



计算机行业研究

买入（维持评级）
行业专题研究报告

证券研究报告

计算机组

 分析师：刘高畅（执业 S1130525120005） 分析师：郑元昊（执业 S1130525120004） 分析师：孙恺祈（执业 S1130526060005）
liugaochang@gjzq.com.cn zhengyuanhao@gjzq.com.cn sunkaiqi@gjzq.com.cn

海外算力 Q3 迎来加速期

行业观点

- **2026 年 Q3 将成为海外算力核心加速节点，英伟达、谷歌、云厂商自研 ASIC 芯片集体进入交付窗口期，需求确定性持续强化。**英伟达 Vera Rubin 平台已全面投产，黄仁勋多次公开表态对冲延期传闻，该平台集成六类全新芯片，AI 训练性能较 Blackwell 提升 3.5 倍，配套 HBM4 内存带宽翻倍，2026 年 Rubin GPU 预计出货 570 万颗，同平台 Groq 3 LPU 同步 Q3 出货；公司 2027 财年 Q1 收入指引 780 亿美元，头部云厂商均规划批量部署 Rubin 设备。谷歌 TPU 产业链双线推进，七代 Ironwood 持续放量，全年 TPU 出货量预计 332.5 万颗领跑 ASIC 赛道，全新双芯片架构 TPU v8 分训练、推理两款芯片，下半年开启客户交付。OpenAI、AWS、Meta、微软同步完成新一代自研 ASIC 迭代，Jalapeno、Trainium3、MTIA400 等产品集中锁定 2026 年下半年部署窗口。博通、Marvell 包揽九成以上定制 AI 芯片协同设计市场，AI 半导体收入翻倍增长，机构预测 2028 年 ASIC 出货量将超越 GPU，全产业链下游算力硬件需求迎来 Q3 集中兑现。
- **英伟达 Rubin NVL144 机柜完成架构级革新，带动机柜、PCB、光模块三大硬件板块同步迎来价值重估。**整机柜层面采用全液冷模块化设计，取消传统线缆改用中央 PCB 中板，工业富联作为首发代工厂，Q2 小批量试产、Q3 启动规模化出货，单柜自供零部件比例提升，毛利率弹性显著。PCB 是本轮价值增幅最高环节，VR200 机柜相比 GB300 新增中板、网卡配套 PCB 等品类，单机柜 PCB 总价值暴涨；M9 及以上高端板材带动上游材料持续涨价。光模块主线切换至 1.6T 产品，Rubin 平台 800GB/s 网卡带宽升级倒逼交换机侧速率迭代，英伟达与康宁合作扩产光纤夯实光互连需求；NPO、CPO 尚处客户验证阶段，2026 下半年仅作为估值催化，2027 年才会实现规模化落地，高速连接器、MT 插芯等配套器件同步受益 1.6T 放量周期。
- **AI 机柜功率密度台阶式上涨成为液冷、电源、电容三大细分共同增长逻辑。**Rubin 代际 VR200 机柜功耗 190-230 千瓦，较上代提升超五成，下一代 Rubin Ultra 功耗将达 600 千瓦，风冷、传统低压配电体系全面触达瓶颈。散热端全液冷成为标配，模块化液冷母排、冷板、管路为机柜带来显著增量价值，行业渗透率与单机价值同步上行。供电端 800VDC 高压直流架构成为下一代数据中心标准方案，可降低转换损耗、缩减运维成本，富士康、CoreWeave 等头部厂商已落地 800V 数据中心，带动高压整流模组、DC/DC 电源柜、独立供电机架需求爆发。电容需求分层分化，官方明确 Rubin 机柜储能容量较上代提升 20 倍，大容量储能电容、BBU 备份单元增量最确定；GPU 侧末级供电电流提升推高 MLCC 板级去耦电容用量；800V 架构虽降低母线电流，但滤波、支撑电容规格同步迭代。三者共享功耗提升底层逻辑，同步受益 Q3 Rubin 机柜批量出货，是算力硬件中确定性较强的配套细分赛道。

相关标的

- 1) 海外算力：东山精密、工业富联、胜宏科技、中际旭创、江海股份、中钨高新、蓝思科技、东阳光、光智科技、先导基电、火炬电子、三环集团、欧科亿、天孚通信、鼎泰高科、新易盛、领益智造、兆易创新、鹏鼎控股、唯科科技、海川智能、天岳先进、大普微、源杰科技、麦格米特、景旺电子、英维克、京东方等；英特尔、SK 海力士、Lumentum、闪迪、铠侠、美光。
- 2) 国内算力：寒武纪、海光信息、东阳光、中芯国际、华虹半导体、禾盛新材、摩尔线程、天数智芯、沐曦股份、壁仞科技、协创数据、华丰科技、杰华特、京基智农、华达科技、中科曙光、利通电子、浪潮信息、利扬芯片、胜蓝股份、华勤技术、国科微、中国长城、晶科科技、罗曼股份、盈峰环境、芯原股份、晶科科技、亿田智能、豫能控股、星环科技、鸿日达、盛视科技、神州数码、润泽科技、大位科技、润建股份、奥飞数据、瑞晟智能、科华数据、潍柴重机、欧陆通、杰创智能、奥尼电子。
- 3) 大模型&云：智谱、MiniMax、阿里巴巴、腾讯控股、金山云、百度集团、优刻得、首都在线、网宿科技、云赛智联、青云科技等。

风险提示

- Rubin 量产爬坡不及预期的风险；ASIC 出货节奏低于预期的风险；数据中心基础设施改造进度的风险；贸易摩擦与出口管制的风险；第三方预测数据时效性的风险。



内容目录

一、海外算力 Q3 迎来加速器，出货节奏集中爆发.....	3
1.1 英伟达：Vera Rubin 全面投产、Q3 客户出货，管理层四地密集表态	3
1.2 谷歌：Ironwood 持续放量，TPU v8 双芯片战略下半年接棒	4
1.3 其他 ASIC 与 LPU：出货窗口集体锁定下半年	5
二、机柜、PCB 与光模块：主流大板块跟随机柜代际切换	7
2.1 服务器机柜：Rubin 整机柜量产启动，工业富联首发受益	7
2.2 PCB：中板替代线缆，机柜级 PCB 价值量系统性抬升	8
2.3 光模块：1.6T 放量主线明确，NPO 进入验证窗口	9
三、液冷、电源与电容：功耗跃升驱动的同源主线	10
3.1 功耗总逻辑：单柜功率密度跃升，散热与供电被迫同步重构	10
3.2 液冷：100%全液冷成为标配，价值量与渗透率双升	10
3.3 电源与电容：800VDC 架构确立，储能需求提升 20 倍	11
四、相关标的	12
五、风险提示	12

