综合

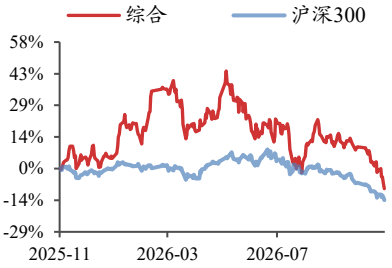
2026 年 11 月 01 日

投资评级：看好（维持）

海外科技：CSP 资本开支韧性延续，CPO 产业化提速

——行业周报

行业走势图



数据来源：聚源

相关研究报告

《AI 应用企业业绩持续兑现，双十一美妆经营表现稳健—行业周报》-2026.10.25

《海外科技：Agent 能力向长程专业任务延伸，SpaceX 与 Meta 业绩均超预期—行业周报》-2026.10.25

《海外科技：AMD 推出 Helion 机架级平台，ASML 评估进一步扩产—行业周报》-2026.10.18

初敏（分析师）
chumin@kysec.cn
证书编号：S0790522080008

叶彬慧（分析师）
yebinhui@kysec.cn
证书编号：S0790526070007

● **AI 硬件：AI 资本开支韧性延续，CPU 与硅光产业链景气上行**

云厂 AI 投入延续，CPU 光通信加速受益。四大云巨头 CY26Q3 业绩分化，但 AI 需求仍强，Google Cloud、Azure、AWS 收入均实现高增长。CSP 资本开支调整更多反映交付节奏差异，并非需求收缩。微软明确 CPU、GPU 及内存存储将驱动后续 CAPEX 加速，谷歌服务器与网络设备投入维持高位。我们认为，2027 年 AI 基础设施景气有望由 GPU 进一步扩散至 CPU 与光通信。**英伟达光引擎扩产提速，硅光产业链受益。**英伟达 CPO 正由技术展示进入产能扩张阶段，计划至 2027 年中将光引擎制造产能提升至当前 15 倍，产业化可见度显著增强。其 CPO 光引擎明确采用硅光 PIC，并通过先进封装与 EIC 集成。Tower 虽不直接供货英伟达，但已确认通过中间厂商间接参与，CPO 放量有望带动硅光 PIC、光源前端等环节需求提升。

● **AI 设备：生产级 AI 负载加速向私有云回流，企业级 AI 基础设施受益**

受 Token 成本、合规治理与数据主权诉求共同驱动，企业生产级 AI 负载正加速向私有云迁移。Cloudera 调研显示当前企业 AI 生产成熟度仍低，EMEA 地区近半数企业已启动 AI 负载跨环境调整，业务韧性、合规要求、成本压降、数据就近部署为核心动因。博通数据验证回迁趋势明确，83%企业已考虑或完成负载从公有云回流，AI 首次成为独立回迁品类，当前 56%企业规划在私有云运行生产级 AI 推理，公有云承载占比同比下滑 15 个百分点。本地部署在 Token 成本上优势显著，中小模型本地化回本周期短，稳态高并发、敏感数据场景性价比突出，私有云开支增速领先公有云，将直接拉动服务器、存储等基建需求，具备生态与交付优势的头部 OEM 厂商优先受益。

● **投资建议：**

AI 硬件：推荐标的：Tower 半导体，受益标的：AMD、台积电；**服务器：**受益标的：联想、Dell、HPE。

● **风险提示：**AI 资本开支增长不及预期风险；Agentic AI 商业化进展不及预期风险；先进制程扩产及技术迭代风险；高速互联技术演进及竞争加剧风险。

目 录

1、 AI 硬件：云厂资本开支分化，算力互联景气共振 3
1.1、 云巨头财报披露：云与 AI 需求延续，CPU 与光通信仍获资本开支支撑 3
1.2、 英伟达计划将 CPO 光引擎产能扩大 15 倍，光通信进入更明确的产业化阶段 4
2、 AI 设备：生产级 AI 负载加速向私有云回流，企业级 AI 基础设施受益 5
3、 港股周度更新 7
4、 模型与算力价格数据更新 9
4.1、 模型信息更新 9
4.2、 算力租赁价格更新 12
5、 投资建议 12
6、 风险提示 13

图表目录

图 1： AI 系统韧性、合规要求、成本压降为 EMEA 企业 AI 跨环境部署核心驱动 6
图 2： 2026 年 AI 首次成为独立的回迁 IT 负载类目 6
图 3： 56%企业计划或正在私有云部署 AI 推理负载 6
图 4： AI 本地部署与商业 LLMAPI 盈亏平衡周期对比 7
图 5： 本周恒生指数（2026.10.26-2026.10.30）在全球主要市场中涨跌幅靠前（单位：%） 7
图 6： 本周（2026.10.26-2026.10.30）恒生科技跌幅较大（%） 8
图 7： 本周（2026.10.26-2026.10.30）恒生航运主题、公共事业类涨幅居前（%） 8
图 8： 2026 年 10 月 13 日以来港股通当日买入成交净额整体为负（单位：亿元） 8
图 9： 恒生沪港通 AH 溢价指数或已触底 9
图 10： OpenRouter 的 token 调用量持续增加 11
图 11： 国产 AI 在 ZenMux 的 Token 市场份额占比超过 65% 11
图 12： B200 租赁价格创阶段性新高（单位：$/小时） 12
图 13： H200 租赁价格高位震荡（单位：$/小时） 12
图 14： H100 租赁价格高位震荡（单位：$/小时） 12
图 15： A100 租赁价格有所下降（单位：$/小时） 12

表 1： 基本面：超预期与不及预期对比 3
表 2： 资本开支：一家上调下限、两家维持、一家因交付顺延而下调 3
表 3： 本周（2026.10.26-2026.10.30）智谱、金斯瑞生物科技与中国移动流入资金居前 8
表 4： 9 月以来新上线模型 9

1、 AI 硬件：云厂资本开支分化，算力互联景气共振

1.1、 云巨头财报披露：云与 AI 需求延续，CPU 与光通信仍获资本开支支撑

云巨头业绩表现不一。截至 10 月 30 日，4 大云巨头（Meta，谷歌，微软，亚马逊）完成了日历年第三季度的财报披露。业绩方面，微软表现最均衡，既有 Azure 的强劲增长，也有正向现金流支撑；谷歌则最能体现“AI 需求超预期不等于集团利润超预期”，云业务高速扩张与搜索广告、非经营性投资损失同时存在。在云业务方面，微软、Meta 集团收入超预期，谷歌和亚马逊小幅不及预期，但 Google Cloud、Azure、AWS 收入分别同比增长 103%、48%、39.3%，均超过市场预期。

