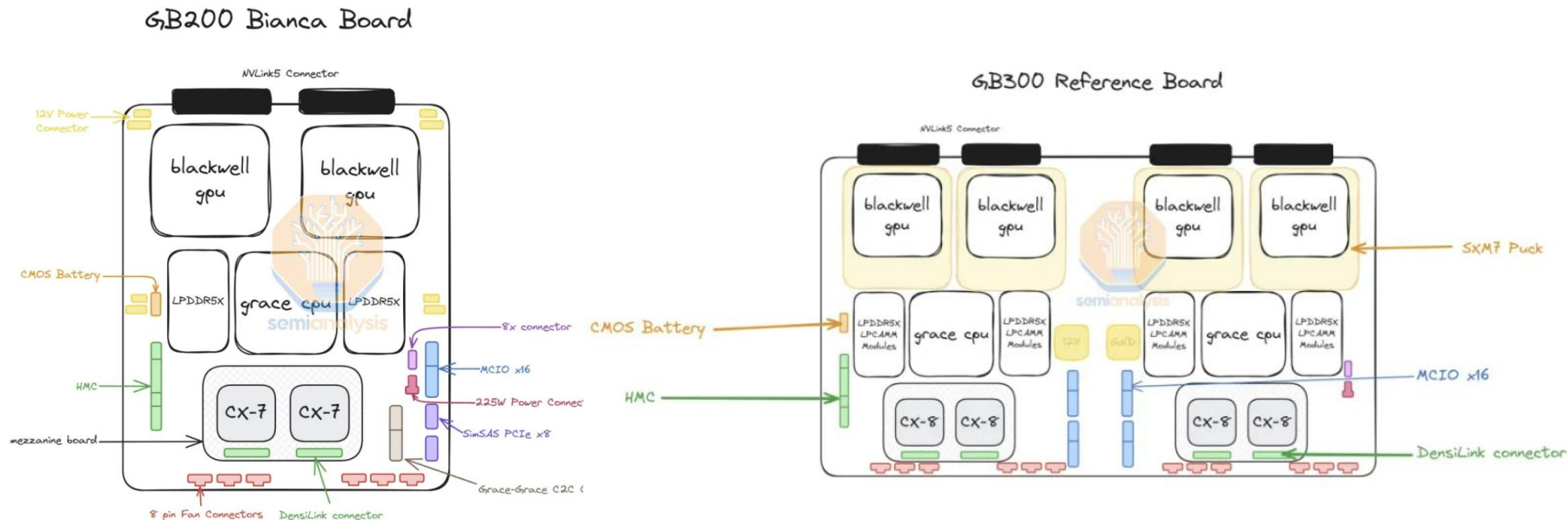
- I Blackwell Ultra平台发布，测试时扩展推理能力优化
- II Rubin Ultra NVL576性能大幅提升，“纵向扩展”推进
- III GB300&Rubin平台架构变化，催生供应链新增量

3.1 架构模块化，助力灵活配置需求

►回归UBB+OAM架构，为客户提供灵活配置可能性。在GB200中，Compute Tray为集成了2颗GPU、1颗Grace CPU、LPDDR5X等器件的Bianca主板；预计英伟达在GB300中将不再采用Bianca主板，而是采用UBB+OAM模式，由英伟达提供GPU SXM Puck，客户配置更加灵活。