图表目录

图表 1：英伟达于 2026 年 1 月 CES 上对 Rubin 平台进行系统性发布	3
图表 2：英伟达在 3 月 GTC 上发布了整合 Groq 团队后的首款 LPU 芯片 Groq 3	4
图表 3：谷歌发布第八代 TPU，这是 TPU 项目十年历史上首次将训练与推理拆分为两颗独立芯片	5
图表 4：MTIA 四代产品路线图（MTIA 300/400/450/500）	6
图表 5：超大规模定制 ASIC 比较	6
图表 6：NVIDIA 的 GB300 计算盘采用了水冷，但由于布线和管道复杂，制造起来更为困难	7
图表 7：超大规模定制 ASIC 比较	8
图表 8：VR200 机柜新增多类 PCB 板，各品类 ASP 与用量双升，带动 PCB 价值量暴涨 233%	9
图表 9：MLCC 结构图	10
图表 10：英伟达 GW 级 AI 工厂演进路线	11
图表 11：英伟达机柜代际功耗与散热供电架构演进	11



一、海外算力 Q3 迎来加速器，出货节奏集中爆发

1.1 英伟达：Vera Rubin 全面投产、Q3 客户出货，管理层四地密集表态

Vera Rubin 已全面投产，Q3 进入客户出货窗口。英伟达于 2026 年 1 月 CES 上对 Rubin 平台进行系统性发布，该平台采用极端协同设计理念，整合六颗全新芯片，包括 Vera CPU、Rubin GPU、NVLink 6 交换芯片、ConnectX-9 SuperNIC、BlueField-4 DPU 以及 Spectrum-6 以太网交换芯片，覆盖计算、网络、存储与安全多个层级；相比前代 Blackwell 架构，Rubin 加速器 AI 训练性能提升 3.5 倍、运行性能提升 5 倍，并配备 88 核的新款 Vera CPU。6 月初，黄仁勋在中国台湾台北举行的 GTC 大会上宣布 Vera Rubin 进入全面生产，生产出货计划于 2026 年秋季启动；他同时透露，公司与全球 30 个国家超过 350 家工厂协作，其中 150 家位于中国台湾，Vera Rubin 的供应链规模达到 Blackwell 的两倍。

图表1：英伟达于 2026 年 1 月 CES 上对 Rubin 平台进行系统性发布



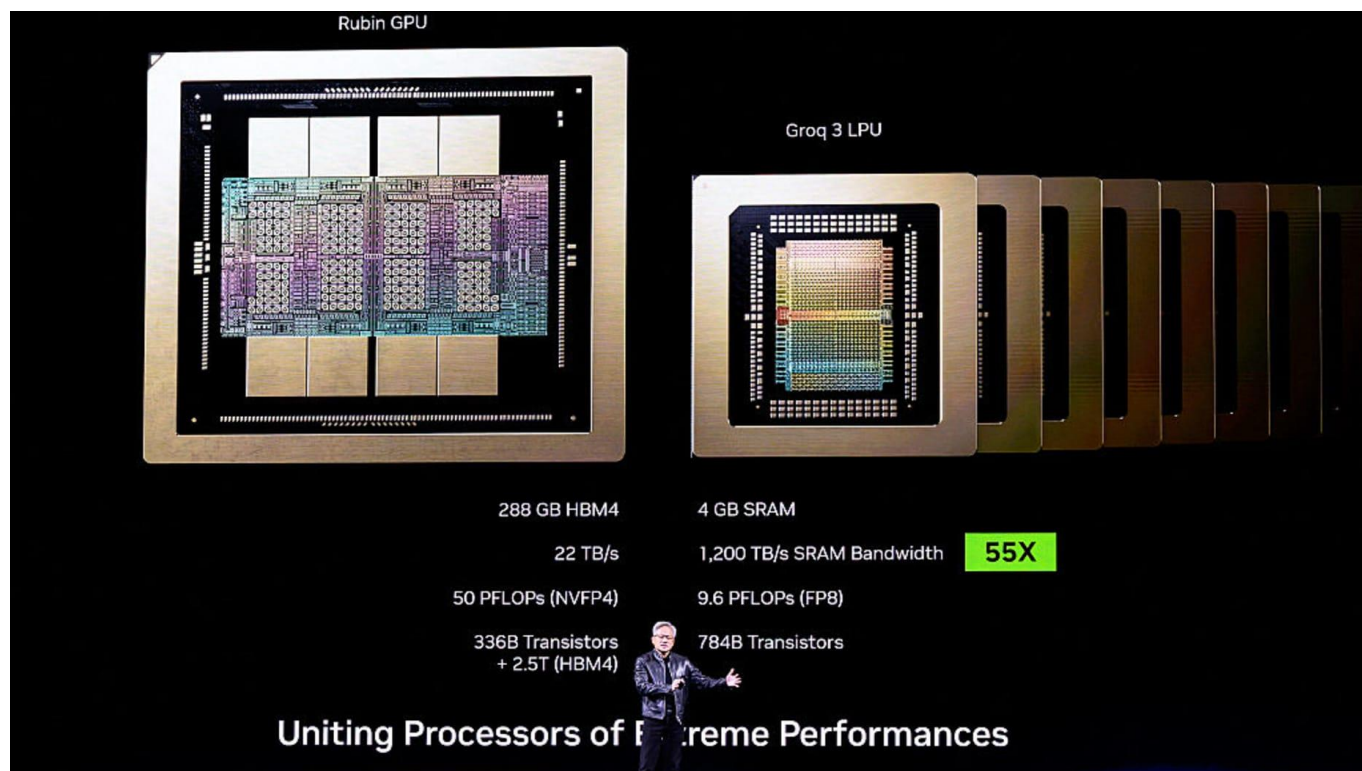
来源：21 世纪经济网，国金证券研究所

黄仁勋四个月内四次递进表态，出货确定性持续且修。6 月 6 日 GTC 大会上，黄仁勋表示 Blackwell 与 Vera Rubin 两代平台至 2027 年的采购订单预计达到 1 万亿美元，高于上年给出的 5000 亿美元收入机会展望；Rubin 系统由 130 万个零部件构成，性能功耗比达上代 Grace Blackwell 的 10 倍。6 月 5 日，黄仁勋在首尔确认三星电子、SK 海力士、美光三家 HBM4 供应商全部通过认证并处于量产状态，全力支持 Rubin 平台的交付时间表；HBM4 将内存接口位宽从 1024bit 翻倍至 2048bit，单堆栈带宽达到 2TB/s 以上。7 月 15 日，针对 SemiAnalysis 有关英伟达下一代机柜系统因生产问题延期的传闻，黄仁勋在东京直接回应称 Vera Rubin 正在生产中、按计划向客户交付，并表示巨量产能正在到来。管理层用持续加码的公开表态对冲市场噪音，Rubin 的 Q3 出货节奏获得反复确认。