表1：基本面：超预期与不及预期对比

公司	超预期／积极信号	不及预期／主要压力
Meta	Q4 收入指引 730 至 760 亿美元，中值 745 亿美元，高于预期 736.5 亿美元；广告及核心业务保持增长。	计提 98.9 亿美元法律费用，GAAP EPS 为 3.83 美元，剔除该项后为 7.40 美元。
谷歌	Google Cloud 收入 307.1 亿美元，同比增长 103%，高于预期 295.7 亿美元；但增长包含 TPU 系统销售贡献。	广告收入约 820 亿美元，低于预期 834 亿美元；GAAP EPS 为 2.12 美元，受到股权投资损失拖累。
微软	调整后 EPS 4.89 美元／4.71 美元，超预期 3.8%；Azure 收入增长 48%；下一季固定汇率增速指引 49%—50%，高于预期 47.6%。	核心云业务未见明显不及预期；Windows OEM 收入同比下降 27%、弱于公司预期；第三方算力交付延期影响扩容节奏。
亚马逊	AWS 收入约 460 亿美元，同比增长 39.3%，高于预期 456.1 亿美元，较上季度进一步提速。	EPS 1.89 美元／1.97 美元，低于预期 4.1%；Q4 收入指引中值 2,435 亿美元，略低于预期 2,441.6 亿美元，集团表现弱于云业务。

资料来源：Meta、Alphabet、微软、亚马逊财报、FactSet、路透社，开源证券研究所

CAPEX 调整存在分化，尚非需求全面收缩。Meta 上调 2026 年指引下限；谷歌、亚马逊维持全年计划，并明确 2027 年显著增长；微软全年指引由 1,750 亿降至 1,700 亿美元，主要反映第三方算力交付延迟，仍预计 FY2027 下半年支出加速。因此，本轮变化更接近扩张期的交付分化，而非四家同步削减 AI 投入。

表2：资本开支：一家上调下限、两家维持、一家因交付顺延而下调

公司	此前 2026 年 CAPEX 计划	最新 2026 年 CAPEX 计划	本次调整判断	2027 年／后续展望
Meta	1,300 至 1,450 亿美元	1,350 至 1,450 亿美元	上调下限；中值上调约 1.8%，上限不变	继续积极投资核心业务及 AI，尚未给出 2027 年量化 CAPEX 指引。
谷歌	1,950 至 2,050 亿美元	1,950 至 2,050 亿美元	维持	明确预计 2027 年 CAPEX 显著增长，具体金额将在后续披露。
微软	约 1,750 亿美元	约 1,700 亿美元	下调约 2.9%，主要反映第三方算力交付延	预计 FY2027 下半年，即 2027 年 1—6 月，

公司	此前 2026 年 CAPEX 计划	最新 2026 年 CAPEX 计划	本次调整判断	2027 年/后续展望
			期	CAPEX 加速；驱动包括 CPU、GPU 及延期交付的第三方算力。
亚马逊	约 2,200 亿美元	约 2,200 亿美元	维持	明确预计 2027 年投入显著增加；2026 年投资主要用于 AWS。

资料来源：Meta、Alphabet、微软、亚马逊财报、开源证券研究所

微软将 CPU 列为 FY2027 下半年 CAPEX 加速的主要来源之一。微软在 2026 年 10 月 28 日的 FY2027Q1（对应 CY26Q3）电话会上披露，当季 CAPEX 为 452 亿美元，约四分之三用于短寿命资产，主要是 CPU 和 GPU，并包含更高的内存及存储成本。管理层预计，FY2027 下半年，即 2027 年 1—6 月，CAPEX 将加速，主要驱动包括 CPU 及其相关内存、存储成本，GPU，以及延期交付的第三方算力。说明 AI 数据中心的开支开始从 GPU 逐步走向 GPU+CPU、内存、存储并行。

谷歌服务器及网络设备投入维持较高强度，为高速光通信需求持续增长提供了客户侧支撑。谷歌在 2026 年 10 月 28 日的 CY26Q3 财报电话会上提及，当季资本开支约 514 亿美元，绝大部分用于支持 AI 的技术基础设施；其中，技术基础设施投资约 65%用于服务器及网络设备，剩余 35%用于数据中心。前者通常使用年限为 6 年，后者为 7—40 年。表明 AI 投入不仅涉及数据中心建设，也包含大规模计算与互联设备部署。

我们认为，CSP 资本开支维持高位，有望推动 AI 基础设施投资的受益范围从 GPU 进一步扩散至 CPU 与光通信，为 2027 年两大环节的需求增长提供支撑。CPU 方面，随着 Agentic AI 应用拓展，任务编排、工具调用、代码执行及数据处理等工作负载增加，新增 CPU 需求有望从 GPU 配套进一步延伸至独立任务执行，带动服务器 CPU 部署规模扩大及高核心数产品渗透。光通信方面，算力集群扩容叠加 Scale-up/Scale-out 网络升级，有望推动 800G 向 1.6T 演进，并随着跨机柜连接增加，带动高速光模块、CW 激光器、硅光芯片及 TIA/Driver 需求增长与产品结构升级。两大环节不仅受益于资本开支规模扩张，也有望受益于计算与互联配置升级。

1.2、 英伟达计划将 CPO 光引擎产能扩大 15 倍，光通信进入更明确的产业化阶段

英伟达 CPO 从“展示”进入“加速扩产”。10 月 21 日，英伟达宣布与多家半导体制造伙伴扩大产能合作，在公告之中提及英伟达与日月光旗下矽品精密计划到 2027 年中，将 CPO 所需光引擎制造产能提高至目前的 15 倍，并以 2028 年底在扩产基础上再翻倍为目标。在过去，英伟达展示的 CPO 相关信息更多有关 CPO 新产品展示、光引擎技术验证，以及硅光 PIC 工艺开发，意味着这些都是未进入量产之前的早期展示。但本次的公告表明 CPO 已经开始有量产供应链、封装测试产能、明确扩产倍数及资金支持。这意味着英伟达不仅在研发 CPO 产品，还在为后续更大规模的制造和交付能力做准备。