图表：GB200 Bianca Board架构

图表：GB300 Reference Board架构

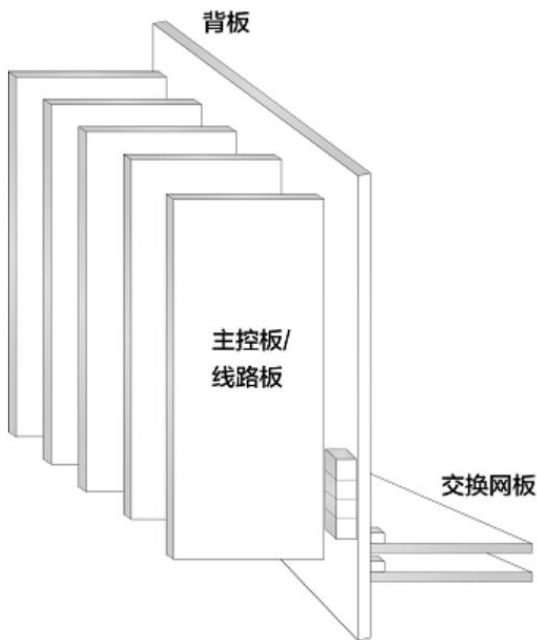


资料来源：SemiAnalysis，太平洋证券

3.2 PTFE材料有望助力高传输低损耗的实现

►PTFE具备超低介电损耗优势，Rubin有望引入PTFE背板架构。考虑到Rubin性能大幅提升，功耗随之增加，Rubin系列中有望迎来架构调整。正交架构或成为机柜方案，从而提升传输效率。正交架构需要高传输速率及低损耗，PTFE材料介电损耗仅为2.1，有望引入PTFE背板。

图表：伪正交架构



图表：不同材料Dk/Df值对比

材料	Dk	Df
环氧树脂 EP	4.5	0.022
聚苯醚 PPO	2.5	0.0007
碳氢树脂 CH	2.6	0.005
聚四氟乙烯 PTFE	2.1	0.0004

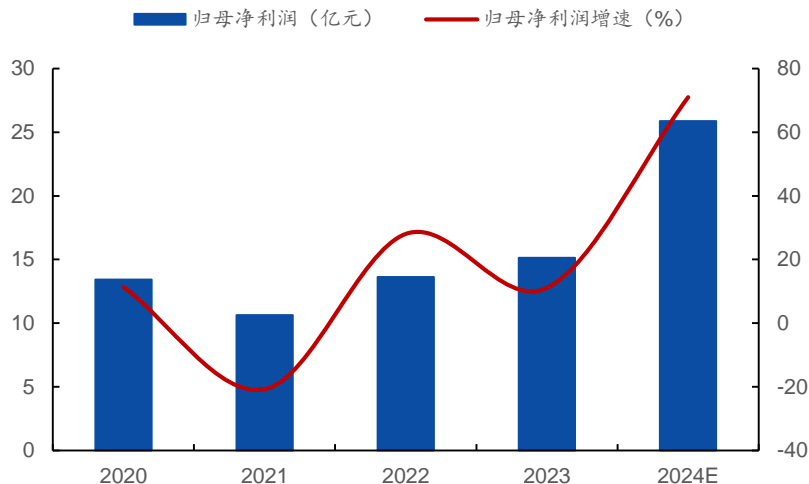
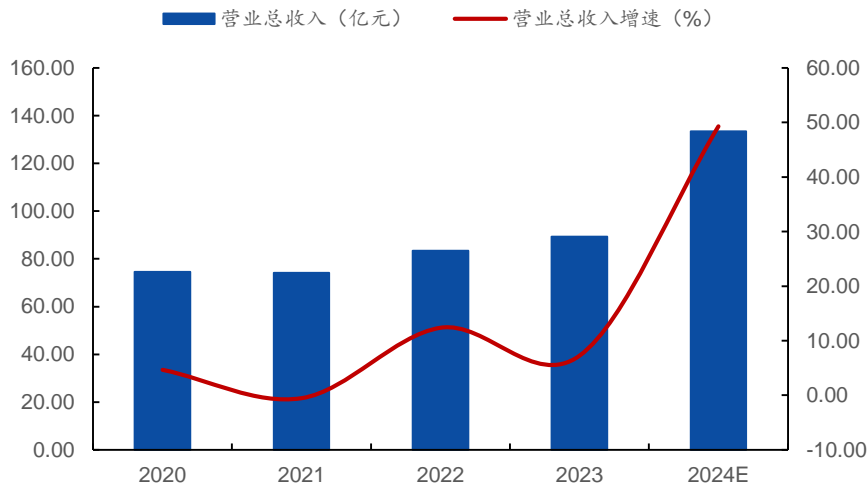
资料来源：华为encyclopedia, materialsnet, 太平洋证券