财务口径同步印证需求强度。英伟达 2026 财年（截至 2026 年 1 月 25 日）实现收入 2159 亿美元，同比增长 65%，其中数据中心收入 1937 亿美元，同比增长 68%；公司指引 2027 财年第一季度收入 780 亿美元，同比增长约 77%。CFO Colette Kress 在财报会上表示，预计每一家云端模型厂商都将部署 Vera Rubin，样片已开始向客户交付，量产计划于下半年启动。第三方分析师预计 2026 年 Rubin GPU 出货量约 570 万颗。此外，英伟达在 3 月 GTC 上发布了整合 Groq 团队后的首款 LPU 芯片 Groq 3，同样计划于第三季度出货，从 GPU 到 LPU 的产品矩阵在 Q3 同步放量，进一步强化了出货节奏集中爆发的判断。



图表2：英伟达在3月GTC上发布了整合Groq团队后的首款LPU芯片Groq 3



来源：CNBC，国金证券研究所

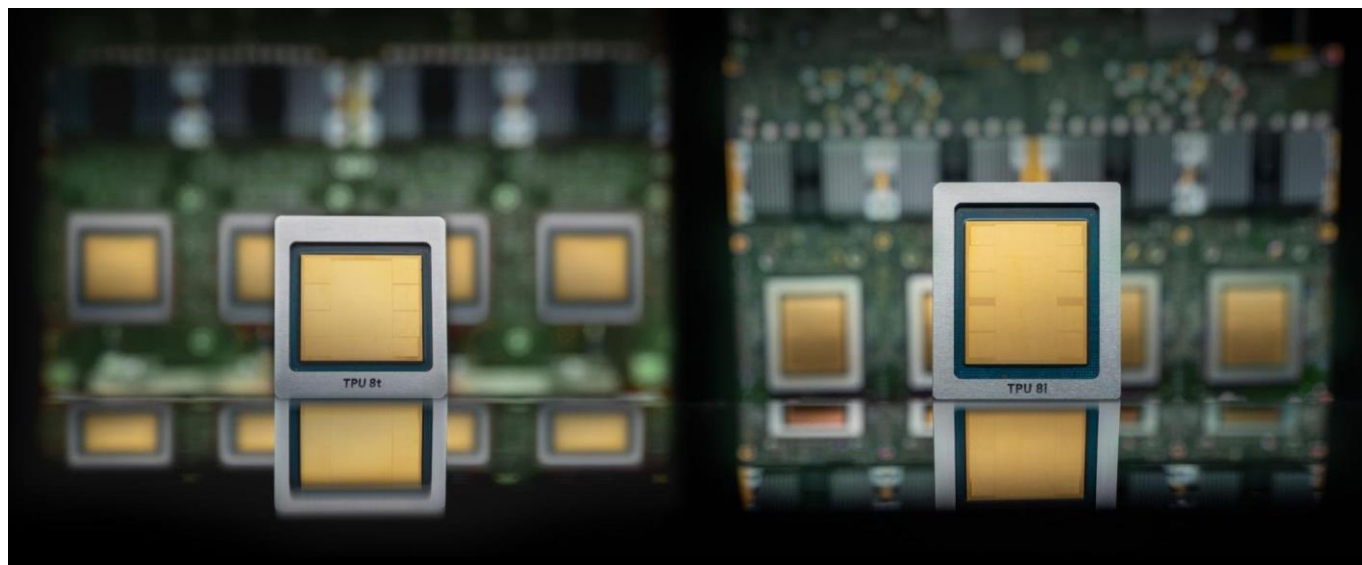
1.2 谷歌：Ironwood 持续放量，TPU v8 双芯片战略下半年接棒

谷歌 TPU 出货量 2026 年跃升，稳居 ASIC 阵营第一。据 DIGITIMES Research 统计，2026 年谷歌 TPU 出货量预计达 332.5 万颗，远超 AWS 的 147 万颗、Meta 的 60 万颗和微软的 33 万颗；ASIC 服务器 2026 年出货量增速预计达 64.2%，超过 GPU 服务器的 43.8%。当前承担放量任务的是第七代 TPU Ironwood：单芯片算力 4.6 PFLOPs、搭载 192GB HBM3E，可扩展至 9216 颗芯片的超级 Pod、合计 42.5 ExaFLOPS 算力。

TPU v8 开创双芯片战略，训练推理首次分立。谷歌于 4 月 22 日 Cloud Next 大会上发布第八代 TPU，这是 TPU 项目十年历史上首次将训练与推理拆分为两颗独立芯片：TPU 8t（代号 Sunfish）由博通参与设计、面向大规模训练；TPU 8i（代号 Zebrafish）由联发科参与设计、面向低时延推理，采用全新 Boardfly 互连拓扑。据 Tom's Hardware 报道，两款芯片均采用台积电 N3 工艺家族、搭载 HBM3E，并将于年内向谷歌云客户开放。联发科的加入终结了博通自 2015 年以来对 TPU 设计的独家参与，谷歌的多设计伙伴策略同时带来议价能力与供应链冗余。需要提示的是，市场对 v8 量产上量的具体时点仍存在口径分歧，部分供应链信息指向更晚的规模化节点，我们建议以 Ironwood 放量为基本盘、v8 下半年进入交付窗口为弹性项来把握谷歌链条的节奏。



