

海内外云厂商发展与现状（二） AI投入、算力建设梳理与ROI测算

行业研究 · 海外市场专题

互联网 · 互联网 II

投资评级：优于大市（维持）

证券分析师：张伦可

0755-81982651

zhanglunke@guosen.com.cn

S0980521120004

证券分析师：刘子谭

liuzitan@guosen.com.cn

S0980525060001

证券分析师：张昊晨

zhanghaochen1@guosen.com.cn

S0980525010001

- **海内外云厂商资本开支复盘：**从Capex投资节奏上，海外从23Q3开始增长，国内约晚一年。从Capex增速来看海内外仍处于50%以上同比增长。当前微软云、亚马逊云与阿里云等头部云厂商资本开支金额与全年云收入基本相当。**海外云厂商与阿里云已将超100%的净利润、60%以上的经营性现金流进行资本开支投入**，且增速指引依然显著高于业务收入/利润增速。为了维持增长需求，各家进行一定的AI基础设施金融化操作或加大融资租赁。从Capex投资周期来看，传统云扩张投资约持续四年、且投入到收入兑现约一年时间差。但此次AI供需错配度更高、建设周期更长、商业模式亦有所变化，导致收入兑现节奏更慢，**中性预期下此次AI投资可能会持续五到六年扩张周期。**
- **云服务提供商约贡献英伟达数据中心收入的50%。**英伟达表示2025年全球数据中心投入预计6000亿美元（两年内翻倍），2030年前有望达到3-4万亿美元，其中1GW英伟达数据中心约500到600亿美元投入。测算各家算力规模与增长如下：
 - ① **微软：**计划本年度（25Q3-26Q2）将AI总容量提升超 80%，**未来两年将数据中心总规模扩大约一倍、约达10GW；**
 - ② **谷歌：**测算2023到2025三年累计投入超1700亿美元，其芯片包括GPU与TPU（算力角度约各占一半），其中TPU自用更多，GPU租外部更多。
 - ③ **亚马逊：**25Q3业绩会表示过去12个月AWS新增3.8吉瓦电力容量，第四季度预计再增1吉瓦以上，**AWS目标2027年底实现算力翻倍。**自研AI芯片Trainium2已发展为几十亿美元业务，相较于其他GPU选项具备30%-40%的价格性能优势。
 - ④ **阿里巴巴：**未来三年将投入至少3800亿元人民币用于AI基础设施建设。**2032年，阿里云栖大会表示全球数据中心的能耗将比2022年提升10倍，测算阿里云能耗2032年将达到15GW、十年总资本开支规模累计约1.5万亿元。**
- **云厂自研芯片布局与进展：ASIC产品预计明后年进入将密集落地期。**当前测算英伟达从实际算力角度预计当前占据80%以上份额，尤其是在训练场景。ASIC等其他玩家对芯片自主可控和降本诉求强烈，**测算GB200和谷歌TPU v6e性价比与能效比基本可比。**
- 从AI商业化变现角度，AI云业务主要包括GPU租赁（主要满足训练需求、类似传统云IaaS业务）、MaaS/API服务（满足推理需求、利润率差异大）、PaaS/SaaS 服务（提供研发工具、利润率高）。**测算微软云、谷歌云和阿里云AI业务2030年现金流贡献转正，2030年ROIC可达10%以上，即六到十年可收回投资周期。**
- **投资建议：**伴随模型能力提升、AI云与算力相关需求快速增长，推荐AI云平台厂商：微软（MSFT.O）、谷歌（GOOG.O）、亚马逊（AMZN.O）、阿里巴巴（9988.HK）、腾讯控股（0700.HK）与芯片供应商英伟达（NVDA.O）。
- **风险提示：**宏观经济波动风险、AI技术进展不及预期风险、行业竞争加剧风险、AI幻觉问题影响等。