Tower 半导体间接供货英伟达。Tower 于 2026 年 2 月 11 日举行的 2025Q4 电话会的分析师问答环节，CEO 表示：公司并不直接向英伟达出货，而是通过中间厂商

交付。

我们认为，这比单纯发布下一代 CPO 规格，更有助于提高产业化进程的可见度。CPO 的量产或推动产业链价值向光源、PIC、模拟前端、光引擎封装及系统集成重新分配。

2、 AI 设备：生产级 AI 负载加速向私有云回流，企业级 AI 基础设施受益

Token 成本、治理要求和主权要求提升，企业或将加速跨环境 AI 部署。根据 Cloudera2026 年 9 月对 1500 位全球 IT 领导者的调查结果：

（1）**实际生产环境中 AI 执行尚不成熟**：近 98%组织承认曾因生产环境中 AI 应用出现幻觉、错误或意外的模型行为而被迫撤回人工智能部署。

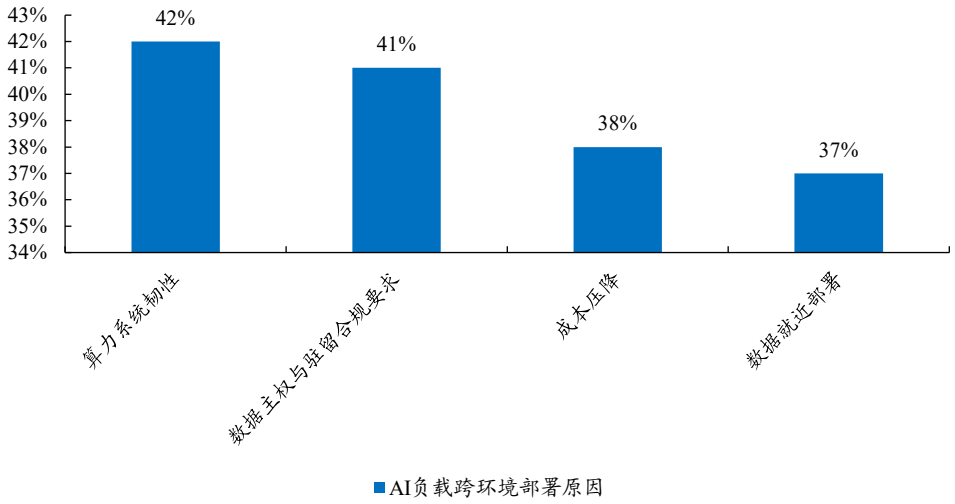
（2）**合规成本限制 AI 规模扩张速度**：75%EMEA 地区 IT 领导者认为数据保护、隐私合规等治理问题是他们扩展 AI 的主要障碍，其中领导者表示合规支出增长速度超过预期。

（3）**基建和推理成本扩张驱动 AI 负载迁移**：使用成本的 47%EMEA 企业将 AI 工作负载迁移到其他环境，而 25%企业已经暂停或彻底取消相关项目。

（4）**主权要求影响 AI 负载部署位置**：82%的 EMEA 决策者认为管辖地或主权的数据驻留要求影响 AI 负载的运行地点，但 58%的决策者承认，在云端或本地平台之间迁移工作负载需要代价高昂的程序重新设计。

欧洲、中东及非洲（EMEA）地区的受访企业将以下几点列为 AI 工作负载跨多环境部署运行的核心动因：业务韧性（42%，算力架构在故障、流量波动、外部冲击下保障业务连续运行的能力）、数据主权与数据驻留合规要求（41%，数据存储、处理需在指定地域内完成）、成本压降（38%，通过混合部署、资源潮汐调度、多云比价等方式实现全链路算力成本优化），以及数据就近部署诉求（37%，算力节点与数据源同域部署以降低数据传输时延、减少跨网带宽成本）。**企业对 AI 业务适配能力的需求持续升级，跨环境部署灵活性的重要性与日俱增**。

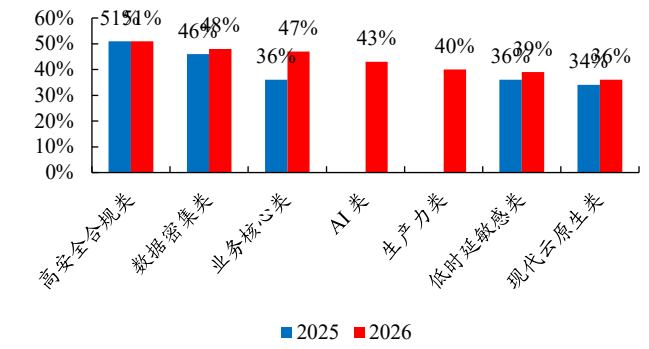
图1：AI 系统韧性、合规要求、成本压降为 EMEA 企业 AI 跨环境部署核心驱动



数据来源：Cloudera、开源证券研究所

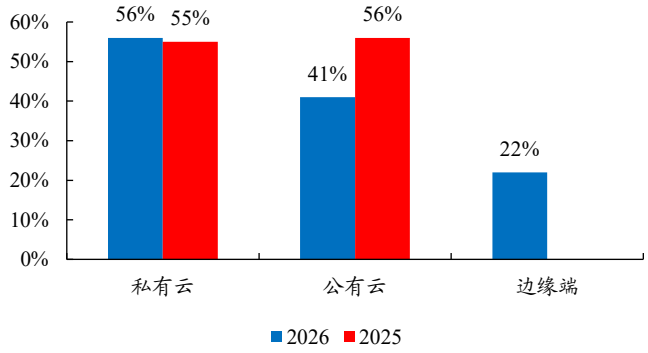
AI 公有云和私有云面临再平衡，生产性 AI 推理负载向私有云迁移。根据博通，2026 年 IT 负载向私有云回迁的趋势进一步升温，97%的 IT 领导者认为他们的部分公有云支出被浪费，52%的领导者认为浪费超过其公有云总预算的 25%。83%的企业正在考虑将工作负载从公有云迁移到私有云，其中 50%的企业已经回流部分工作负载，2026 年 AI 首次成为独立的回迁负载品类，公有云成本过高且治理不足问题在企业 AI 从试点和培训阶段进入规模扩展阶段时迫使工作负载回归企业内部。截至 2026 年 6 月，56%的调查企业正在私有云上运行或计划运行生产 AI 推理，而用于相同工作负载的公共云使用率同比下降 15pct。