3.3 产业链相关公司：沪电股份

► **公司AI相关高端产能扩张，助力长期增长。**公司聚焦高多层、高频高速板业务，下游覆盖AI服务器、交换机等高增速行业，受益下游大客户需求增加，公司订单饱满。公司已在24年Q4规划新建人工智能芯片配套高端印制电路板扩产项目，生产高层高密度互连积层板，以满足高速运算服务器、人工智能等新兴计算场景对高端印制电路板的中长期需求，项目将分两阶段实施，投资总额预计约为43亿元人民币，计划年产约29万平方米人工智能芯片配套高端印制电路板。

图表：沪电股份营业收入及增速

图表：沪电股份归母净利润及增速

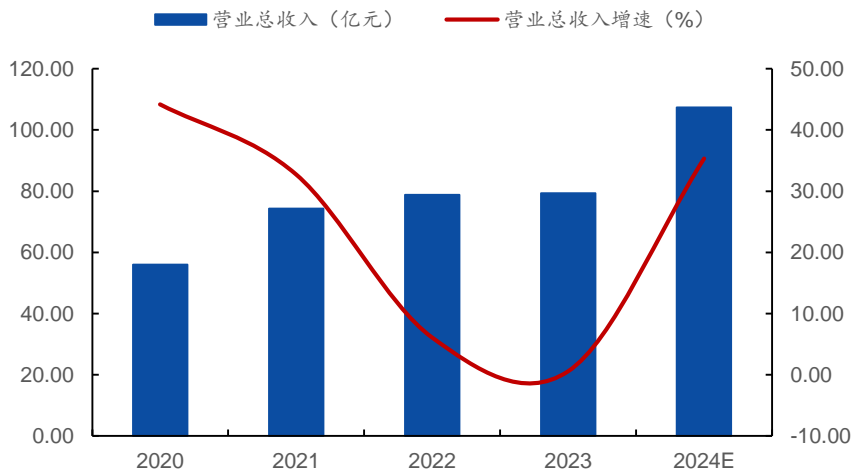


资料来源：公司公告，太平洋证券

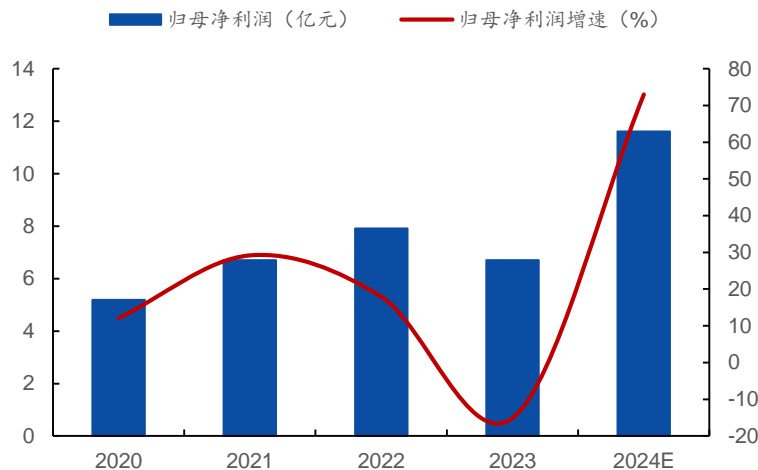
3.3 产业链相关公司：胜宏科技

►公司在高阶HDI+高多层均有布局，有望迎来放量。公司深耕行业20余年，终端客户覆盖NVIDIA、AMD、Intel、Tesla、微软、谷歌、思科等行业头部企业，充分受益下游客户需求爆发。公司目前在惠州、泰国和越南均有新的产能规划。惠州工厂会在现有产能基础上新扩50%HDI和30%高多层；泰国目前以多层板为主，电源和汽车产线率先投产，后续结合HDI工艺的复合型高多层产能也将投产。越南生产基地也在持续推进，规划布局HDI产品。

