



通信行业研究

买入（维持评级）

行业周报

证券研究报告

通信组

分析师：张真桢（执业 S1130524060002）

zhangzhenzhen@gjq.com.cn

阿里、百度 capex 高增，AI 算力硬件β受益持续

通信周观点：

1) 阿里 26Q2（自然季度）资本开支 677 亿元，同比+75%，云收入增 45%创近年新高；26Q2 百度 GPU 云收入猛增 283%，资本开支 114 亿元，同比约 2 倍，占总营收比重达到 45.2%。国内云厂商 AI 算力收入加速兑现，验证 CSP 算力需求景气度持续高位，国产 AI 服务器、光模块及 IDC 等硬件产业链有望持续受益。2) DeepSeek API 正式调价，国产大模型从低价竞争转向价值定价。8 月 17 日起，DeepSeek V4 Pro 高峰时段输出价格涨 350%、缓存命中输入涨 1100%，空闲时段为高峰一半。本次调价为年内国产大模型密集涨价潮的延续，腾讯云、阿里云、百度智能云及智谱此前均已上调 API 价格。国产大模型在国际竞争中展现出强议价能力，模型能力与商业化正同步进入收获期。3) 中兴通讯上半年利润承压但政企算力业务高增，中际旭创 AI 光模块业绩爆发式增长，产业链分化中算力硬件景气度持续兑现。1H26 中兴营收 780.25 亿元（+9.05%），归母净利润 27.53 亿元（-45.57%），但政企业务收入 276.31 亿元（+43.51%），服务器及存储在多个行业突破；中际旭创同期营收 417.78 亿元（+182.49%），归母净利润 136.51 亿元（+241.70%），海外收入占比超 94%，1.6T 硅光模块规模放量。AI 算力投资加速向硬件端传导，光模块等互联环节业绩弹性最为突出，国产服务器及光通信产业链高景气有望延续。4) SK 海力士承诺三年将 50%以上自由现金流用于股东回报，三星或将推出高达 110 万亿韩元回购计划，存储巨头集体转向强化股东回报。SK 海力士宣布 2025-2027 年将累计自由现金流的 50%以上用于股东回报，并实施 40 万亿韩元股份回购及注销；三星电子宣布规模达 90-110 万亿韩元的股东回报计划，有望创韩国企业历史之最。两大存储巨头将未来现金流优先用于分红和回购而非激进扩产，行业资本配置趋于理性，有助于稳定存储芯片供需格局和价格预期，对股价形成正面支撑。5) SK 海力士在《自然·电子学》发表 CPO 技术蓝图，提出“以光为中心”架构将光互联延伸至内存接口。公司联合多所高校研究机构指出，AI 集群扩展至数千 GPU 时铜缆受功耗和信号衰减制约形成“带宽墙”，CPO 被视为破局关键。内存厂商角色正从零部件供应商转向 AI 基础设施架构共建者，AI 竞争已从单一芯片性能比拼升级为系统级互联与整体架构的全面竞争。6) 三星上调先进制程晶圆代工价格 10%-15%，AI 芯片需求激增致产能供不应求。三星 7 月起将 4nm/5nm 制程新订单价格上调 10%-15%，8nm 制程涨近 10%。供需结构性失衡推动代工价格进入上行通道，AI 算力投资仍处激进扩张期。7) Nebius 发行 45 亿美元可转换债券用于 AI 数据中心建设，新兴云厂商加速算力基础设施布局。Nebius Group 将发行两批共 45 亿美元可转债，资金用于建设数据中心及发展 AI 云业务。除头部 CSP 厂商持续扩张外，新兴 AI 云平台亦在大规模投入算力基建，全球 AI 硬件需求正从头部集中向全行业扩散，景气度维持高位。8) OpenAI CFO 在全员大会上表示公司“将在 2027 年成为上市公司”。二季度 OpenAI 营收 67 亿美元，环比仅增 18%且亏损扩大，今年 IPO 可能性愈发渺茫。9) OpenAI 旗舰模型 GPT-5.6 SoI 在 OpenRouter 和 Vercel AI Gateway 两大平台推出为期一年 50%独家折扣，覆盖全部计费类型及服务层级。此次促销与 7 月官方降价策略形成差异化，折扣有望刺激更多开发者调用旗舰模型，推动全球 Token 用量进一步提升，进而带动推理算力及 AI 硬件需求持续上行。10) 长江存储 IPO 辅导状态变更为“辅导验收”，国产存储龙头上市进程再进一步。公司成立于 2016 年，为集芯片设计、制造、封测于一体的存储器 IDM 企业，今年一季度营收突破 200 亿元（+100%），二季度全球 NAND 闪存市场份额首次跻身前三。

投资建议与估值

建议关注海外 AI Capex 持续超预期带动的光模块、光纤光缆与服务器板块；国内 CSP 云业务高增，Capex 超预期，同时继续关注国产链。

风险提示

AI 资本开支节奏不及预期；地缘政治与出口管制不确定性；存储价格大幅波动。



内容目录

一、细分行业观点.....	3
二、核心数据更新.....	5
三、本周行情.....	7
四、本周重要新闻.....	8
风险提示.....	13

图表目录

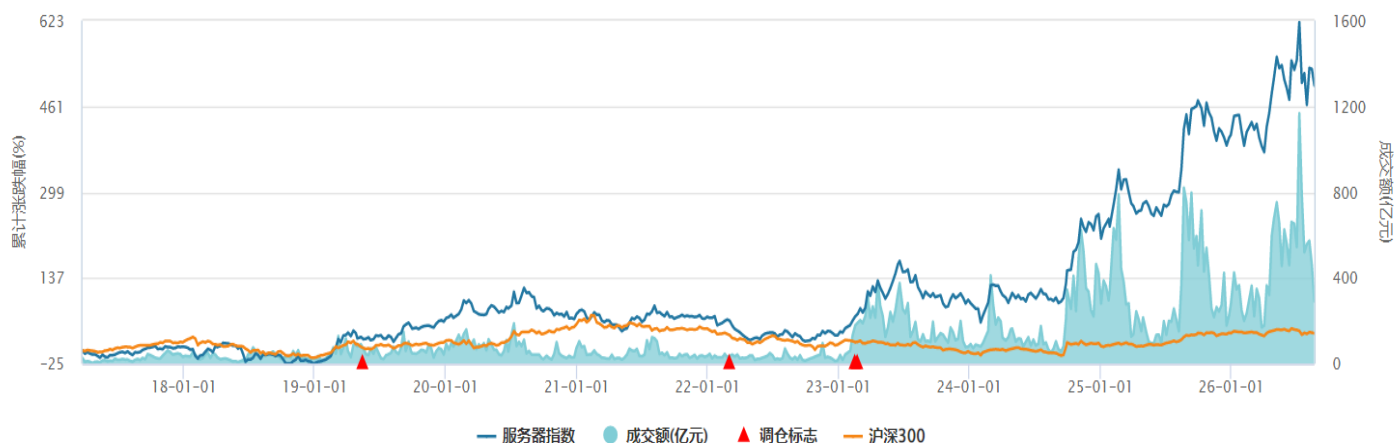
图表 1: 服务器指数 (8841058.WI) 走势	3
图表 2: 光模块 (CP0) 指数 (8841258.WI) 走势	3
图表 3: IDC 指数 (866052.WI) 走势.....	4
图表 4: 光纤指数 (884944.WI) 走势	4
图表 5: 本周通信板块景气度	5
图表 6: 千兆用户占比超三成	5
图表 7: 截至 6 月末 5G 用户占比约七成	5
图表 8: 上半年移动互联网累计流量同比增长 17.7%.....	6
图表 9: 6 月当月 DOU 达 23.57GB/户·月	6
图表 10: 电话通话量持续下滑	6
图表 11: 移动短信业务量大幅下降	6
图表 12: 千兆光纤宽带网络建设稳步推进	6
图表 13: 5G 网络建设持续深化.....	6
图表 14: 5 月光模块出口金额累计同比增加 18.6%	7
图表 15: 5 月光模块出口金额当月同比增加 50.7%	7
图表 16: 截至 6 月末蜂窝物联网终端用户数同比增长 5.76%.....	7
图表 17: 26Q1 全球蜂窝物联网模组出货量同比增长 16%	7
图表 18: 板块周涨跌幅排序 (%)	8
图表 19: 通信板块个股周涨跌幅 (剔除 ST 公司)	8



一、细分行业观点

服务器：本周服务器指数-4.94%，本月以来，服务器指数+6.44%。阿里云收入增速 45%创近年新高，百度 GPU 云收入猛增 283%，AI 服务器产业链景气度持续验证。