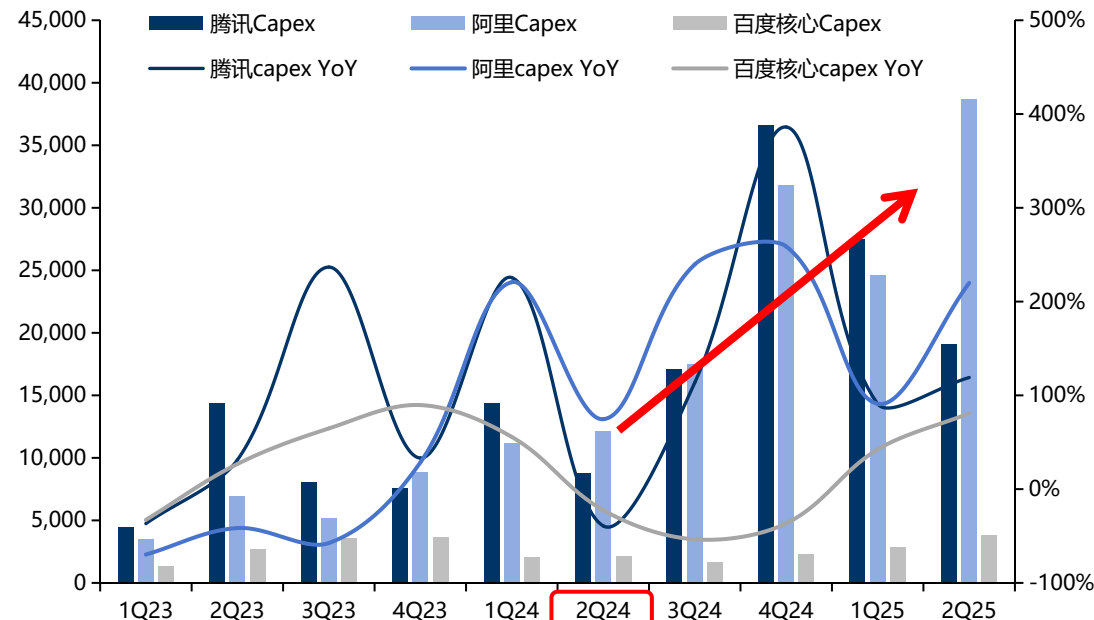
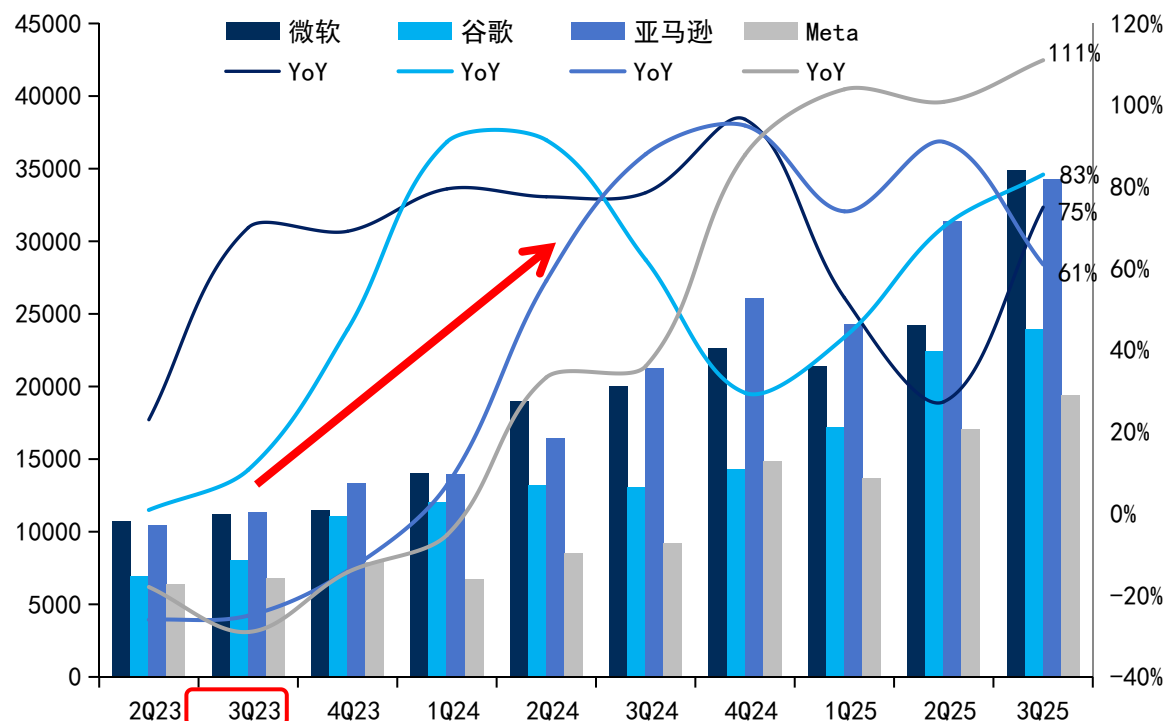
- 【 01 】 海内外云厂商资本开支复盘分析
- 【 02 】 云厂商算力规模与建设计划
- 【 03 】 云厂自研芯片布局与进展
- 【 04 】 云厂AI云收入、AI ROI测算与估值