图2：2026 年 AI 首次成为独立的回迁 IT 负载类目



数据来源：Broadcom、开源证券研究所

图3：56%企业计划或正在私有云部署 AI 推理负载



数据来源：Broadcom、开源证券研究所

本地部署在 Token 成本优势明显，整体 TCO 需负载类型及综合使用量。根据 DreamFactory 分析，本地运行的开放权重模型每百万 Token 的成本大约比高级商业 API 低 18 倍，而在卡内基梅隆大学一项针对 54 种部署场景的比较结果中，中小型模型本地部署优势更为明显：

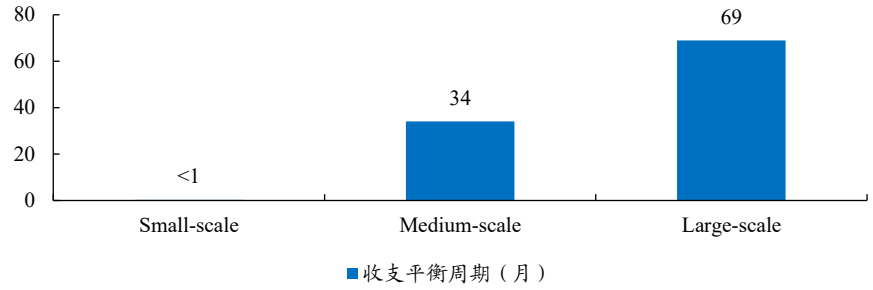
（1）**小型开源部署**：在单个消费级 GPU 上运行的 320 亿参数模型（售价约为 2000 美元），只需 0.3 个月即可与 ClaudeOpus 等高级 API 实现收支平衡，；

（2）**中等规模部署**：实现盈亏平衡的时间范围更广，大约为 2.3 到 34 个月，具体取决于部署规模和硬件选择。

（3）**大规模部署：**可能需要长达 69.3 个月才能与最便宜的商业供应商实现收支平衡。

对于 AI 推理工作中涉及高容量、稳定的工作负载（分类、提取、预测、内部搜索），受监管或敏感数据（PII、财务记录、健康数据、源代码）等，本地迁移必要性较为充分；而对于新颖的推理、开放式的讨论和前沿研究任务，租用封闭式模型仍然有意义。根据博通，2026 年企业私有云三年展望期支出意愿增长 21 个百分点，增速约为公有云的两倍，AI 基建本地化将直接拉动通用服务器、AI 推理服务器、配套存储与超融合一体机的采购需求释放，头部 OEM 厂商凭借与 VMware 等云平台的深度生态认证、全场景产品矩阵、行业定制化交付能力及本地化服务网络，将率先受益。

图4：AI 本地部署与商业 LLMAPI 盈亏平衡周期对比

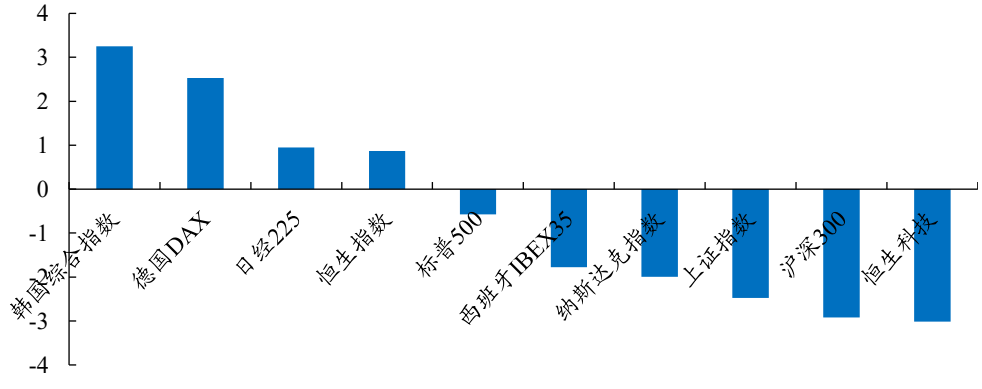


数据来源：卡内基梅隆大学、开源证券研究所

3、港股周度更新

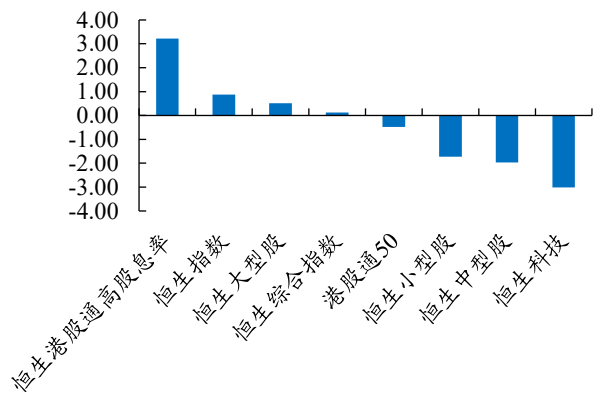
本周恒生指数上涨 0.87%，在全球主要市场中涨跌幅靠前，从成分上看，恒生航运主题等涨幅居前，恒生科技本周下跌 3.02%。