图表：胜宏科技营业收入及增速



图表：胜宏科技归母净利润及增速



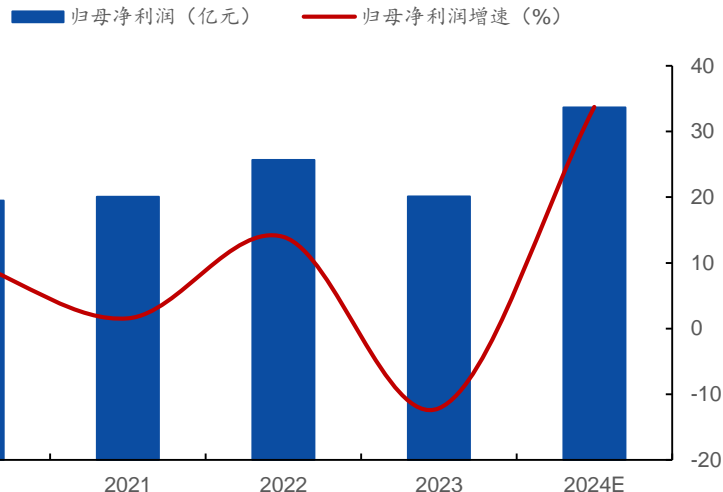
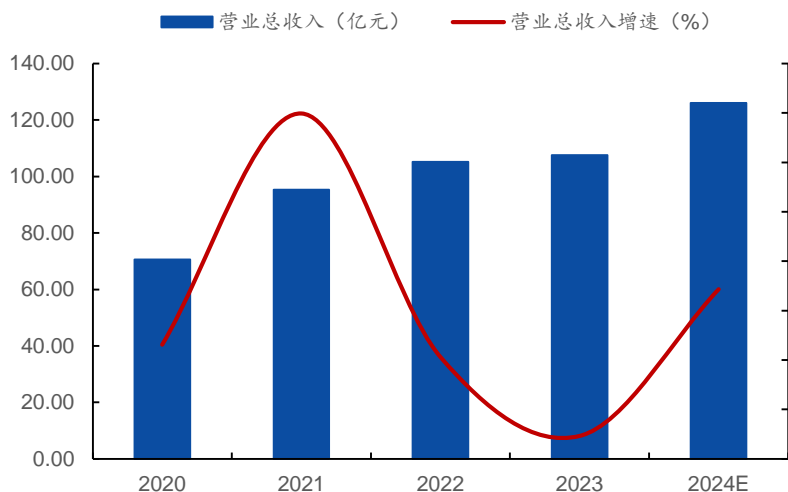
资料来源：公司公告，太平洋证券

3.3 产业链相关公司：景旺电子

►公司PTFE板业务进展显著，打开新增长曲线。在AI服务器领域，公司成功开拓软板和软硬结合板在数据中心的应用产品，同时在高阶HDI、高多层PTFE板等产品上实现了重大突破。在通用服务器领域，公司已实现EGS/Genoa平台高速PCB稳定量产，同时在Birth stream平台高速PCB等产品技术上取得重大突破。高速通信领域，公司800G光模块、通信模组高阶HDI等产品实现批量出货，112G交换路由PCB取得重大技术突破。

图表：景旺电子营业收入及增速

图表：景旺电子归母净利润及增速



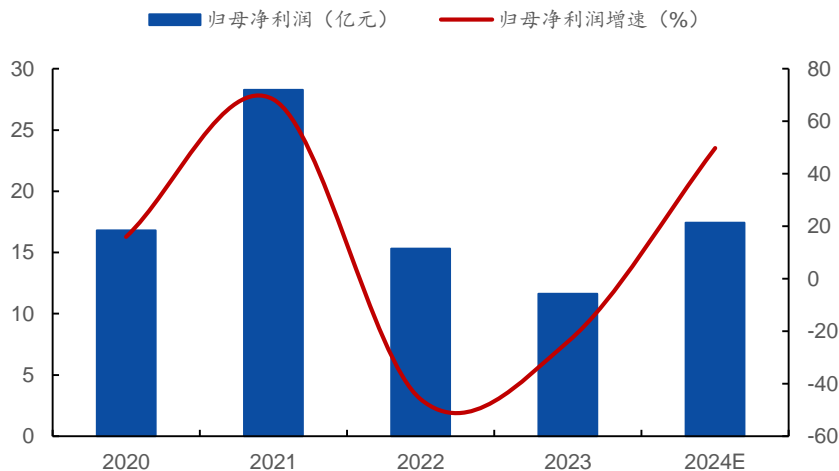
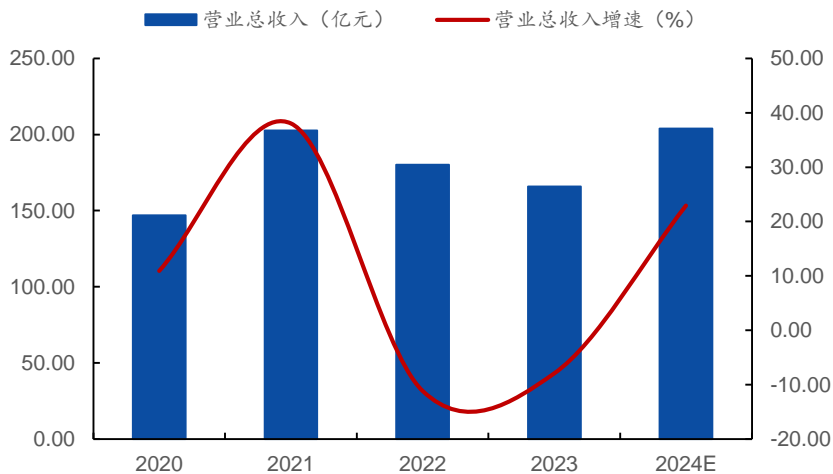
资料来源：公司公告，太平洋证券