图表3：谷歌发布第八代 TPU，这是 TPU 项目十年历史上首次将训练与推理拆分为两颗独立芯片



来源：Google Cloud blog，国金证券研究所

TPU 代际切换对硬件外围的带动路径清晰。TPU v8 代际将为 OCS 全光交换机、液冷、电源及光通信等外围环节创造一系列升级机会；谷歌自研的 OCS 光路交换架构是其区别于英伟达链条的独立增量，直接受益于 TPU 集群规模扩张。TPU 服务器的机柜组装环节由中国台湾 ODM 厂商主导，广达、富士康、纬颖、纬创、英业达等均在积极争取 ASIC 服务器订单，ASIC 服务器单价低于 GPU 服务器但定制化设计灵活度更高、对 ODM 的利润率更为友好。

1.3 其他 ASIC 与 LPU：出货窗口集体锁定下半年

OpenAI 首颗自研芯片样品已交付，机柜部署下半年启动。OpenAI 与博通于 2025 年 10 月 13 日官宣合作部署 10GW 自研 AI 加速器，OpenAI 负责加速器与系统设计、博通负责开发与部署，机柜级部署目标自 2026 年下半年启动、2029 年底前完成。2026 年 6 月 24 日，双方公布首颗芯片 Jalapeno 并交付实物样品，目标 2026 年底实现初步部署；博通 CEO Hock Tan 表示 2026 年底为小规模原型部署，2027 年开始爬坡，2028 年上半年全速放量。OpenAI 总裁 Greg Brockman 透露，该芯片从设计到交付仅用九个月完成端到端设计。

AWS、Meta、微软的自研芯片同步处于代际切换期。AWS 方面，Trainium3 是 AWS 首颗 3nm 芯片，每个 AWS Trainium3 芯片提供 2.52 千万亿次浮点 (PFLOP) 的 FP8 计算，与 Trainium2 相比，内存容量增加 1.5 倍，带宽增加 1.7 倍，达到 144 GB 的 HBM3e 内存，以及 4.9 TB/s 的内存带宽。Trainium3 专为密集型和专家并行工作负载而设计，具有高级数据类型 (MXFP8 和 MXFP4)，并改善了实时、多模态和推理任务的内存与计算平衡。下一代 Trainium4 已于 2025 年 12 月官宣，计划 2026 年底或 2027 年初可用。Meta 方面，公司于 2026 年 3 月通过官方技术博客罕见公开 MTIA 四代产品路线图 (MTIA 300/400/450/500)，其中 MTIA400 已完成测试、进入生产部署；据 TrendForce 援引工商时报信息，MTIA-2 已投产并计划 2026 年上半年亮相，MTIA-3 计划 2026 年下半年亮相，均采用台积电 3nm 工艺。微软方面，Maia 200 已从得梅因数据中心承接 OpenAI GPT-5.2 模型与 Microsoft 365 Copilot 推理负载，微软称其为迄今性能最强的超大规模厂商第一方芯片、性价比较现有有机群最优硬件提升 30%。



图表4: MTIA 四代产品路线图 (MTIA 300/400/450/500)

Metric	MTIA 300	MTIA 400	MTIA 450	MTIA 500
Workload Focus	R&R Training	General	GenAI Inference	GenAI Inference
Module TDP	800 W	1200 W	1400 W	1700 W
HBM Bandwidth	6.1 TB/s	9.2 TB/s	18.4 TB/s	27.6 TB/s
HBM Capacity	216 GB	288 GB	288 GB	384-512 GB
MX4 Performance	—	12 PFLOPs	21 PFLOPs	30 PFLOPs
FP8/MX8 Performance	1.2 PFLOPs	6 PFLOPs	7 PFLOPs	10 PFLOPs
BF16 Performance	0.6 PFLOPs	3 PFLOPs	3.5 PFLOPs	5 PFLOPs
Scale-up Domain Size	16	72	72	72
Scale-up Network (unidirectional bandwidth*)	1 TB/s	1.2 TB/s	1.2 TB/s	1.2 TB/s
Scale-out Network (unidirectional bandwidth*)	200 GB/s**	100 GB/s	100 GB/s	100 GB/s

*Some vendors report bidirectional bandwidth. Multiply the value in the table by two to obtain the corresponding bidirectional bandwidth.

**MTIA 300 is configured with a scale-out network with higher bandwidth (200 GB/s) due to its relatively small scale-up domain size and the target R&R workloads.

来源: chipstrat, 国金证券研究所

图表5: 超大规模定制 ASIC 比较

Company	Custom Chip	Purpose	Design Partner	Status (2026)
Google	TPU v7 + v8 (Trillium)	Training + Inference	Broadcom + MediaTek	✅ In Production
Meta	MTIA 400	Internal inference	TSMC (self-designed)	✅ Entering Deploy
Amazon	Trainium 3	Training + Inference	Marvell + internal	➡️ Ramping Q2 2026
Microsoft	Maia 100	Azure inference	Broadcom	✅ In Production
OpenAI	Titan (in dev)	Training frontier	Broadcom	📄 Dev Phase



来源: oplexa, 国金证券研究所

设计与代工环节的景气度为出货节奏提供交叉验证。博通 2026 财年第一财季 AI 半导体收入 84 亿美元、同比增长 106%，第二财季指引 107 亿美元；6 月 3 日第二财季财报会上，公司形容 AI 芯片需求处于 insatiable（永不满足）状态，并披露谷歌、OpenAI、Meta、Anthropic、字节跳动等六大定制芯片核心客户的进展；Hock Tan 此前表示 2027 年仅芯片的 AI 收入就有望超过 1000 亿美元。Marvell 预计 2026 年 AI ASIC 收入最高达 110 亿美元，博通与 Marvell 合计占据定制 AI ASIC 协同设计市场约 95% 份额。更长维度看，Counterpoint 预计 2028 年全球数据中心 AI ASIC 出货量将突破 1500 万颗、历史性超越数据中心 GPU 出货量。Anthropic 的自研芯片计划仍处早期，作为博通六大客户之一暂无明确出货时点。综合来看，英伟达 Rubin 与 Groq 3 LPU、谷歌 TPU、OpenAI Jalapeno、AWS Trainium3、Meta MTIA 新代际的出货或部署窗口全部落在 2026 年下半年，海外算力需求在 Q3 起进入集中爆发阶段。