海内外Capex投资节奏：海外从23Q3开始增长，国内约晚一年

- **从时间点来看，国内约晚海外一年开始增加投资。**海外资本开支显著加速从2023年下半年开始（ChatGPT发布半年后），其中以微软最为迅速，然后谷歌、亚马逊与META陆续加大投入。国内整体资本开支规模增长从2024年中旬开始，并且由于低基数增幅更为明显。
- **从节奏来看，但海外节奏更平稳、规模领先。**国内则是经历22年与23年资本开支调整（负增长）后激进扩张，也由于算力限制和AI策略原因增速波动性更为明显。

图：各互联网与云厂Capex季度投资与变化情况（百万美元、%）

图：国内云厂商资本开支与同比变化预测（百万元/%）



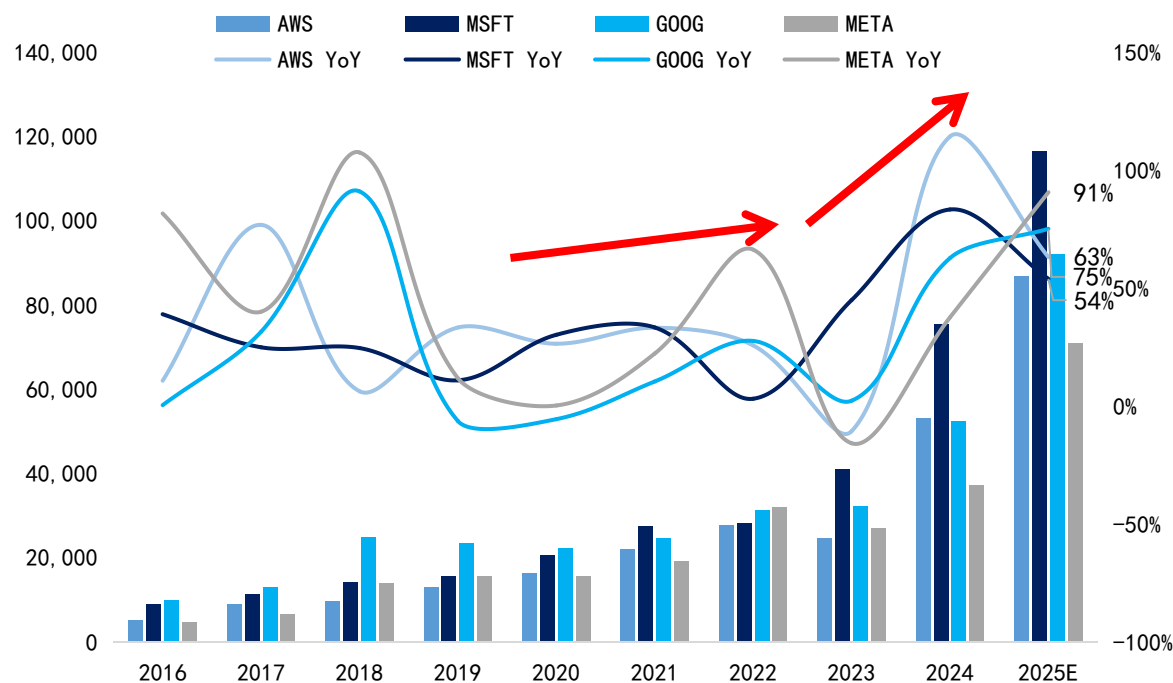
资料来源：公司财报、彭博一致性预期、国信证券经济研究所整理