图5：本周恒生指数（2026.10.26-2026.10.30）在全球主要市场中涨跌幅靠前（单位：%）



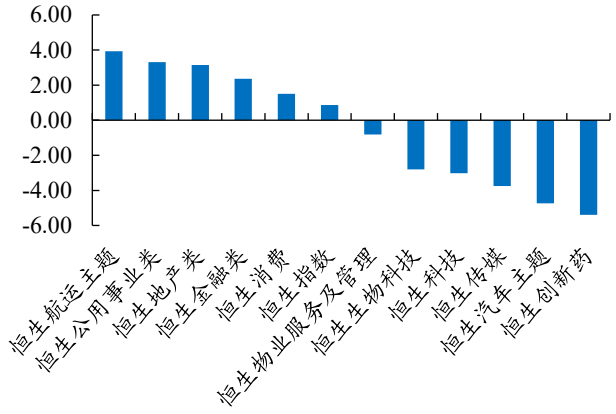
数据来源：Wind、开源证券研究所

图6：本周（2026.10.26-2026.10.30）恒生科技跌幅较大（%）



数据来源：Wind、开源证券研究所

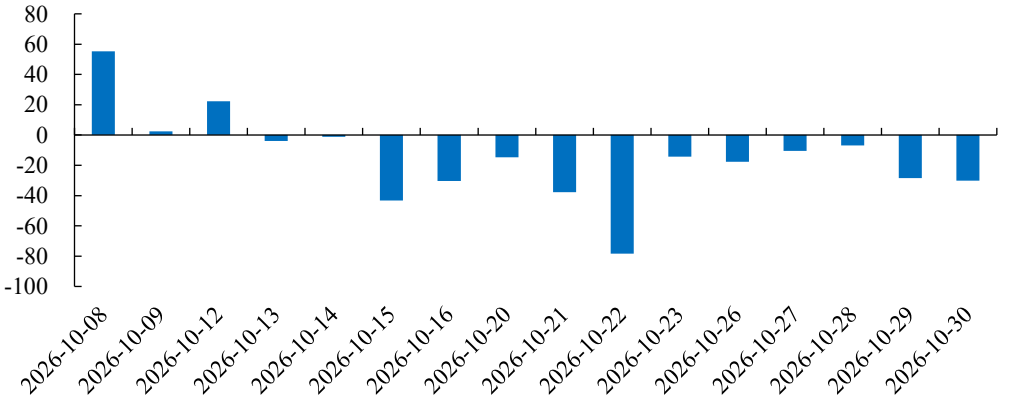
图7：本周（2026.10.26-2026.10.30）恒生航运主题、公共事业类涨幅居前（%）



数据来源：Wind、开源证券研究所

本周港股通成交净额为流出 93.65 亿元，环比增加，从头部活跃个股交易层面来看，智谱（+16.5 亿港元）、金斯瑞生物科技（+4.3 亿港元）、中国移动（+4.2 亿港元）资金净流入居前。

图8：2026 年 10 月 13 日以来港股通当日买入成交净额整体为负（单位：亿元）



数据来源：Wind、开源证券研究所

表3：本周（2026.10.26-2026.10.30）智谱、金斯瑞生物科技与中国移动流入资金居前

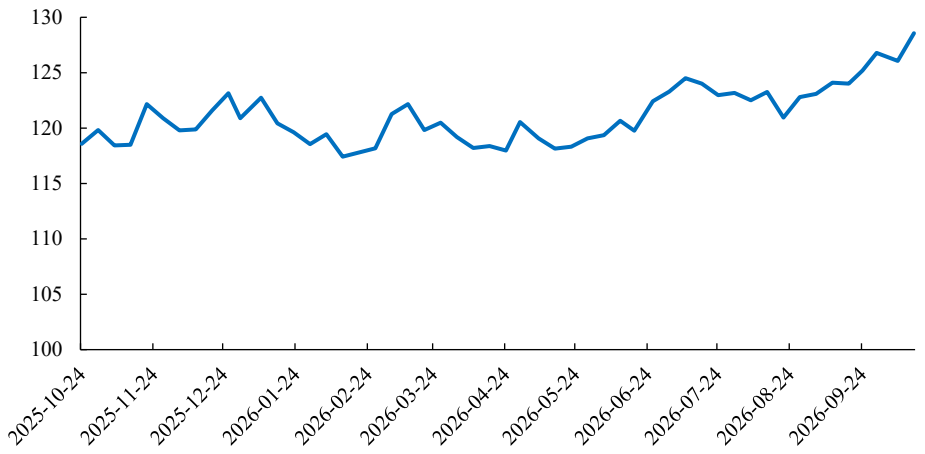
股票代码	股票名称	成交净买入（亿港元）	港股通持股数占总股本比例
2513.HK	智谱	16.5	16.1%
1548.HK	金斯瑞生物科技	4.3	19.4%
0941.HK	中国移动	4.2	13.2%
3308.HK	中际旭创	4.1	9.4%
0100.HK	MINIMAX-W	3.5	12.2%
9888.HK	百度集团-W	2.7	3.9%
6681.HK	脑动极光-B	0.2	19.6%
1398.HK	工商银行	0.1	40.9%
6651.HK	五一视界	-0.1	5.8%
0148.HK	建滔集团	-0.4	20.5%
2269.HK	药明生物	-0.4	23.4%

股票代码	股票名称	成交净买入（亿港元）	港股通持股数占总股本比例
0992.HK	联想集团	-0.9	6.9%
3690.HK	美团-W	-1.3	18.8%
1888.HK	建滔积层板	-3.1	10.5%
1347.HK	华虹宏力	-3.2	14.3%
0700.HK	腾讯控股	-3.5	12.0%
1810.HK	小米集团-W	-4.0	19.4%
9992.HK	泡泡玛特	-6.5	23.1%
0883.HK	中国海洋石油	-8.5	22.4%
6869.HK	长飞光纤光缆	-13.3	41.3%
0981.HK	中芯国际	-18.0	26.6%

数据来源：Wind、开源证券研究所

本周恒生沪港通 AH 溢价指数为 132.49，环比有所下降，AH 股溢价触底回弹，随着后续更多 A 股科技类标的赴港上市，或带动港股板块的“成长性溢价”持续向 A 股靠拢。

图9：恒生沪港通 AH 溢价指数或已触底



数据来源：Wind、开源证券研究所

4、 模型与算力价格数据更新

4.1、 模型信息更新

9 月以来全球大模型行业进入新品密集发布期，海内外头部厂商先后推出十余款新模型，产品定位与定价梯队分化明显。本周（2026.10.26-2026.10.30）新上线模型包括闭源模型 GPT-6.1 Terra 与 Gemini 4 Carbon。