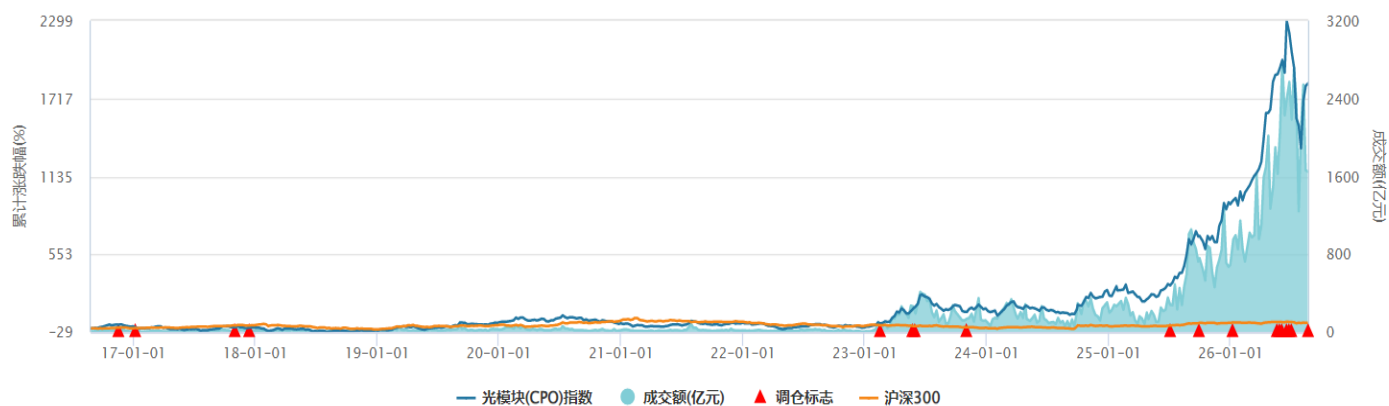
图表1：服务器指数 (8841058.WI) 走势



来源：wind，国金证券研究所

光模块：本周光模块（CPO）指数+0.96%，本月以来，光模块（CPO）指数+33.49%。中际旭创同期营收 417.78 亿元（+182.49%），归母净利润 136.51 亿元（+241.70%），海外收入占比超 94%，1.6T 硅光模块规模放量。AI 算力投资加速向硬件端传导，光模块等互联环节业绩弹性最为突出，光通信产业链高景气延续。

图表2：光模块(CPO)指数 (8841258.WI) 走势

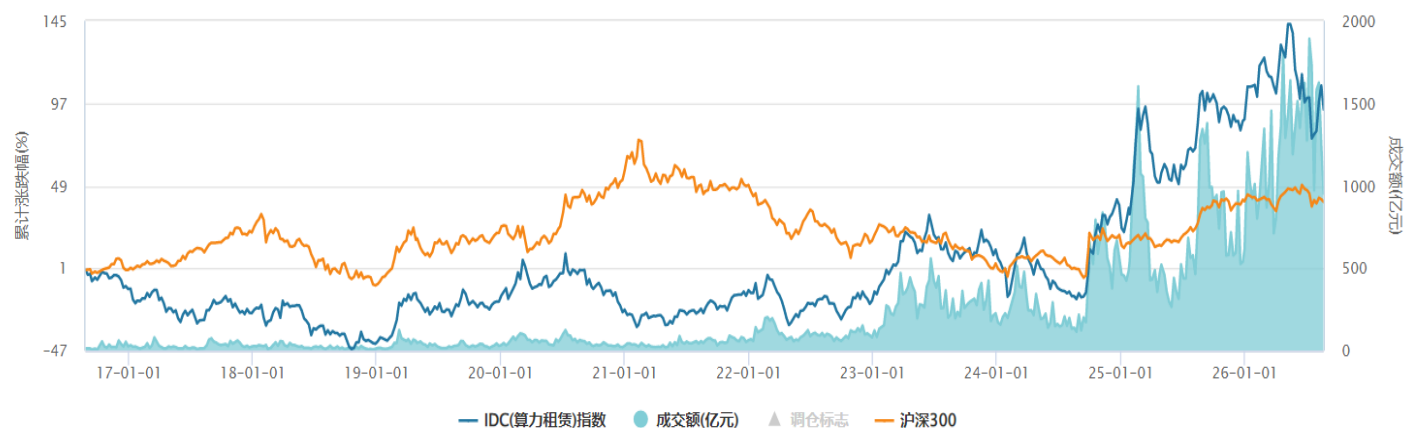


来源：wind，国金证券研究所

IDC：本周 IDC 指数-6.68%，本月以来，IDC 指数+6.94%。百度 26Q2 资本开支 114 亿元，同比约+200%，占总营收比重达到 45.2%，主要用于 GPU 采购和智算中心建设；阿里 26Q2（自然季度）资本开支 677 亿元，同比+75%。国内 IDC 需求景气度持续受国内 csp 大厂 Capex 增长拉动。



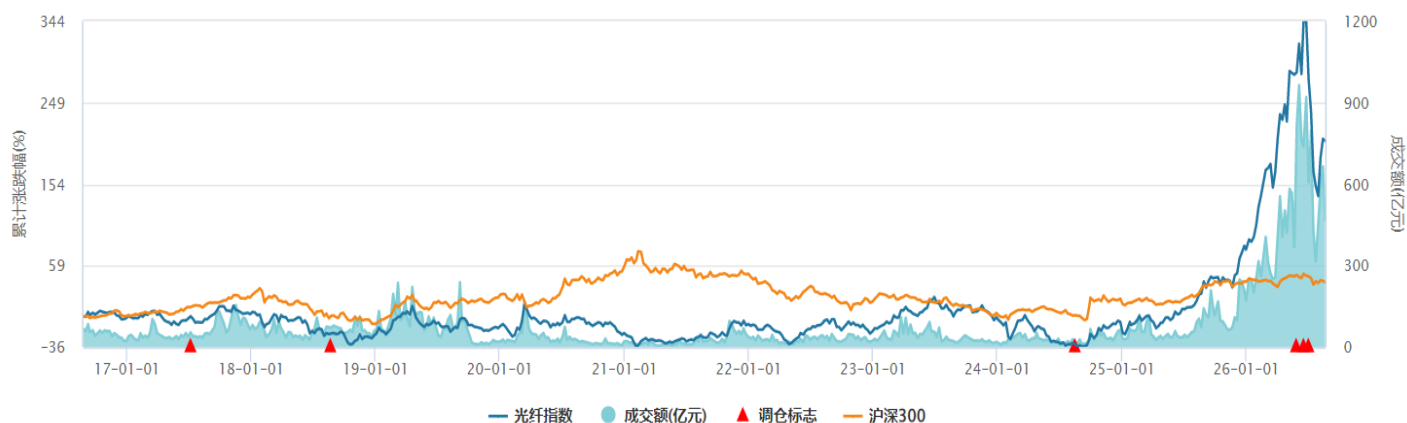
图表3: IDC指数 (866052.WI) 走势



来源: wind, 国金证券研究所

光纤: 本周光纤指数-0.75%, 本月以来, 光纤指数+26.69%。长飞光纤 2Q 业绩预告大超预期, 预计上半年归母净利润同比增长 711%-914%, 主要受益于北美长协涨价及低价运营商订单占比下降。公司作为全球光纤龙头, 北美涨价周期有望持续, 同时多芯光纤、空芯光纤等新技术变革将进一步打开成长空间。

图表4: 光纤指数 (884944.WI) 走势



来源: wind, 国金证券研究所



图表5：本周通信板块景气度

板块	景气度指标	本期景气度说明
运营商	稳健向上	5G 投资周期结束，电信行业端承压，但云与 IDC 业务放量接力成长，整体景气度稳健向上。
光模块	稳健向上	Lumentum 单季营收同比增长 109.3%，调整后净利润同比增长 415%，下一季度收入指引亦明显超预期。AI 数据中心需求持续拉动光芯片、光器件及系统产品放量，光通信产业链稳健向上。
服务器	稳健向上	工业富联 1H26 归母净利润同比增长 95.99%，云服务商 AI 服务器收入同比增长 2.3 倍，GPU/ASIC AI 机柜出货分别增长 3.2 倍/3 倍。中兴通讯 26H1 服务器及存储营收同比增长近 60%。GPU 与 ASIC 需求同步放量，AI 服务器产业链景气度持续验证。
交换机	稳健向上	26H2 及 2027 年超节点成为国产芯片主力出货方向，交换网络的层级和带宽要求持续提升，交换机板块需求旺盛。
连接器	稳健向上	康宁计划将在美国的光连接制造能力扩大 10 倍、光纤生产能力扩大 50%以上，看好光无源器件释放业绩。
IDC	加速向上	百度 26Q2 资本开支 114 亿元，同比近+200%，占总营收比重达到 45.2%，主要用于 GPU 采购和智算中心建设；阿里 26Q2（自然季度）资本开支 677 亿元，同比+75%。国内 IDC 需求景气度持续受国内 csp 大厂 Capex 增长拉动。
物联网	加速向上	苹果加码推进端侧 AI，依托自研芯片在 iPhone、Mac 等终端设备本地运行 AI 模型，同时正考虑收购端侧 AI 初创公司 LiquidAI。
液冷	高景气维持	2026 年 5 月四部门《促进人工智能与能源双向赋能行动方案》明确：新建万架级智算中心 100%液冷、28 年前完成全部风冷存量机房改造；20kW 以上机柜强制液冷，液冷板块景气度向上。