二、机柜、PCB 与光模块：主流大板块跟随机柜代际切换

2.1 服务器机柜：Rubin 整机柜量产启动，工业富联首发受益

Vera Rubin NVL144 机柜是一次架构级重构而非常规升级。据英伟达官方博客，Vera Rubin NVL144 计算托盘采用 100% 液冷的模块化设计，PCB 中央板（midplane）取代了传统的线缆连接方式，装配与维护效率显著提升。ConnectX-9 800GB/s 网络与 Rubin CPX 海量上下文推理模组的扩展仓位；机柜设计包含 45 摄氏度液冷、全新液冷母排以及较上代提升 20 倍的机柜储能。该机柜基于 MGX 架构，将获得超过 50 家 MGX 系统与零部件伙伴支持，英伟达并计划将升级后的机柜与计算托盘设计作为开放标准贡献给 OCP 开放计算项目。标准化与模块化的意义在于，产业链各环节可以并行开发、快速上量，这正是 Q3 出货节奏得以兑现的架构基础。

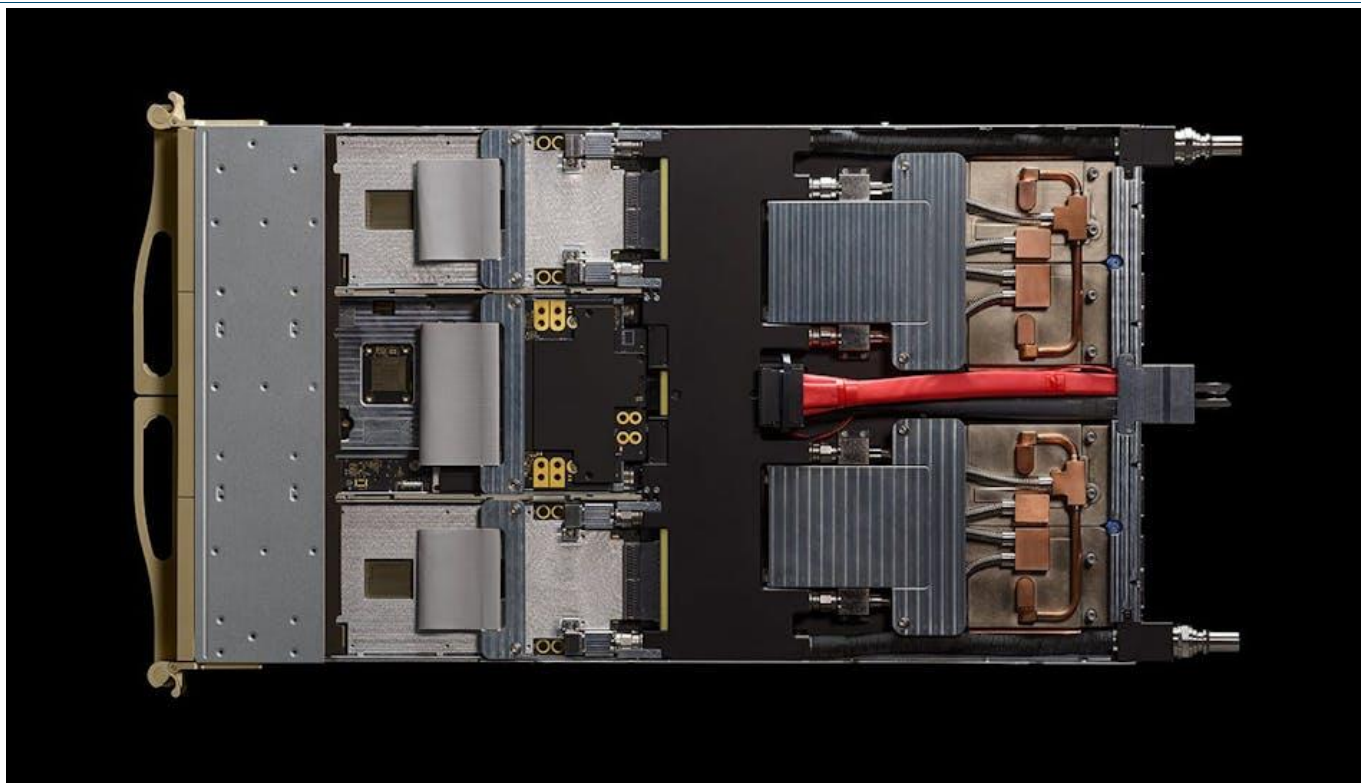
图表6: NVIDIA 的 GB300 计算盘采用了水冷，但由于布线和管道复杂，制造起来更为困难



来源: Electronic Design, 国金证券研究所



图表7：超大规模定制 ASIC 比较



来源：Electronic Design，国金证券研究所

工业富联是 Rubin 整机柜的首发制造商，率先卡位出货窗口。工业富联在英伟达 AI 服务器代工体系中份额领先，深度参与 GB200、GB300 及 Rubin 平台；根据公司投资者关系活动披露口径，Rubin 平台自 2026 年 6 月起小批量出货客户测试，公司自 Q2 起进行小批量试产、开展制造验证，规模化放量预计在 2027 年上半年。英伟达 2026 年服务器整机出货指引中 Rubin 机柜约 2.3 万台，预计 2026 年 Q3、Q4 开始生产，工业富联作为首发供应商有望占据主要份额；同时 Rubin 机柜的液冷冷板、管路、结构件、电源等核心零部件自供比例较 GB 系列进一步提升，单台加工利润与毛利率弹性同步放大。整机柜环节的投资逻辑是量价利三击：Q3 量产启动带来量的拐点，单柜价值量较 GB300 大幅抬升带来价的弹性，自供率提升带来利润率的上修。

2.2 PCB：中板替代线缆，机柜级 PCB 价值量系统性抬升

PCB 是 Rubin 代际中价值量提升逻辑最硬的环节之一。英伟达官方明确，Vera Rubin NVL144 以中央 PCB 中板替代传统线缆连接，这意味着机柜内部互连从线缆方案向大尺寸、高层数 PCB 方案迁移，属于从无到有的增量价值；叠加计算托盘、交换托盘、ConnectX-9 与 Rubin CPX 模组板的高多层与 HDI 用量，单机柜 PCB 价值量较 GB 代际系统性抬升。展望下一代，Rubin Ultra（2027 年）整机柜正交背板方案有望将 PCB 价值量曲线进一步推陡，延续我们此前提出的斜率之王叙事——每一次机柜代际切换，PCB 都是单位价值量增幅居前的环节。