3.3 产业链相关公司：生益科技

► **国产高端CCL头部厂商，有望迎来放量。** AI算力相关产品对信号传输速率和带宽形成新要求，承载信号通道的覆铜板材料需要更低损耗，产品形态也同以往有线产品有较大差异，配合客户提出的新架构和新的材料需求，在公司强大的管理体系，技术平台上，能够快速给予解决方案，在新品验证上给予及时交付。公司正积极同国内外各大终端就GPU和AI展开相关项目开发合作，并已有产品在批量供应。

图表：生益科技营业收入及增速

图表：生益科技归母净利润及增速



资料来源：公司公告，太平洋证券

风险提示

- 下游需求不及预期风险；
- 行业竞争加剧风险；
- 地缘政治风险。

投资评级说明

1、行业评级

看好：预计未来6个月内，行业整体回报高于沪深300指数5%以上；

中性：预计未来6个月内，行业整体回报介于沪深300指数-5%与5%之间；

看淡：预计未来6个月内，行业整体回报低于沪深300指数5%以下。

2、公司评级

买入：预计未来6个月内，个股相对沪深300指数涨幅在15%以上；

增持：预计未来6个月内，个股相对沪深300指数涨幅介于5%与15%之间；

持有：预计未来6个月内，个股相对沪深300指数涨幅介于-5%与5%之间；

减持：预计未来6个月内，个股相对沪深300指数涨幅介于-5%与-15%之间；

卖出：预计未来6个月内，个股相对沪深300指数涨幅低于-15%以下。

太平洋证券股份有限公司

云南省昆明市盘龙区北京路926号同德广场写字楼31楼



投诉电话： 95397

投诉邮箱： kefu@tpyzq.com

免责声明

太平洋证券股份有限公司（以下简称“我公司”或“太平洋证券”）具备中国证券监督管理委员会核准的证券投资咨询业务资格。

本报告仅向与太平洋证券签署服务协议的客户发布，为太平洋证券签约客户的专属研究产品，若您并非太平洋证券签约客户，请取消接收、订阅或使用本报告中的任何信息；太平洋证券不会因接收人收到、阅读或关注媒体推送本报告中的内容而视其为太平洋证券的客户。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机构和个人的投资建议，投资者应自主作出投资决策并自行承担投资风险，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。

本报告信息均来源于公开资料，我公司对这些信息的准确性和完整性不作任何保证。负责准备本报告以及撰写本报告的所有研究分析师或工作人员在此保证，本研究报告中关于任何发行商或证券所发表的观点均如实反映研究人员的个人观点。报告中的内容和意见仅供参考，并不构成对所述证券买卖的出价或询价。我公司及其雇员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。我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业务服务。本报告版权归太平洋证券股份有限公司所有，未经书面许可任何机构和个人不得以任何形式翻版、复制、刊登。任何人使用本报告，视为同意以上声明。