表4：9 月以来新上线模型

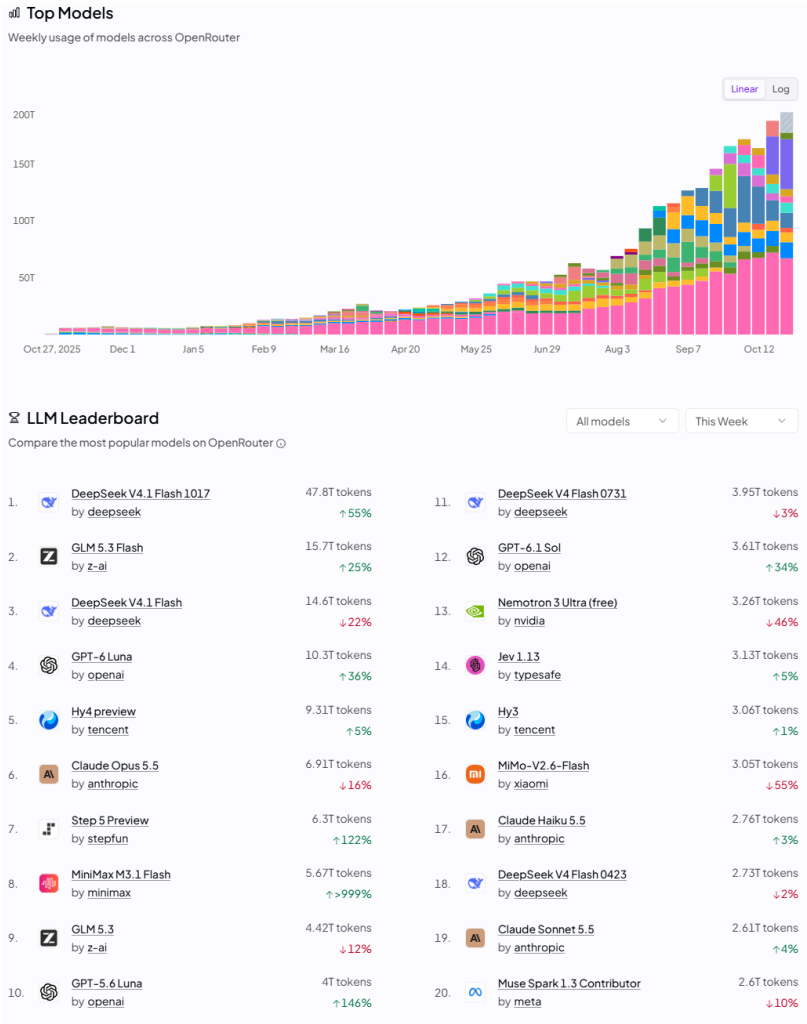
已发布模型	发布时间	厂商	输入价格（/M）	缓存命中输入价格（/M）	输出价格（/M）	上下文窗口
Claude Fable 5.1	2026-09-01	Anthropic	$10.00	$0.25	$50.00	1M
Claude Mythos 5.1	2026-09-01	Anthropic	$10.00	$0.25	$50.00	1M

已发布模型	发布时间	厂商	输入价格(/M)	缓存命中输入价格(/M)	输出价格(/M)	上下文窗口
Gemini 3.8 Flash	2026-09-02	Google	$0.75	$0.08	$3.75	1M
Muse Spark 1.3	2026-09-02	Meta	$1.25	$0.15	$4.25	1M
GPT-6 Astra	2026-09-03	OpenAI	$10.00	$1.00	$50.00	1.05M
Kimi K2.8 Preview	2026-09-11	月之暗面	目前只在 Kimi Code 与 Kimi Work 上线			1M
GLM-5.3-FlashX	2026-09-18	智谱	¥2.00	¥0.57	¥7.00	1M
Step 5 Preview	2026-09-20	阶跃星辰	¥7.00	¥0.35	¥20.00	1M
MiMo-V2.6-Pro	2026/9/22	小米	¥3.00	¥0.03	¥6.00	1M
MiMo-V2.6-Flash	2026/9/22	小米	¥1.00	¥0.02	¥2.00	1M
Grok 4.7	2026/9/22	SpaceXAI	$2.00	$0.50	$6.00	500K
GPT-6 Sol	2026/9/23	OpenAI	$2.00	$0.20	$10.00	1.05M
GPT-6 Luna	2026/9/23	OpenAI	$0.10	$0.01	$0.50	1.05M
Claude Opus 5.5	2026/9/23	Anthropic	$4.00	$0.20	$20.00	1M
MiniMax M3.1 Flash Preview	2026/9/27	MiniMax	目前只在 MiniMax Code 上线			1M
Claude Sonnet 5.5	2026/9/28	Anthropic	$2.00	$0.20	$10.00	1M
GPT-6.1 Sol	2026/9/29	OpenAI	$2.00	$0.10	$10.00	1.05M
Gemini 4 Argon	2026/10/01	Google	$2.00	$0.10	$10.00	1M
Beam	2026/10/05	Reflection AI	暂无			1M
Claude Haiku 5.5	2026/10/8	Anthropic	$0.10	$0.01	$0.50	1M
Muse Wonder Full	2026-10-12	Meta	目前只在 Muse Code 订阅中开放			1M
Muse Wonder	2026-10-12	Meta	$4.00	$0.20	$20.00	1M
Muse Spark Open	2026-10-15	Meta	$1.25	$0.15	$4.25	1M
DeepSeek-V4.1-Flash-1017(高峰)	2026-10-17	DeepSeek	$0.30	$0.006	$1.20	1M
DeepSeek-V4.1-Flash-1017(空闲)	2026-10-17	DeepSeek	$0.15	$0.003	$0.60	1M
MiniMax M3.1 Flash	2026/10/19	MiniMax	$0.15	$0.02	$0.60	1M
Kimi K3.1 Pro Highspeed	2026/10/24	月之暗面	$4.00	$0.20	$20.00	1M
Kimi K3.1 Pro	2026/10/24	月之暗面	$2.00	$0.20	$10.00	1M
GPT-6.1 Terra	**2026-10-27**	**OpenAI**	**$1.00**	**$0.05**	**$5.00**	**1.05M**
Gemini 4 Carbon	**2026-10-28**	**Google**	**$2.00**	**$0.10**	**$10.00**	**1M**