图表8: VR200 机柜新增多类 PCB 板, 各品类 ASP 与用量双升, 带动 PCB 价值量暴涨 233%

PCB ASP per board (US\$)	GB300	VR200
Compute PCB	\$650	\$1,400
Switch PCB	\$800	\$1,450
Midplane PCB	\$0	\$1,500
BlueField PCB	\$0	\$255
ConnectX PCB	\$0	\$270
Other peripheral PCB	\$50	\$50
PCB Units per rack	GB300	VR200
Compute PCB	36x	36x
Switch PCB	9x	9x
Midplane PCB	0x	18x
BlueField PCB	18x	18x
ConnectX PCB	0x	72x
Other peripheral PCB	90x	45x
Total PCB Content per rack	GB300	VR200
Compute PCB	\$23,400	\$50,400
Switch PCB	\$7,200	\$13,050
Midplane PCB	\$0	\$27,000
BlueField PCB	\$0	\$4,590
ConnectX PCB	\$0	\$19,440
Other peripheral PCB	\$4,500	\$2,250
Total PCB Content per rack	\$35,100	\$116,730

来源: 新浪财经, 国金证券研究所

高端产能与材料环节同步受益, 涨价与缺口并存。高多层板、HDI 所需的高端覆铜板、低轮廓铜箔等上游材料在 2026 年至 2028 年存在结构性供给缺口, 材料端涨价向板厂传导的过程中, 具备海外大客户交付能力的头部板厂议价能力增强。1.6T 光模块配套的高阶 PCB 亦对材料与工艺提出更高要求, 相关口径待进一步取证后在后续报告中展开。需要提示的是, PCB 环节的份额分配仍存在不确定性, 海外交付能力与高阶产品良率是决定个股弹性的核心变量。

2.3 光模块: 1.6T 放量主线明确, NPO 进入验证窗口

Rubin 代际将 1.6T 光模块推入放量通道。Vera Rubin NVL144 预留 ConnectX-9 800GB/s 网络扩展仓位, 网卡侧带宽翻倍直接牵引交换机侧与光模块侧的速率升级, 1.6T 光模块在 Rubin 集群组网中进入规模部署阶段, 800G 向 1.6T 的切换斜率决定 2026 年下半年至 2027 年光模块板块的业绩弹性。基础设施投资同步印证光互连景气度: 据 21 世纪经济报道, 当地时间 5 月 6 日, 英伟达与康宁宣布达成多年期商业与技术合作伙伴关系。根据协议, 康宁计划将其在美国的光连接制造能力提升 10 倍, 光纤产量提升超 50%, 并在北卡罗来纳州和得克萨斯州新建三座先进制造工厂。龙头芯片厂商直接下场投资光纤产能, 侧面反映 AI 集群对光互连需求的紧迫程度。

NPO 与 CPO 处于验证窗口, 2027 年才是更关键的采用周期。在可插拔光模块持续放量的同时, NPO (近封装光学) 与 CPO (共封装光学) 正处于工程样机与客户验证阶段, 英伟达路径与博通、Marvell 白盒路径出现分化; 谷歌侧的 OCS 全光交换则是独立于英伟达链条的额外增量。延续我们在光通信系列报告中的判断: CPO/NPO 的规模化采用节点在 2027 年及以后, 2026 年下半年的主线仍是 1.6T 可插拔放量, NPO 作为技术储备与估值催化看待, 无需在当期业绩中过度定价。光器件与连接器环节 (FAU、MT/MPO 插芯、高速连接器等) 跟随



1.6T 与 NPO 双线演进，是弹性排序中容易被忽视的细分。

三、液冷、电源与电容：功耗跃升驱动的同源主线

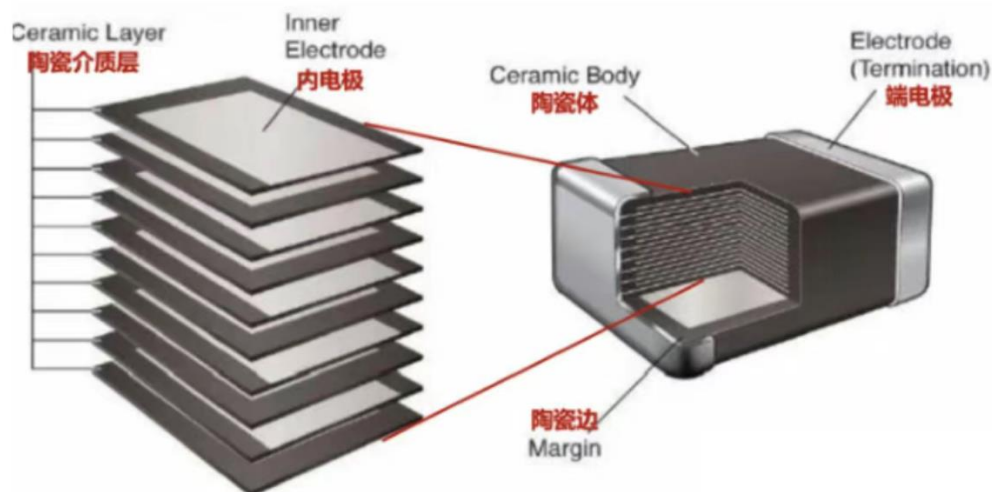
3.1 功耗总逻辑：单柜功率密度跃升，散热与供电被迫同步重构

功耗是理解液冷、电源、电容三个环节的同一把钥匙。据产业链信息，Rubin 代际 VR200 机柜功耗约 190 至 230 千瓦，较 GB300 代际的约 120 至 130 千瓦提升超过五成，较 Hopper 代际的约 40 千瓦提升数倍；2027 年 Rubin Ultra Kyber 机柜功耗将达约 600 千瓦，1MW 级机柜已在路线图上。单机柜功率密度的台阶式跃升，使传统风冷散热与 415/480V 交流配电体系双双触及物理极限，散热体系向全液冷迁移、供电体系向 800VDC 高压直流迁移，成为 Rubin 及后续代际的确定性方向。液冷、电源、电容并非三条独立逻辑，而是同一条功耗曲线在散热端与供电端的两个投影，这是本章将三者合并论述的原因。

3.2 液冷：100%全液冷成为标配，价值量与渗透率双升

MLCC（多层陶瓷电容）在主板/基板层面承担纳秒级的高频去耦与滤波，对应存储层级里的“SRAM 缓存”——数量庞大、密集分布在 GPU 周边。AI GPU 服务器用 MLCC 的价值量明显高于消费电子和传统 CPU 服务器，对高可靠、高性能产品的需求显著提升。但 MLCC 存在压电效应、温度特性及机械脆性等局限，在部分高容值、高可靠场景下并不完全适合，这为下一层 MLPC 留出了替代空间。