来源：经济参考报，新浪财经，华尔街见闻，东方财富网，今日头条，光纤在线，新浪科技，科技圈观察，搜狐网，国金证券研究所

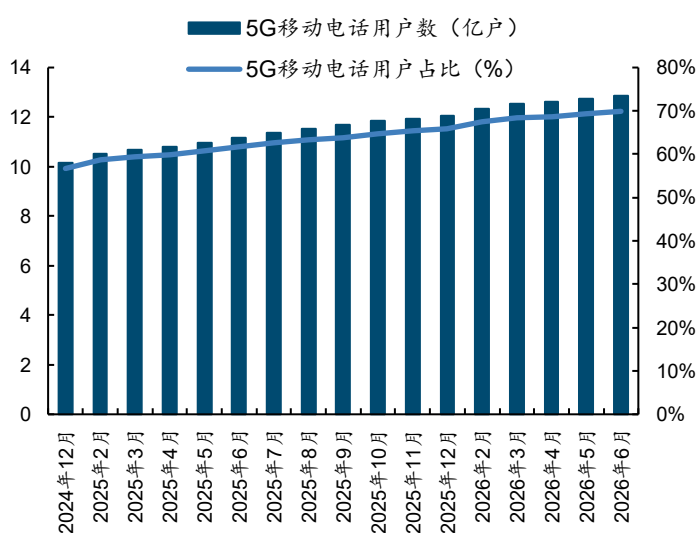
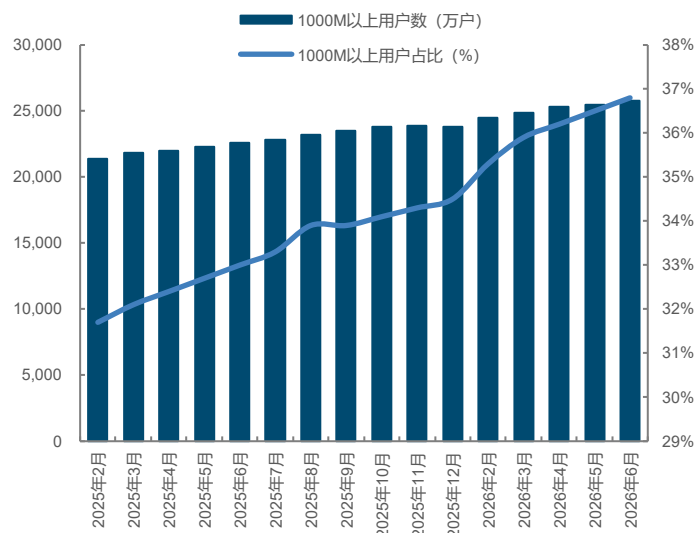
二、核心数据更新

运营商：运营商数据维持稳健增长

截至 6 月末，三家基础电信企业及中国广电的移动电话用户总数达 18.44 亿户，比上年末净增 1704 万户。其中，5G 移动电话用户达 12.88 亿户，比上年末净增 8434 万户，占移动电话用户的 69.9%。

图表6：千兆用户占比超三成

图表7：截至 6 月末 5G 用户占比约七成



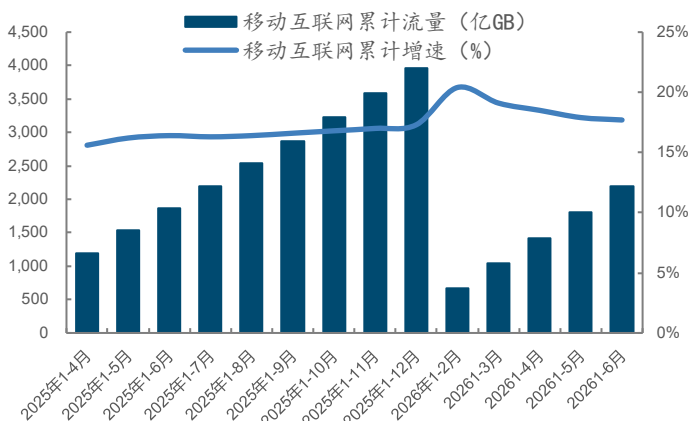
来源：工信部，国金证券研究所

来源：工信部，国金证券研究所

上半年，移动互联网累计流量达 2197 亿 GB，同比增长 17.7%。截至 6 月末，移动互联网用户数达 16.36 亿户，比上年末净增 2591 万户。6 月当月户均移动互联网接入流量（DOU）达到 23.57GB/户·月，同比增长 13.6%，比上年底高 0.53GB/户·月。

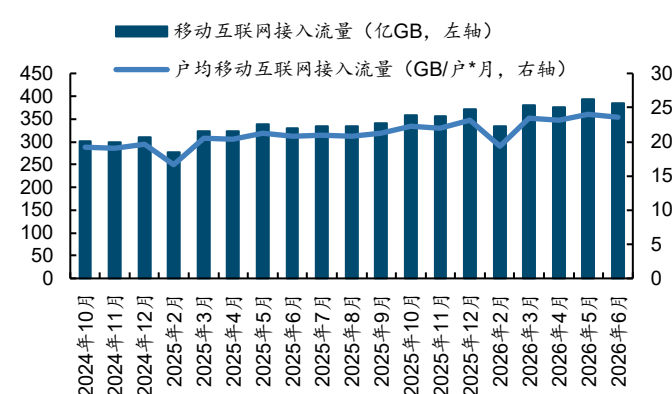


图表8：上半年移动互联网累计流量同比增长17.7%



来源：工信部，国金证券研究所

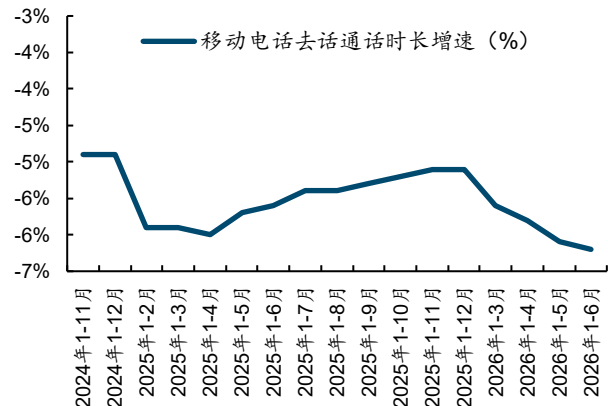
图表9：6月当月DOU达23.57GB/户·月



来源：工信部，国金证券研究所

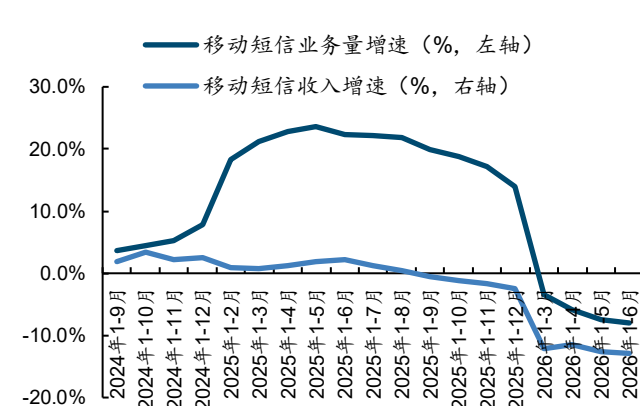
上半年，移动电话去话通话时长完成9469亿分钟，同比下降6.2%；固定电话主叫通话时长完成279.3亿分钟，同比下降20.4%。上半年，全国移动短信业务量同比下降8%；移动短信业务收入同比下降12.9%。

