

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

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/135324203314012122>