资料来源：各模型厂商官网、IT 之家、Openrouter、开源证券研究所

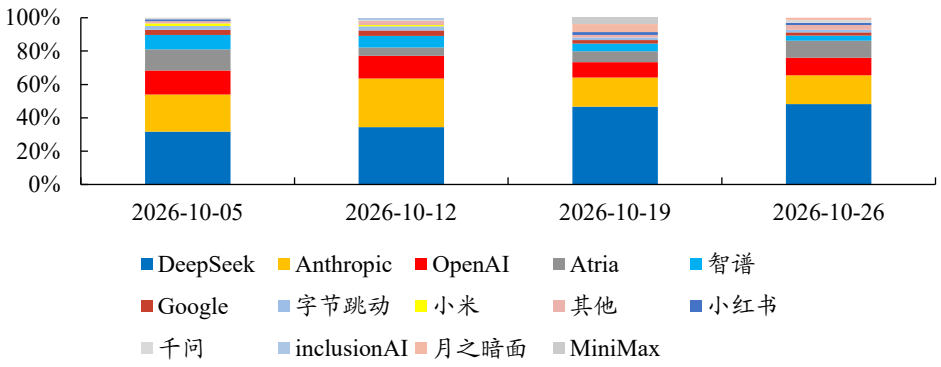
本周（2026.10.26-2026.10.30）OpenRouter 的 token 调用量持续提升，其中国产模型 DeepSeek V4.1 Flash、GLM 5.3 Flash 与 Hy4 preview 位居前列。截至 10 月 26 日，在 ZenMux 统计的 Token 市场份额中，国产 AI 占比超过 65%，DeepSeek 市场份额保持第一，占比为 48.3%。依托开源生态及出色的性价比优势，国产模型持续扩大市场影响力。

图10：OpenRouter 的 token 调用量持续增加



资料来源：OpenRouter 官网

图11：国产 AI 在 ZenMux 的 Token 市场份额占比超过 65%



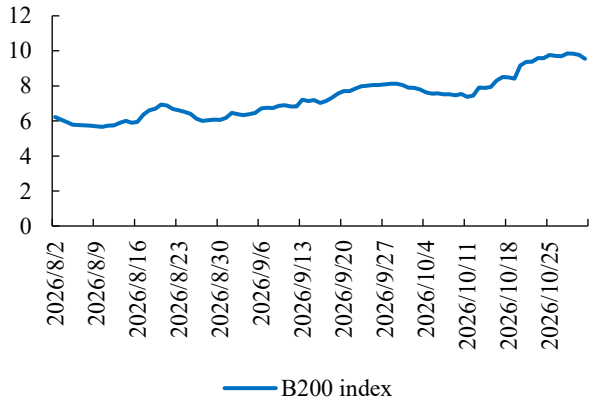
数据来源：ZenMux、开源证券研究所

4.2、 算力租赁价格更新

GPU 租赁价格整体高位运行，不同性能 GPU 价格走势有所分化。根据 Ornn 统计的基于 neocloud 市场的租赁交易价格，GPU 租赁价格整体保持高位运行，但不同性能 GPU 价格走势有所分化，高性能型号 B200 租赁价格走势坚挺，创下阶段性新高；而 H200、H100 与 A100 租赁价格走势则偏震荡。

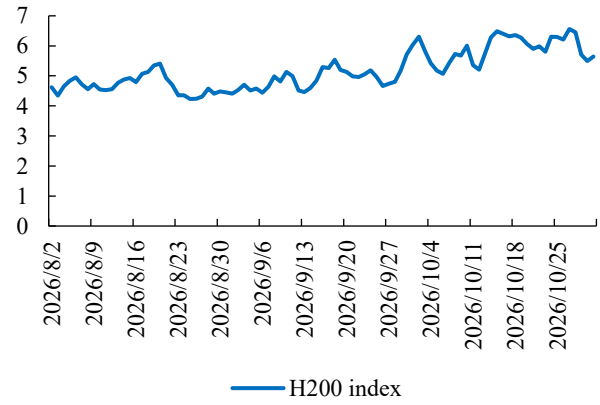
截至 10 月 31 日，B200 租赁价格为 9.54 $/小时；H200 租赁价格为 5.64 $/小时；H100 租赁价格为 3.00$/小时；A100 租赁价格为 1.04 $/小时。

图12：B200 租赁价格创阶段性新高（单位：$/小时）



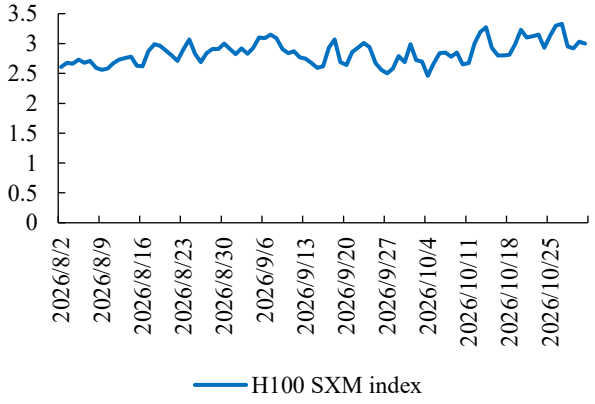
数据来源：Ornn、开源证券研究所

图13：H200 租赁价格高位震荡（单位：$/小时）



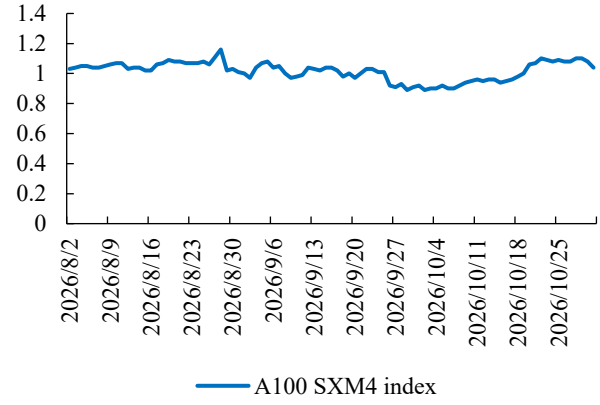
数据来源：Ornn、开源证券研究所

图14：H100 租赁价格高位震荡（单位：$/小时）



数据来源：Ornn、开源证券研究所

图15：A100 租赁价格有所下降（单位：$/小时）



数据来源：Ornn、开源证券研究所

5、 投资建议

AI 硬件：我们认为，本周 CSP 披露的财报表明 CAPEX 处于相对企稳，但开始分化的情况，而后续的 CAPEX 或将继续投入 CSP 更加认可的发展方向，本次微软明确提及了 CPU，以及谷歌明确提及了光互联，再叠加英伟达扩产 CPO 光引擎的印证，更加说明未来发展方向偏向于 CPU 和光互联。推荐标的：在“互联”方向的硅光/