图表10：电话通话量持续下滑



来源：工信部，国金证券研究所

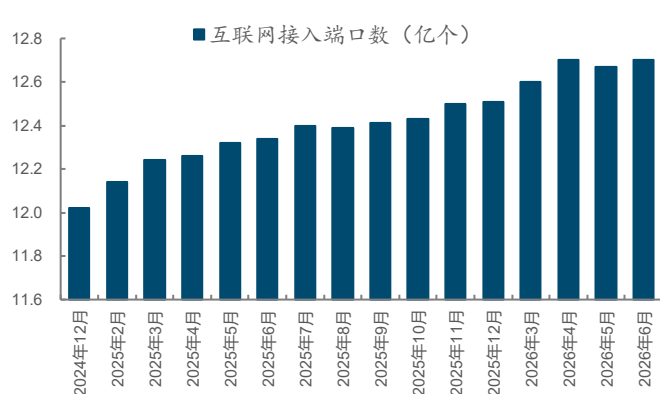
图表11：移动短信业务量大幅下降



来源：工信部，国金证券研究所

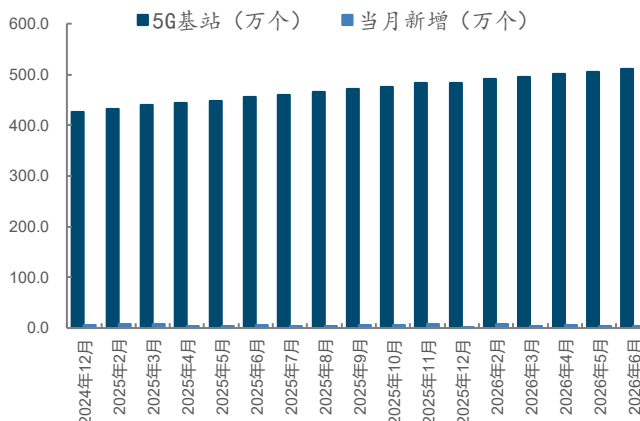
截至6月末，全国互联网宽带接入端口数量达12.7亿个，比上年末净增1790万个。其中，光纤接入（FTTH/O）端口达到12.3亿个，比上年末净增1552万个，占互联网宽带接入端口的96.6%。截至6月末，具备千兆网络服务能力的10G PON端口数达3286万个，比上年末净增124.1万个。

图表12：千兆光纤宽带网络建设稳步推进



来源：工信部，国金证券研究所

图表13：5G网络建设持续深化



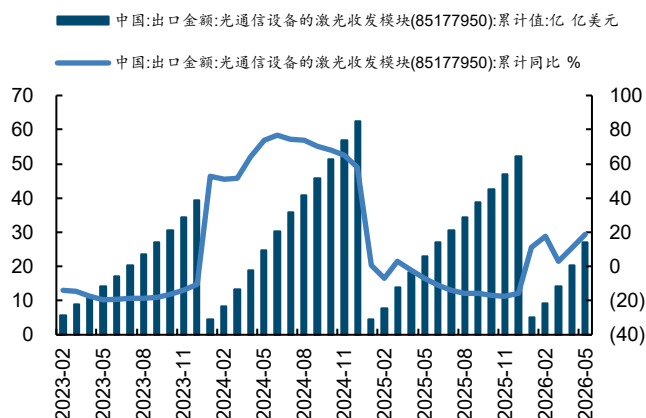
来源：工信部，国金证券研究所



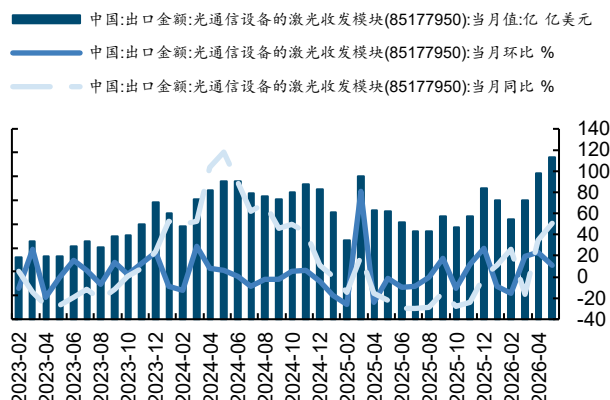
光模块数据：2026 年 5 月我国光模块出口数据同比+50.67%；1-5 月累计同比+18.6%。

图表14：5 月光模块出口金额累计同比增加 18.6%

图表15：5 月光模块出口金额当月同比增加 50.7%



来源：wind，国金证券研究所

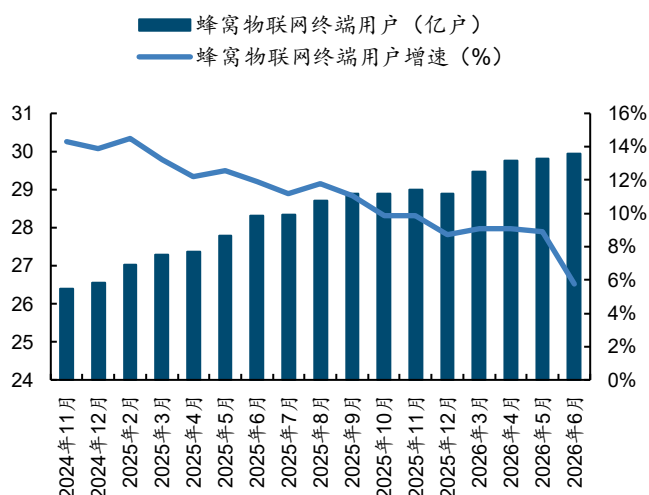


来源：wind，国金证券研究所

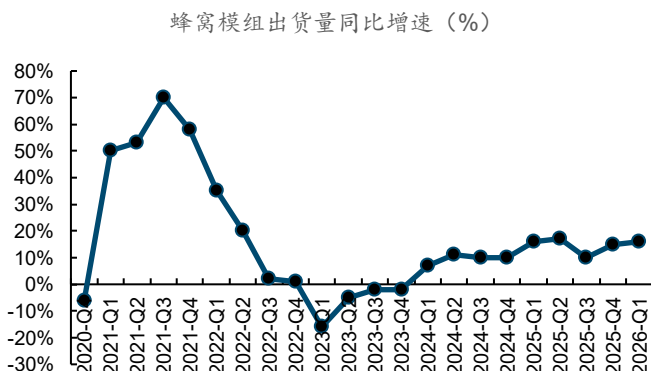
物联网数据：截至 6 月末，三家基础电信企业发展移动物联网终端用户 29.94 亿户，比上年末净增 1.06 亿户。互联网电视（IPTV、OTT）用户数达 4.09 亿户，比上年末净增 99 万户。

图表16：截至 6 月末蜂窝物联网终端用户数同比增长 5.76%

图表17：26Q1 全球蜂窝物联网模组出货量同比增长 16%



来源：工信部，国金证券研究所



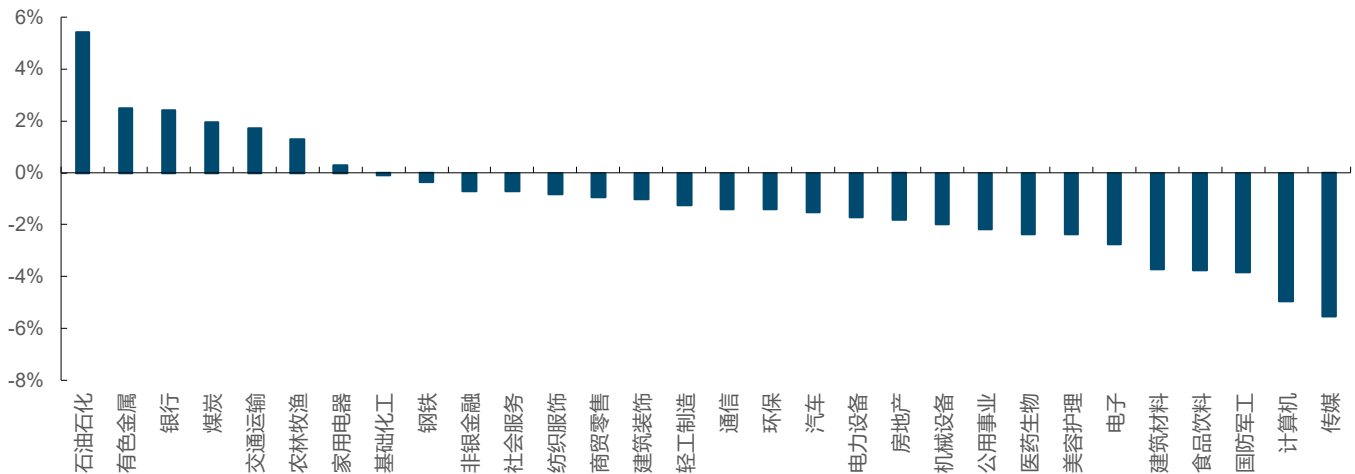
来源：国金证券数字未来实验室，RFID 世界网，C114 通信网，国金证券研究所