图表9：MLCC 结构图



来源：财联学研投微信公众平台，国金证券研究所

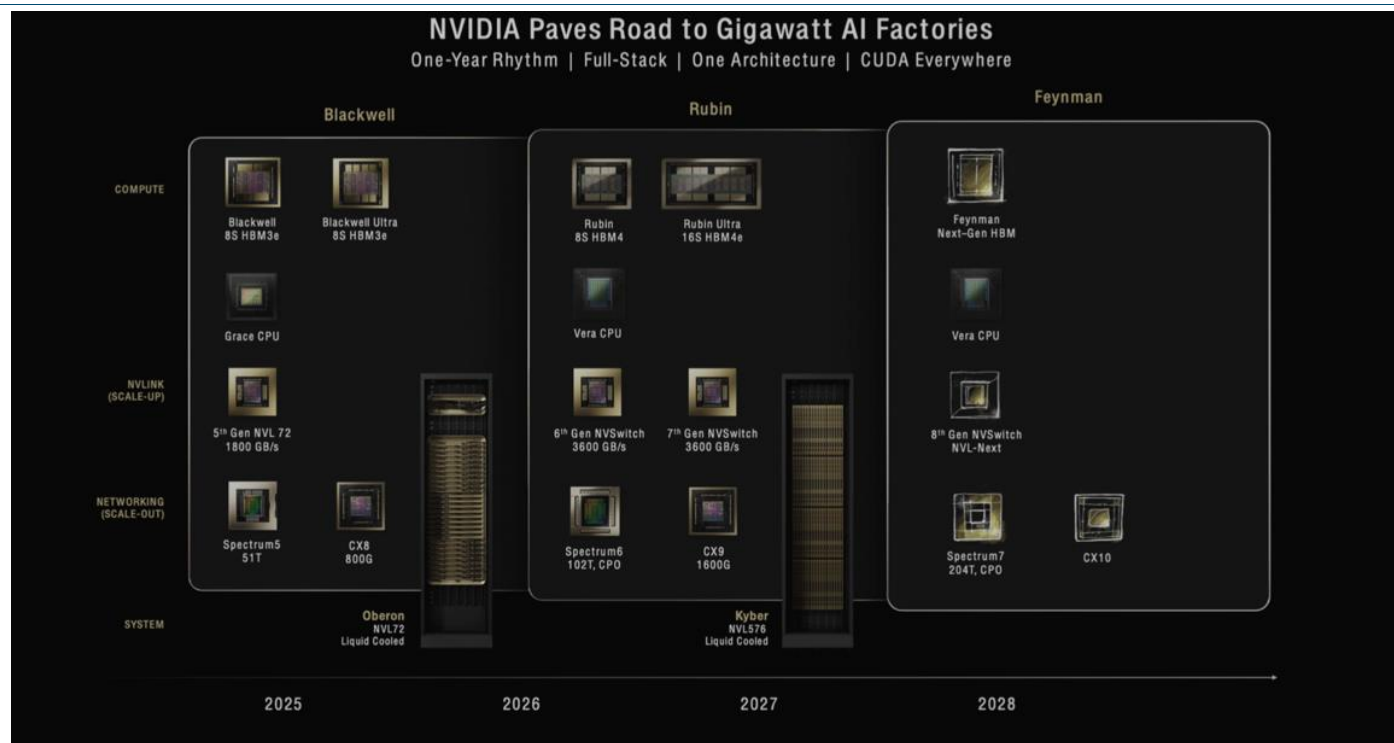


3.3 电源与电容：800VDC 架构确立，储能需求提升 20 倍

800VDC 高压直流确立为下一代数据中心参考架构。英伟达已将 800VDC 确立为面向 Rubin Ultra Kyber 机柜体系的参考架构，从传统 415/480V 三相交流向 800VDC 迁移，可消除机柜级 AC/DC 转换环节、端到端效率最高提升 5%、维护成本最高下降 70%。生态推进速度超预期：富士康披露其位于中国台湾高雄的 40MW 数据中心 Kaohsiung-1 按 800VDC 建设；CoreWeave、Lambda、Oracle 云等多家云厂商已在设计 800V 数据中心；光宝科技展出 800VDC Power Rack 与 110 千瓦 Power Shelf，供电架构从 power shelf 向独立 power rack 形态演进。电源环节的增量来自三个层面：机柜级 800VDC 电源柜与整流模组、机柜内的高压 DC/DC 变换、以及 UPS/BBU 体系的重构，单柜电源价值量随功耗与架构复杂度同步上行。

电容的正确打开方式是分层拆解，储能层是最硬的官方口径。英伟达官方明确，Vera Rubin NVL144 机柜储能较上代提升 20 倍，以平抑 GPU 负载剧烈波动带来的供电扰动，机柜级储能（大容量电容与 BBU 电池备份单元）是 Rubin 代际电容需求最确定的增量层。需要辨析的是，800VDC 通过提高电压降低了配电轨上的电流，这一层面的载流与滤波需求趋缓；但 GPU 侧末级供电的电流规模仍随芯片功耗刚性上升，板级去耦电容（MLCC）的用量与规格随功率密度同步增长；输入侧滤波、母线支撑等中间层级则随电压平台切换出现规格迁移。分层看，储能层增量最确定、板级去耦层跟随功耗增长、配电层结构性迁移，避免将高压化简单线性外推为所有电容环节同幅受益。

图表10：英伟达 GW 级 AI 工厂演进路线



来源：英伟达官方博客，国金证券研究所

图表11：英伟达机柜代际功耗与散热供电架构演进

代际	功耗与散热供电架构
Hopper 代际	约 40kW，风冷为主
GB300 代际	约 120-130kW，液冷渗透提升
VR200 (2026 H2)	约 190-230kW，100%液冷、45°C、800VDC 生态起步


Rubin Ultra Kyber (2027)

约 600kW, 800VDC 参考架构、双机柜占位 (算力柜+供电散热 sidecar)

来源：英伟达官方博客，Datacenter Dynamics 等，国金证券研究所

四、相关标的

1) 海外算力：东山精密、工业富联、胜宏科技、中际旭创、江海股份、中钨高新、蓝思科技、东阳光、光智科技、先导基电、火炬电子、三环集团、欧科亿、天孚通信、鼎泰高科、新易盛、领益智造、兆易创新、鹏鼎控股、唯科科技、海川智能、天岳先进、大普微、源杰科技、麦格米特、景旺电子、英维克、京东方等；英特尔、SK 海力士、Lumentum、闪迪、铠侠、美光。