硅锗“卖铲者”Tower 半导体；受益标的：因 Agent 而 CPU 受益的 AMD，和不可或缺的先进制程代工厂台积电。

服务器：服务器 OEM 行业进入 AI 算力+传统 IT 换机双轮驱动的上行周期，需求从单一超大规模云厂扩散至新云/企业/主权算力，产品形态从单机硬件向整机柜液冷一体化方案升级；行业盈利模型重构，具备客户结构、全栈方案、供应链调度、高密度交付能力的头部 OEM 厂商业绩确定性更强，受益标的：联想、Dell、HPE。

6、 风险提示

AI 资本开支增长不及预期风险：若全球 CSP 及 AI 厂商放缓资本开支，可能导致 AI 服务器、先进制程及高速互联芯片需求不及预期。

Agentic AI 商业化进展不及预期风险：若 Agent 应用渗透率、并发规模及 CPU 计算负载增长低于预期，可能影响独立 CPU 机架的市场需求。

先进制程扩产及技术迭代风险：若 2nm 制程良率爬坡、CoWoS 先进封装产能扩张不及预期，可能制约 AI 芯片交付，并影响相关产业链盈利能力。

高速互联技术演进及竞争加剧风险：若 1.6T、3.2T 产品渗透不及预期，或 CPO 等新架构改变高速模拟芯片价值分配，可能影响 SiGe 代工需求及相关厂商盈利能力。

特别声明

《证券期货投资者适当性管理办法》、《证券经营机构投资者适当性管理实施指引（试行）》已于2017年7月1日起正式实施。根据上述规定，开源证券评定此研报的风险等级为R4（中高风险），因此通过公共平台推送的研报其适用的投资者类别仅限定为境内专业投资者及风险承受能力为C4、C5的普通投资者。若您并非境内专业投资者及风险承受能力为C4、C5的普通投资者，请取消阅读，请勿收藏、接收或使用本研报中的任何信息。
因此受限于访问权限的设置，若给您造成不便，烦请见谅！感谢您给予的理解与配合。

分析师承诺

本研究报告的署名人员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，以勤勉的职业态度，独立、客观地出具本报告，并对内容和观点负责。本报告清晰准确地反映了署名人员的研究观点，所包含的分析基于各种假设，不同假设可能导致分析结果出现重大不同。本报告采用的各种估值方法及模型均有其局限性，估值结果不保证所涉及证券能够在该价格交易。本报告署名人员保证他们报酬的任何一部分不曾与，不与，也将不会与本报告中具体的推荐意见或观点有直接或间接的联系。

股票投资评级说明

	评级	说明
证券评级	买入（Buy）	预计相对强于市场表现 20%以上；
	增持（outperform）	预计相对强于市场表现 5%～20%；
	中性（Neutral）	预计相对市场表现在－5%～＋5%之间波动；
	减持（underperform）	预计相对弱于市场表现 5%以下。
行业评级	看好（overweight）	预计行业超越整体市场表现；
	中性（Neutral）	预计行业与整体市场表现基本持平；
	看淡（underperform）	预计行业弱于整体市场表现。
备注：评级标准为以报告日后的 6~12 个月内，证券相对于市场基准指数的涨跌幅表现，其中 A 股基准指数为沪深 300 指数、港股基准指数为恒生指数、新三板基准指数为三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）、美股基准指数为标普 500 或纳斯达克综合指数。我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重建议；投资者买入或者卖出证券的决定取决于个人的实际情况，比如当前的持仓结构以及其他需要考虑的因素。投资者应阅读整篇报告，以获取比较完整的观点与信息，不应仅仅依靠投资评级来推断结论。		

分析、估值方法的局限性说明

本报告所包含的分析基于各种假设，不同假设可能导致分析结果出现重大不同。本报告采用的各种估值方法及模型均有其局限性，估值结果不保证所涉及证券能够在该价格交易。

法律声明

开源证券股份有限公司是经中国证监会批准设立的证券经营机构，已具备证券投资咨询业务资格。

本报告仅供开源证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为客户。本报告是发送给开源证券客户的，属于商业秘密材料，只有开源证券客户才能参考或使用。

本报告是基于本公司认为可靠的已公开信息，但本公司不保证该等信息的准确性或完整性。本报告所载的资料、工具、意见及推测只提供给客户作参考之用，并非作为或被视为出售或购买证券或其他金融工具的邀请或向人做出邀请。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动，过往的业绩表现不应作为其日后表现的预示。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。客户应当考虑到本公司可能存在可能影响本报告客观性的利益冲突，不应视本报告为做出投资决策的唯一因素。本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。本公司未确保本报告充分考虑到个别客户特殊的投资目标、财务状况或需要。本公司建议客户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。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人的投资建议。在任何情况下，本公司不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。若本报告的接收人非本公司的客户，应在基于本报告做出任何投资决定或就本报告要求任何解释前咨询独立投资顾问。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。

开源证券在法律允许的情况下可参与、投资或持有本报告涉及的证券或进行证券交易，或向本报告涉及的公司提供或争取提供包括投资银行业务在内的服务或业务支持。开源证券可能与本报告涉及的公司之间存在业务关系，并无需事先或在获得业务关系后通知客户。

本报告的版权归本公司所有。本公司对本报告保留一切权利。未经本公司事先书面授权，本报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

开源证券研究所

上海
地址：上海市浦东新区世纪大道1788号陆家嘴金控广场1号楼3层
邮编：200120
邮箱：research@kysec.cn

深圳
地址：深圳市福田区金田路2030号卓越世纪中心1号楼45层
邮编：518000
邮箱：research@kysec.cn

北京
地址：北京市西城区西直门外大街18号金贸大厦C2座9层
邮编：100044
邮箱：research@kysec.cn

西安
地址：西安市高新区锦业路1号都市之门B座5层
邮编：710065
邮箱：research@kysec.cn