三、本周行情

回顾本周行情（8 月 17 日-8 月 21 日），参考申万一级行业划分，通信板块涨跌幅为-1.39%，排名全行业第 16。



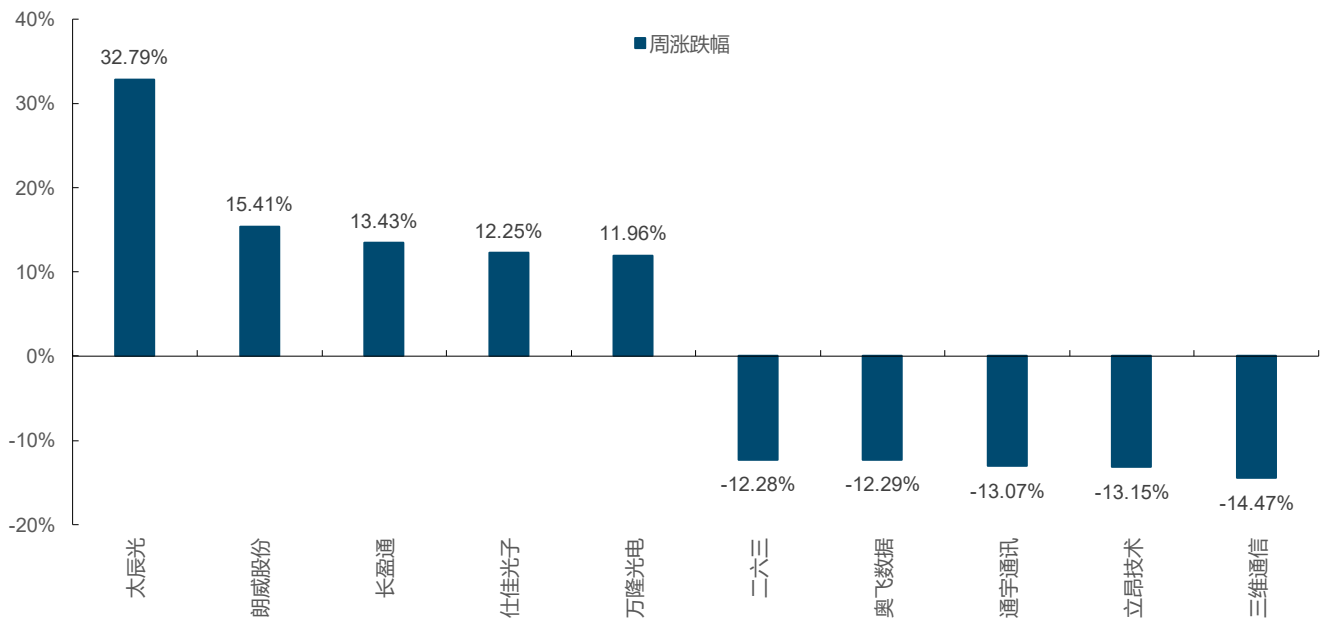
图表18：板块周涨跌幅排序（%）



来源：同花顺 ifind，国金证券研究所

从个股情况来看，本周太辰光、朗威股份、长盈通、仕佳光子、万隆光电为通信（申万）涨幅前五大公司，涨跌幅分别为+32.79%、+15.41%、+13.43%、+12.25%、+11.96%。通信（申万）跌幅前五为三维通信、立昂技术、通宇通讯、奥飞数据、二六三，涨跌幅分别为-14.47%、-13.15%、-13.07%、-12.29%、-12.28%。

图表19：通信板块个股周涨跌幅（剔除ST公司）



来源：同花顺 ifind，国金证券研究所

四、本周重要新闻

4.1 行业新闻

英伟达缩减对 Open AI 数据中心项目提供 2500 亿美元担保的计划

财联社 8 月 15 日电，英伟达与 Open AI 正接近达成融资协议以在俄亥俄州建设大型数据中心园区，但双方已重新调整交易结构，英伟达的财务担保金额从 2500 亿美元降至不到 1200 亿美元，调整旨在回应投资者对这资产负债表风险的担忧。根据目前讨论的方案框架，英伟达将为第一阶段（总计约 5 吉瓦电力容量）提供财务担保，后续阶段是否及如何融资将再做决定。

英伟达与 SB Energy 签订重大协议

环球市场播报 8 月 16 日消息，英伟达正与软银支持的 SB Energy 就最高 30 亿美元投资展开谈判，这笔投资与英伟达



向 Open AI 提供约 1000 亿美元信贷支持的讨论直接相关，旨在推动俄亥俄州大型数据中心项目落地，其中一半约 15 亿美元在项目交易签署时投入，另一半在 SB Energy 计划中的 IPO 中认购，若完成则英伟达将成为该 IPO 最大投资者之一，SB Energy 最早可能于下月上市，目标募资至少 50 亿美元。该项目预计配套 100 亿瓦新发电能力，总开发成本约 5000 亿美元，Open AI 将使用该数据中心并采用英伟达芯片。

美国敦促 Apple 不要购买中国存储芯片

TechWeb 8 月 17 日报道，近日美国商务部长霍华德·卢特尼克在参观休斯敦苹果先进制造中心后公开表态，政府完全不赞成该合作计划，强调美国头部科技企业使用中国本土内存芯片“并非值得鼓励的好事”，并确信企业能找到其他替代方案解决供应链紧张问题。尽管全球存储芯片供应短缺持续发酵，苹果出于缓解供货压力及压减物料成本考量，早已启动对长鑫存储产品的适配测试，并于今年 6 月 27 日正式向特朗普政府提交采购请愿申请，但该计划因触及中美半导体技术合作高敏感领域，随即遭遇巨大政治阻力。在美国国内半导体保守派势力及传统美系存储厂商联手施压下，该计划几无获批可能，苹果被迫放弃这一备选供应链方案。

特斯拉中国车机上线豆包大模型

钱江晚报 8 月 19 日消息，特斯拉宣布在中国市场全系车型车机系统上线豆包大模型，实现语音交互能力的全面升级，车主可通过语音助手获得更智能的交互体验。此次升级覆盖特斯拉在华所有在售车型，标志着豆包大模型在智能汽车领域的重要落地。

Marvell 与谷歌达成定制芯片合作并发行 121.8 亿美元认股权证

智通财经 8 月 22 日报道，Marvell 与谷歌达成定制芯片合作，同时向谷歌发行一份允许购买价值约 121.8 亿美元 Marvell 股份的认股权证，受此消息提振，迈威尔科技周三美股盘中大涨。根据双方声明，两家公司将合作开发 AI 推理加速器、存储控制器、网络接口控制器、内存接口控制器及内存计算技术；Marvell 于 7 月 29 日与谷歌签署商业协议提供定制芯片产品，认股权证最多可购买 5897 万股，行权价每股 206.58 美元，其中协议签署后第一年内每季度归属 136.0867 万份，剩余股份则根据 2027 财年第三季度至 2033 财年结束期间的“自主购买”情况归属，每产生 5 亿美元收入即有一批股份归属。

Meta 成为微软 AI 最大客户之一

财联社 8 月 20 日电，Meta 每年花费数亿美元通过微软 Azure 云服务访问 AI 模型，每周通过 Azure 使用的计算量已达数万亿 Token，Meta 已成为微软最大 AI 客户之一。微软 AI 战略的重要一环是通过 Foundry 模型市场提供来自不同供应商的 AI 模型，截至今年 7 月已拥有 10 万名客户。Meta 正在大举投资 AI 以协助软件开发工作，开发人员曾通过 Foundry 使用 Open AI 的技术帮助评估自研模型输出效果，公司也会在研发过程中租用其他领先模型；同时 Meta 也在打造自己的 AI 模型服务 API 业务，未来可能与 Foundry 形成竞争，这有助于降低对微软等外部服务的支出。过去 Meta 也曾依赖微软技术后再以自研方案取代，2000 年代末微软 Bing 曾为 Facebook 提供搜索支持，但到 2014 年底 Meta 已停止使用，如今 Meta 已成为全球最大的 AI 数据中心建设者之一。

4.2 公司新闻

长江存储：

8 月 19 日，证监会官网显示，长江存储控股股份有限公司 IPO 辅导状态变更为“辅导验收”。长江存储今年一季度营收突破 200 亿元，同比增长 100%。据市场研究机构发布的 2026 年第二季度数据，按出货量计算，长江存储在全球 NAND 闪存市场的市场份额首次跻身前三。

快手科技：

8 月 19 日，快手科技披露 2026 年第二季度业绩，财报显示当期总收入达 355 亿元，经调整净利润达 39 亿元，盈利能力保持稳健。用户基本盘方面，第二季度月度活跃用户数达 1.97 亿，同比增长超 20%，AI 技术从产品工具向商业全链路渗透，被视为驱动业绩增长的重要变量。公司表示未来将持续加大 AI 投入，推动技术在多领域落地，同时优化用户体验以巩固用户粘性，实现可持续发展。作为快手 AI 业务的核心载体，可灵 AI 交出了亮眼的商业化成绩单。第二季度，可灵 AI 营收超 8.5 亿元，同比增长超 200%。截至 6 月末，可灵 AI 全球用户突破 1 亿，较 2025 年末增长约 67%，业务覆盖 224 个国家和地区；企业客户数量接近 5 万家，同样较 2025 年末增长约 67%。