2) 国内算力：寒武纪、海光信息、东阳光、中芯国际、华虹半导体、禾盛新材、摩尔线程、天数智芯、沐曦股份、壁仞科技、协创数据、华丰科技、杰华特、京基智农、华达科技、中科曙光、利通电子、浪潮信息、利扬芯片、胜蓝股份、华勤技术、国科微、中国长城、晶科科技、罗曼股份、盈峰环境、芯原股份、晶科科技、亿田智能、豫能控股、星环科技、鸿日达、盛视科技、神州数码、润泽科技、大位科技、润建股份、奥飞数据、瑞晟智能、科华数据、潍柴重机、欧陆通、杰创智能、奥尼电子。

3) 大模型&云：智谱、MiniMax、阿里巴巴、腾讯控股、金山云、百度集团、优刻得、首都在线、网宿科技、云赛智联、青云科技等。

五、风险提示

■ Rubin 量产爬坡不及预期的风险

Rubin 机柜零部件数量与集成复杂度大幅提升，若量产良率、HBM4 供给或液冷部件配套出现瓶颈，实际出货量可能低于指引。

■ ASIC 出货节奏低于预期的风险

谷歌 TPU v8 的规模化上量时点在市场上存在口径分歧，OpenAI Jalapeno 等新入者芯片处于首次量产爬坡期，实际部署节奏可能慢于官宣目标。

■ 数据中心基础设施改造进度的风险

100%液冷与 800VDC 架构对存量数据中心提出重建级要求，若下游液冷机房与高压配电改造进度慢于预期，机柜出货与安装节奏可能受到牵制。

■ 贸易摩擦与出口管制的风险

中美科技领域政策变化可能影响产业链公司的订单获取、原材料采购与产品交付。

■ 第三方预测数据时效性的风险

本报告引用的出货量预测、订单展望等数据来自公司表态与第三方研究机构，AI 硬件领



域数据更新迭代快，相关数字可能随时间推移发生显著修正。



行业投资评级的说明:

买入：预期未来 3—6 个月内该行业上涨幅度超过大盘在 15%以上；
增持：预期未来 3—6 个月内该行业上涨幅度超过大盘在 5%—15%；
中性：预期未来 3—6 个月内该行业变动幅度相对大盘在 -5%—5%；
减持：预期未来 3—6 个月内该行业下跌幅度超过大盘在 5%以上。



特别声明：

国金证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本报告版权归“国金证券股份有限公司”（以下简称“国金证券”）所有，未经事先书面授权，任何机构和个人均不得以任何方式对本报告的任何部分制作任何形式的复制、转发、转载、引用、修改、仿制、刊发，或以任何侵犯本公司版权的其他方式使用。经过书面授权的引用、刊发，需注明出处为“国金证券股份有限公司”，且不得对本报告进行任何有悖原意的删节和修改。

本报告的产生基于国金证券及其研究人员认为可信的公开资料或实地调研资料，但国金证券及其研究人员对这些信息的准确性和完整性不作任何保证。本报告反映撰写研究人员的不同设想、见解及分析方法，故本报告所载观点可能与其他类似研究报告的观点及市场实际情况不一致，国金证券不对使用本报告所包含的材料产生的任何直接或间接损失或与此有关的其他任何损失承担任何责任。且本报告中的资料、意见、预测均反映报告初次公开发布时的判断，在不作事先通知的情况下，可能会随时调整，亦可因使用不同假设和标准、采用不同观点和分析方法而与国金证券其它业务部门、单位或附属机构在制作类似的其他材料时所给出的意见不同或者相反。

本报告仅为参考之用，在任何地区均不应被视为买卖任何证券、金融工具的要约或要约邀请。本报告提及的任何证券或金融工具均可能含有重大的风险，可能不易变卖以及不适合所有投资者。本报告所提及的证券或金融工具的价格、价值及收益可能会受汇率影响而波动。过往的业绩并不能代表未来的表现。

客户应当考虑到国金证券存在可能影响本报告客观性的利益冲突，而不应视本报告为作出投资决策的唯一因素。证券研究报告是用于服务具备专业知识的投资者和投资顾问的专业产品，使用时必须经专业人士进行解读。国金证券建议获取报告人员应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。报告本身、报告中的信息或所表达意见也不构成投资、法律、会计或税务的最终操作建议，国金证券不就报告中的内容对最终操作建议做出任何担保，在任何时候均不构成对任何人的个人推荐。

在法律允许的情况下，国金证券的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，并可能为这些公司正在提供或争取提供多种金融服务。

本报告并非意图发送、发布给在当地法律或监管规则下不允许向其发送、发布该研究报告的人员。国金证券并不因收件人收到本报告而视其为国金证券的客户。本报告对于收件人而言属高度机密，只有符合条件的收件人才能使用。根据《证券期货投资者适当性管理办法》，本报告仅供国金证券股份有限公司客户中风险评级高于 C3 级（含 C3 级）的投资者使用；本报告所包含的观点及建议并未考虑个别客户的特殊状况、目标或需要，不应被视为对特定客户关于特定证券或金融工具的建议或策略。对于本报告中提及的任何证券或金融工具，本报告的收件人须保持自身的独立判断。使用国金证券研究报告进行投资，遭受任何损失，国金证券不承担相关法律责任。

若国金证券以外的任何机构或个人发送本报告，则由该机构或个人为此发送行为承担全部责任。本报告不构成国金证券向发送本报告机构或个人的收件人提供投资建议，国金证券不为此承担任何责任。

此报告仅限于中国境内使用。国金证券版权所有，保留一切权利。

上海

电话：021-80234211

邮箱：researchsh@gjzq.com.cn

邮编：201204

地址：上海浦东新区芳甸路 1088 号

紫竹国际大厦 5 楼

北京

电话：010-85950438

邮箱：researchbj@gjzq.com.cn

邮编：100005

地址：北京市东城区建国内大街 26 号

新闻大厦 8 层南侧

深圳

电话：0755-86695353

邮箱：researchsz@gjzq.com.cn

邮编：518000

地址：深圳市福田区金田路 2028 号皇岗商务中心

18 楼 1806



【小程序】
国金证券研究服务



【公众号】
国金证券研究