稀宇科技：

财联社 8 月 20 日电，稀宇科技推出 MiniMax Design。MiniMax Design 是一款把多模态模型能力变成真实可用的生产力的 Harness。用户提出创作需求后，它能够理解目标、拆解任务、调用相应的模型与 Skills，完成从素材处理、生成和编辑到最终交付的完整过程。MiniMax 原生多模态视频模型 H3 为 MiniMax Design 提供了新的能力基础，它已经验证了 MiniMax 在多模态理解、视频生成、视频编辑、参考控制等方面的模型能力，尤其适合 Video Editing 等对理解、控制与修改要求更高的场景。

智谱：

上海证券报 8 月 19 日讯，智谱正式上线新一代基座模型 GLM-5.3 的 API，该模型擅长复杂编码、防御性网络安全与长程任务，在 Artificial Analysis Intelligence Index 中取得 60 分，进入全球前沿模型能力区间，与 Claude Fable 5、GPT-5.6 Sol 等闭源旗舰处于同一水平，并与 Kimi K3 并列开源模型第一。GLM-5.3 以更小参数规模、更高参数效



率和更低调用成本达到前沿水平，实现前沿旗舰模型中最低的单任务成本，将“智能—成本”帕累托前沿显著推进，利好 AI 应用加速落地与规模化普及。更具突破性的是，GLM-5.3 与上一代 GLM-5.2 沿用同一基础模型，性能提升几乎全部来自后训练，表明前沿能力推进不必依赖更大参数规模和更高算力投入。目前该模型已接入 ZCode、GLM Coding Plan 等官方产品，API 定价与 GLM-5.2 保持一致，模型权重将于下周五开源。业内认为，资本市场将关注其能否依托 API、编程工具和企业级市场，加速模型能力向开发者规模化使用和商业收入转化，推动国产大模型竞争从能力突破进入应用兑现阶段。

百度：

8月18日，百度披露2026年第二季度财报，当期总收入达313亿元，其中AI业务收入占比达50%，已连续两个季度过半。AI云基础设施中GPU云同比增长283%，较上一季度184%的增速进一步加快，连续四个季度实现三位数增长；百度AI云基础设施整体收入为73亿元，同比增长50%。据IDC报告，百度智能云在金融、游戏、具身智能等行业市场份额位居第一。AI应用收入达25亿元，同比增长3%。8月百度完成AI办公产品线整合，形成以百度搭子、库库AI、秒哒为代表的AI办公产品矩阵——百度搭子7月MAU环比增速达1063.79%，库库AI办公MAU超2500万位居行业第一，秒哒3.5版本以33.4%市场份额居行业第一，自我演化智能体百度伐谋已有3000家企业申请试用。此外，AI原生营销服务收入为26亿元，百度APP月活跃用户数达6.4亿。百度创始人李彦宏表示，AI业务的持续增长印证了百度从以互联网为核心向以AI为先的公司转型，进一步增强了公司对长期增长潜力的信心。

Deepseek：

1) 8月17日，DeepSeek API 正式执行新价格体系，采用峰谷定价机制：高峰时段价格为空闲时段的两倍。旗舰模型 DeepSeek-V4-Pro 高峰时段最高涨幅达1100%。据官方介绍，峰谷定价的核心初衷在于通过市场化价格杠杆引导开发者错峰调度任务，分流高峰算力压力以提升平台运行稳定性——今年以来 DeepSeek 已多次发生服务异常，其中5月8日、21日、24日、28日连续崩溃，直接原因系用户量暴涨66.7%而算力储备仅增长8.3%，供需严重失衡。在此之前，腾讯云、阿里云、百度智能云已相继上调API算力产品价格，智谱年内亦三次提价。价格普涨的背后，是全球Token调用量的爆发式增长：据OpenRouter周度榜单，8月3日至8月9日全球AI大模型总调用量达69万亿Token，环比增长21.48%，中国模型调用量达34.25万亿，连续15周超越美国位居全球第一；当周调用量排名前四的模型均出自中国，其中 DeepSeek-V4-Flash 正式版以8.83万亿Token居首，环比暴涨570%。

2) 8月21日华尔街见闻消息，DeepSeek 上线实验性视觉理解模型 DeepSeek-V4-Flash-Vision-Exp，在保留原有文本能力的基础上增加图像理解功能，并向开发者开放API。据公司公布的测试结果，新模型在Agent、推理、世界知识等纯文本任务上的表现与 DeepSeek-V4-Flash 基本相当；在视觉理解类Agent基准测试中，性能较前代明显提升，已逼近Opus-4.8。模型定位为实验性版本，重点在于增加视觉能力的同时保持文本性能，DeepSeek 在测试中使用Harness极简模式框架，采用max档位、temperature=1.0、top-p=0.95等参数。公司同时展示了新模型在多类Agent任务中的应用，体现其对图像和文本信息的综合理解能力及多轮交互任务中的应用潜力。API计费方面，V4-Flash-Vision-Exp 沿用V4-Flash定价，图片输入转换为token计费，单张图片最多占用384个token，支持Chat Completions、Messages 和 Responses 三种调用格式，图片可通过base64、外部URL或Files API传入。同期上线的Files API目前免费，开发者可提前上传图片并在后续请求中通过file_id调用以减少重复上传。

宇树科技：

中国经济网8月19日讯，宇树科技（688836.SH）正式登陆上交所科创板，发行价格为150.80元/股，发行股票数量为4044.6434万股，募集资金总额约60.99亿元，扣除发行费用后净额约59.17亿元，较原计划420171.12万元超募约17.15亿元。上市首日，该股开盘报1100元，截至收盘报845元，涨幅460.34%。招股书显示，募集资金拟分别用于智能机器人模型研发项目、机器人本体研发项目、新型智能机器人产品开发项目及智能机器人制造基地建设项目。

新易盛：

8月20日晚，光通信龙头新易盛发布2026年限制性股票激励计划（草案），拟向激励对象授予限制性股票总量648.12万股，占公司股本总额的0.4648%，授予价格为212元/股（较当日收盘价414.02元/股折价约五成），首次授予518.52万股，激励对象为公司董事、高管、中层管理人员及核心骨干，合计不超过961人，其中中层管理人员及核心骨干共920人，获授比例达78.27%。预留129.60万股，预留部分对应考核年度为2027年至2029年，各年度目标分别为2027年营收不低于900亿元、2027年至2028年累计不低于2400亿元、2027年至2029年累计不低于4500亿元。以新易盛2025年营收248.42亿元为基准测算，2026年目标需在2025年基础上翻倍，2027年同比增长不低于80%，2028年同比增长不低于66.7%。此番“高门槛”考核充分彰显公司对CPO市场空间的坚定信心。经营层面，公司7月中旬预告上半年净利润达70亿至80亿元，同比增长77.56%至102.93%，二季度预计净利润环比增长51.80%至87.77%。公司披露当前800G光模块为主力出货产品，1.6T光模块二季度出货量较一季度显著增长且后续出货规模预计持续提升，硅光方案产品出货占比稳步提高，下半年及2027年订单能见度较为明确。针对市场流传的“1.6T订单下修”传闻，公司回应称产品需求端未发生变化，整体市场需求预期保持强劲，正在持续扩大产能。

中际旭创：

8月21日，中际旭创披露2026年半年度报告，当期实现营业收入417.78亿元，同比增长182.49%；归属于上市公司股东的净利润136.51亿元，同比增长241.70%；扣非净利润130.92亿元，同比增长229.32%。第二季度单季营收222.82亿元，环比增长14.3%；归母净利润79.17亿元，环比增长38.0%，利润增速显著跑赢营收。海外市场为核心增长引擎，上半年海外收入达396.15亿元，同比增长209.9%，占总营收逾94%；国内市场同样保持增长，400G、800G



产品交付规模稳步提升。产品结构方面，1.6T 硅光模块进入规模放量阶段，出货量逐季提升，800G 仍为主力出货产品；公司同时在 OFC 2026 期间展示了 12.8T 8×DR8 XPO 模块、6.4T 4×DR8 NPO 模块等下一代光互连技术，围绕 XPO、NPO、CPO、OCS 持续布局。公司宣布每 10 股派发现金红利 12 元。研发费用约 11.53 亿元，同比增长 96.9%，重点投向硅光与相干技术。

Cerebras:

8 月 19 日智通财经消息，AI 芯片公司 Cerebras Systems 发布新一代机架级 AI 系统 CS-4，宣称推理速度达 GPU 方案 30 倍，直接对标英伟达，预计本季度开始出货。CS-4 由三颗 WSE-3T 晶圆级处理器驱动，每颗集成 4 万亿晶体管和 90 万个 AI 核心，系统提供 750PFLOPs 算力。在 GPT-OSS-120B 测试中每用户每秒可生成超 4400 个 Token。公司于 5 月以 185 美元/股上市，首日收涨 68%，但股价已从高点回落约 35%；二季度核心营收 2.099 亿美元，同比增长 103%，云推理业务同比增长 287%。行业认为该产品标志着 AI 芯片竞争正向推理市场延伸，能否撼动英伟达地位取决于独立基准验证与规模化交付能力。

Nebius:

8 月 19 日环球市场播报消息，总部位于阿姆斯特丹的 Nebius Group NV 宣布将通过发行可转换债券筹集 45 亿美元，用于建设数据中心及发展人工智能云业务。此次发行分为两批——27.5 亿美元于 2030 年到期，17.5 亿美元于 2034 年到期。该公司此前在与 Meta Platforms 达成交易后，已于今年早些时候通过一次上调规模的可转换债券发行筹集了约 43 亿美元。受此消息影响，Nebius 股价周三美国盘前交易中一度下跌 9.2%，此前截至周二收盘该股今年已累计上涨 197%。

华邦电子:

据 8 月 20 日华尔街见闻报道，全球利基型存储 IDM 龙头企业华邦电子宣布提前启动中国台湾高雄路竹厂扩产规划，将原定分期推进的二、三期工程整并为 Module B 工厂同步开发，预计 2027 年动工兴建无尘室，最快 2029 年初开始装机。客户已在抢订三年后产能，存储产能洽谈已排至 2030 年。财务数据显示，华邦电子第二季度合并营收 598.43 亿元新台币，季增 56.4%、年增 184.7%，DRAM 价格单季上涨约 100%、Flash 上涨约 43%，毛利率升至 66.2%，每股税后纯益达 5.40 元；进入第三季度，机构预计利基型 DRAM 与 SLC NAND 价格季涨幅约 50%，NOR Flash 受 CSP 客户拉货拉动涨幅约 30%，第四季度涨幅虽预计收窄至 2%至 5%，但营收与获利仍可维持季增。Module B 产品规划覆盖 Standard DRAM、CUBE DRAM、Wafer-on-Wafer 及矽电容，将支持 14 纳米及未来 12 纳米制程并规划导入 EUV 设备，设备导入将依客户需求预测及长期协议分阶段推进以控制资本支出节奏。机构人士指出，代理式 AI 正推动存储容量与先进封装需求持续增加，成长趋势预期延续相当长时间，这正是华邦电子决定提前启动扩产的关键原因。

博通:

8 月 20 日每日经济新闻报道，博通正与黑石、阿波罗全球管理洽谈，通过特殊目的载体 (SPV) 为 AI 芯片基础设施筹集超过 600 亿美元债务融资，受益方包括 Anthropic 等 AI 企业。按照目前讨论的方案，融资可能包括 600 亿至 700 亿美元优先担保债务及约 300 亿美元次级债务，总规模最高可达 1000 亿美元。此次交易延续了博通、黑石和阿波罗今年 6 月建立的“AI XpV 合作”框架，此前三方已完成一笔 350 亿美元债务融资。新融资中，博通还可能为部分优先债务提供担保。对于 Anthropic 而言，此举有助于提前锁定模型训练和推理所需的芯片与算力资源；对博通则可借助外部资本扩大 AI 芯片及数据中心设备销售，与英伟达争夺 AI 基础设施市场。分析人士指出，这笔潜在千亿美元融资再度表明，AI 竞争正从“谁能造出更好的芯片”延伸至“谁能为芯片采购提供融资”。

4.3 海内外大厂重点跟踪

阿里巴巴:

1) 8 月 20 日阿里巴巴集团披露 2027 财年第一季度业绩，当期收入达 2689.5 亿元人民币（约 396.3 亿美元），同比增长 9%。财报显示，阿里云外部收入同比增长 45%，创下 22 个季度以来的最快增速；AI 相关产品收入达 123.76 亿元，连续第 12 个季度实现三位数同比增长；AI 云及计算服务经调整 EBITDA 同比猛增 133%。阿里巴巴集团首席执行官吴泳铭表示，在全栈 AI 战略驱动下，阿里巴巴已做好充分准备，以把握人工智能及 AI 计算需求的显著增长机遇。

2) 8 月 18 日韩联社消息，阿里云宣布在韩国正式启用第三座数据中心，以应对当地市场对人工智能及云计算日益增长的需求，该新设施系阿里巴巴集团 530 亿美元 AI 基础设施投资计划的一部分，将提供计算、存储、网络、安全、数据库、云容器及云原生服务等企业级云计算服务。阿里云表示，新设施投入运营后将帮助韩国客户提升服务可用性与运营稳定性、降低网络延迟并增强灾难恢复能力，保障企业核心业务稳定运行；同时顺应企业级 AI 智能体需求增长，将在韩国拓展 AgentRun 等服务，覆盖从 AI 智能体开发、测试到部署、运营及安全管理的全流程。至此，阿里云全球基础设施已扩展至 30 个地域和 104 个可用区。阿里云智能韩国区总经理尹容俊表示，韩国是阿里云长期持续投资的核心市场，第三座数据中心不仅是基础设施扩容，更将支持韩国企业更加稳定地运营 AI 及云服务。

字节跳动:

8 月 21 日科技日报报道，字节跳动旗下豆包完成“工作任务”模式系列升级，新增手机远程控制电脑、本地与云电脑双模式运行、界面协同编辑、模块化技能组合等功能，标志着豆包 AI 能力从基础问答交互向自主落地执行工作任务全面进阶。今年 6 月豆包上线“工作任务”模式，支持 AI 在授权后直接操作本地电脑和浏览器。此次更新将操控扩展至手机端，支持关联多台电脑并跨设备调取文件。同时，新版推出“豆包虚拟桌面”功能，人与 AI 可在同一台电



脑上同时操作互不干扰，AI 不抢占鼠标、键盘或打断当前工作，用户可实时查看操作过程并随时暂停接管任务。新增的本地与云电脑双模式允许按需选择：本地模式数据保留在自有机器上，适用于文件整理、文档操作等本地任务；云电脑模式具备持久在线能力，可运行大型任务而不影响本机性能，适合行业信息整理、文献综述、招聘信息分析等数据采集与调研场景，用户亦可使用手机远程控制云电脑完成 PPT 制作等任务。

寒武纪：

8 月 17 日晚，寒武纪发布公告称已于近日完成员工股票激励相关股份登记工作，激励股票将于 8 月 20 日上市流通。本次拟归属股票的激励对象共 124 名，合计分享 59.76 万股股票，以 8 月 17 日收盘价 1156 元/股估算，对应股票价值合计约 6.9 亿元，人均约 557 万元。其中，陈帅、刘少礼、刘毅等 6 名董事、高管及核心技术人员获授 7.15 万股，剩余 118 人分享 52.61 万股。寒武纪成立于 2016 年，是一家提供云边端一体、软硬件协同、训练推理融合智能芯片产品和平台化基础系统软件的公司，2025 年首次实现盈利并成为首家退出科创板科创成长层的企业，去年 8 月股价一度超越茅台成为 A 股“股王”。

Open AI：

1) 8 月 17 日起，Open AI 旗下旗舰模型 GPT-5.6Sol 在 OpenRouter 与 Vercel AI Gateway 两大第三方平台推出为期一年的 50% 价格优惠。Open AI 官方直营渠道及 Azure、Bedrock 等其他合作渠道未跟进。此次促销引发行业广泛讨论，SemiAnalysis 等机构质疑其动机或为影响市场占有率判断，截至目前 Open AI 及合作平台均未公开说明具体促销原因，这场价格调整正在引发关于市场份额计算方法与商业策略的深层讨论。

2) 8 月 20 日，据财联社报道，Open AI 首席财务官 Sarah Friar 在周三全员大会上对员工表示，公司“将在 2027 年成为一家上市公司”，若“业务发展态势持续向好”亦有可能更早登陆资本市场。此前一日，媒体报道 Open AI 二季度营收 67 亿美元、环比仅增 18%，不及部分股东预期，同时运营亏损扩大至 123 亿美元；竞争对手 Anthropic 二季度营收达 116 亿美元，首次在单季营收上超越 OpenAI 并实现小幅运营盈利。两家公司均于 6 月宣布已秘密向美国证券交易委员会提交招股书草案。与会人士透露，Friar 在会上明确表示即便 Anthropic 比 OpenAI 更早上市，公司也无需为此担心。

Anthropic：

1) 8 月 18 日，据彭博社报道，Anthropic 向投资者披露截至今年 7 月底，公司年化营收运行率（ARR）已达到 650 亿美元，较 2025 年末的逾 90 亿美元增长至 7 倍以上。今年 5 月该数字刚跨过 470 亿美元，仅两个月便再增约 38%。公司最近已完成季度（第二季度）初步收入超过 115 亿美元，去年同期仅为 7.87 亿美元，同比增长约 14.6 倍，且该季度已录得调整后营业利润为正，标志着规模扩张正逐步显现经营杠杆。收入暴涨的关键变量在于企业市场，尤其是编程智能体 Claude Code——它使模型调用从偶发对话转变为持续运行、大量消耗 Token 的软件工程工作流，显著拉动了收入增长。按彭博社此前报道，Open AI 最近 ARR 已超 400 亿美元，但统计口径可能不完全一致。上市进程方面，据《金融时报》此前报道，Anthropic 投资者预计公司或于今年 10 月登陆公开市场，IPO 估值有望达到 2 万亿美元甚至更高，若实现将超越 Space X 成为史上估值最高的 IPO。650 亿美元 ARR 对即将走向资本市场的 Anthropic 而言是一张重要筹码，表明其已从 OpenAI 身后的追赶者转变为具备强劲收入增长与初步盈利能力的头部玩家。不过需注意，今年 6 月公司曾因政府管制要求暂时关闭 Claude Fable 5 和 Mythos 5 两款先进模型访问，经约两周协商后恢复，即便如此收入仍在快速增长，使这条增长曲线更显突出。

2) 8 月 22 日新浪财经报道，Anthropic 宣布聘请谷歌定制芯片项目创始人 Amir Salek 加入公司算力团队，向 James Bradbury 汇报，为这家 AI 实验室进军自研半导体领域做准备。Salek 曾负责谷歌 TPU 业务直至 2022 年，并参与推出了前七代 TPU 芯片。Anthropic 目前从英伟达、谷歌和亚马逊等多个供应商采购芯片，但近期已释放出建立自有芯片业务的明确意愿，已开始为相关项目招聘并发布职位信息。与其他 AI 巨头类似，Anthropic 正加紧获取足够的数据中心基础设施以支撑发展目标，自研芯片有望帮助公司应对供应短缺，同时根据自身需求定制芯片设计。竞争对手 Open AI 亦采取类似行动，其与博通共同开发的 Jalapeno 芯片计划于今年晚些时候开始使用。

SK 海力士：

1) 8 月 19 日澎湃财经消息，SK 海力士宣布一项价值 40 万亿韩元（约合 286 亿美元）的股票回购并全部注销计划，创下韩国上市公司历史之最。根据向韩国交易所提交的公告，回购金额为 40.0043 亿韩元，按董事会决议前一日收盘价 166.2 万韩元计算，对应约 2407 万股，占已发行股份总数的 3.3%，预计从 8 月 20 日起实施，周期约三个月，完成后全部注销。公司明确表示此举基于当前股价未能充分反映其业务竞争力、现金创造能力及中长期增长潜力的判断。

2) 8 月 20 日，SK 海力士联合机构研究人员，在顶级期刊《自然·电子学》发表共封装光学（CPO）论文，系公司首次在该级别期刊公开 AI 互联系统级技术蓝图。论文指出，HBM 虽解决芯片内部内存瓶颈，但当 AI 集群扩展至数千颗 GPU 时，机架至 Pod 层级的数据传输形成“带宽墙”，传统铜缆受功耗与信号衰减制约难以持续扩容，CPO 被视作破局关键。短期 CPO 聚焦处理器与机架之间的光信号传输；长期提出“以光为中心”架构，借助光子中介层直接打通算力资源池与内存资源池，将光互联延伸至内存接口，实现多 AI 加速器共享大容量内存池，突破封装物理限制。研究给出技术目标：单节点带宽超过 100Tb/s，功耗低于 1pJ/bit，芯片间距延迟低于 10ns，并完整梳理从 2D、2.5D 中介层到 3D 异质集成的演进路径，同时列明产业化需攻克光器件集成、一致性协议、可靠性等课题。弗吉尼亚大学李圭相教授表示，即便计算芯片性能提升，若芯片间数据吞吐能力跟不上，系统性能依旧无法提高，摆脱物理受限的铜布线、改用光传输数据是最具可行性的扩展路径。SK 海力士 AI 基础设施团队洪承勋组长将 CPO 阐释为把光收发器与处理器封装在一起，让芯片之间不再依靠长距离铜线而改用光完成数据交互，并指出 AI 时代内存厂商角色正从零部件



供应商转向参与整套 AI 基础设施架构共建，后续将推进产学研开放合作推动技术落地。

3) 8 月 21 日，据财联社援引知情人士透露，韩国存储芯片巨头 SK 海力士正筹划在日本东北部宫城县建设一座大型存储芯片生产工厂，投资规模初步讨论为数十万亿韩元（约合数百亿美元），不过公司尚未官方宣布，具体投资规模和选址仍可能有变数。SK 海力士方面对此回应称“任何具备配套基础设施的地区都可能纳入评估范围，但关于在日本建厂以及具体选址，尚未做出任何决定”。

三星：

1) 8 月 19 日财联社消息，三星电子晶圆代工部门近期已向部分客户提价，针对新订单的价格上调最高达 15%。具体来看，5 纳米 SF5 制程晶圆价格上涨 10% 至 15%，较成熟的 8 纳米制程涨幅接近 10%。三星位于韩国平泽工厂的 SF4 生产线自去年年底以来一直满负荷运行，为高通等客户生产逻辑芯片，同时生产用于三星多层 HBM 芯片的基础裸片。由于 AI 芯片需求激增导致全球芯片制造产能持续紧张，台积电先进制程产能趋于饱和，三星议价能力随之增强。不过公司仍需为美国客户预留产能，同时保留部分产能用于自身芯片生产，因此无法满足全部订单需求。研究机构 Counterpoint 数据显示，2026 年第一季度三星占全球晶圆代工收入的份额仅为 7%，而台积电份额超过 70%。

2) 8 月 19 日，三星电子宣布一项规模最高达 110 万亿韩元（约合 790 亿美元）的股东回报计划，相当于公司自由现金流的五成左右，创下韩国企业史上最高股东回报纪录。据三星向监管机构提交的文件，公司将派发约 30 万亿韩元的第三季度现金股息，并回购约 15 万亿韩元股票用于员工薪酬计划。

3) IT 之家 8 月 19 日消息，据韩媒《The Bell》报道，三星电子目标于 2026 年 9 月率先完成 1d nm DRAM 内存研发，2026 年 12 月进入量产转移阶段，2027 年上半年启动生产线建设，有望在 2027 年底实现首批量产。1d nm 制程属于 DRAM 制程节点中 1a、1b、1c 之后的下一代技术，目前仍处于较早期阶段，整体产品完成度与目标之间仍有较大优化空间。三星电子正加速推进 DRAM 技术迭代，以巩固其在存储芯片市场的竞争优势。

风险提示

1、AI 商业价值不及预期的风险：目前 AI 市场应用仍处于初级阶段，盈利模式仍需探索，市场尚未成熟。若商业模式无法持续发展新客户，需求大幅减弱，或市场接受度偏低，可能对营业收入造成较大负面影响，损害相关公司的盈利能力及产品或服务的商业价值。

2、技术发展速度不及预期的风险：目前 AI 模型的使用仍受限于诸多因素，在特定领域无法达到预期的提高生产力效果。该领域目前仍面临较大的技术挑战，包括模型训练效果不稳定、算法不成熟等问题。若技术落地不及预期，可能影响 AI 的应用领域和运行效率，造成较大的投资损失。

3、供应链集中度过高的风险：AI 行业基础设施建设目前高度依赖某几家核心供应商，极易受到相关供应商供应短缺的影响。此外在训练方面，AI 技术依赖于大量优质数据的输入。不可靠、低质量的数据来源会一定程度上影响 AI 模型训练的性能，同时提高训练过程中的不可控成本。

4、行业监管加剧的风险：目前生成式 AI 工具仍存在法律、伦理、安全风险。AI 生成内容的产权问题仍存在较大争议。各国可能针对 AI 的使用及 AI 生成内容进行更严格的监管及抵制，影响投资预期，并阻碍 AI 技术在产业上进一步落地。公司面临法律诉讼和声誉受损等负面影响风险。

5、市场竞争加剧的风险：在如今巨头科技公司加大 AI 投入，大量创业公司涌入竞争的大环境下，技术的迅速迭代及新算法的涌现可能使得公司技术迅速落后竞争对手，影响相关公司的市场份额和投资回报的稳定性。



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。

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	北京	深圳
电话：021-80234211	电话：010-85950438	电话：0755-86695353
邮箱：researchsh@gjzq.com.cn	邮箱：researchbj@gjzq.com.cn	邮箱：researchsz@gjzq.com.cn
邮编：201204	邮编：100005	邮编：518000
地址：上海浦东新区芳甸路1088号 紫竹国际大厦5楼	地址：北京市东城区建国门内大街26号 新闻大厦8层南侧	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