资料来源：公司财报、彭博一致性预期、国信证券经济研究所整理测算

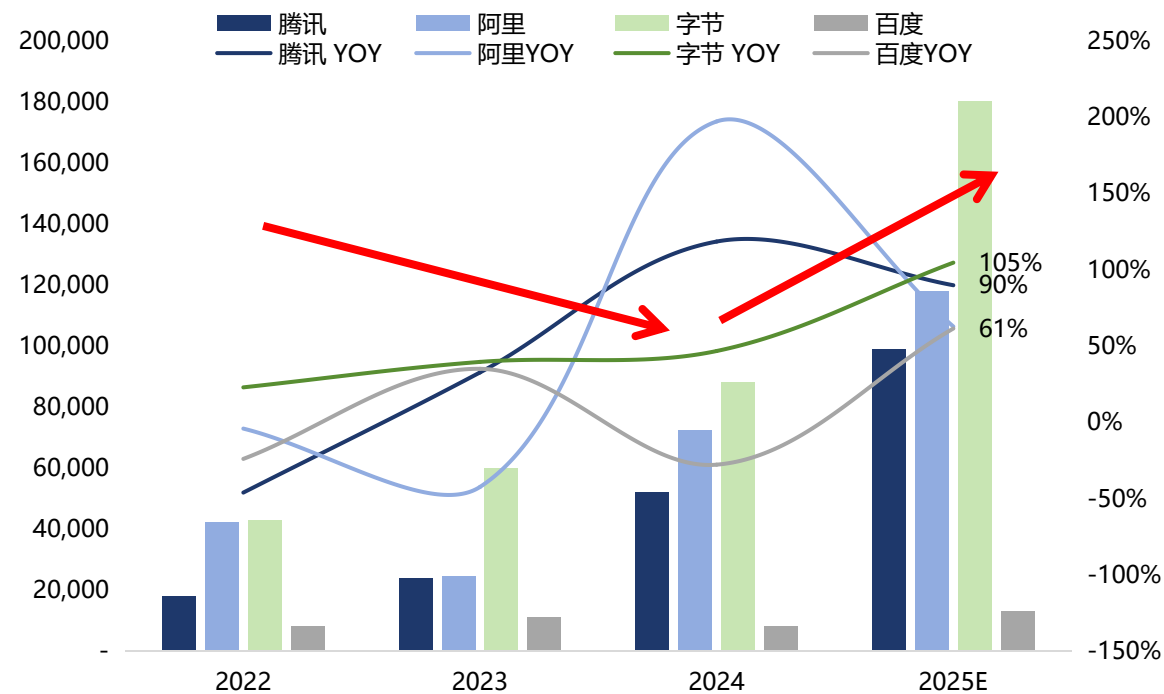
海内外Capex增速：仍处于50%以上同比增长，Capex总量显著提升

- 2025年海外厂商各家全年Capex投入均处于50%以上同比增长：
- 微软：约1160亿美元(同比+54%)；亚马逊：1250亿美元(同比+61%)；谷歌：910-930亿美元(同比+73%-77%)；Meta:700-720亿美元(同比+78%-84%)；
- 国内厂商在23年低资本开支基数的基础上，连续两年实现各家全年Capex投入平均约80%同比增长：
- 阿里：约1180亿元(同比+63%)；腾讯：约1000亿元(+90%)；字节：约1800亿元(同比+105%)；百度核心：约130亿元(+60%)。

图：海外互联网与云厂Capex年度投资与增速变化情况（百万美元、%）



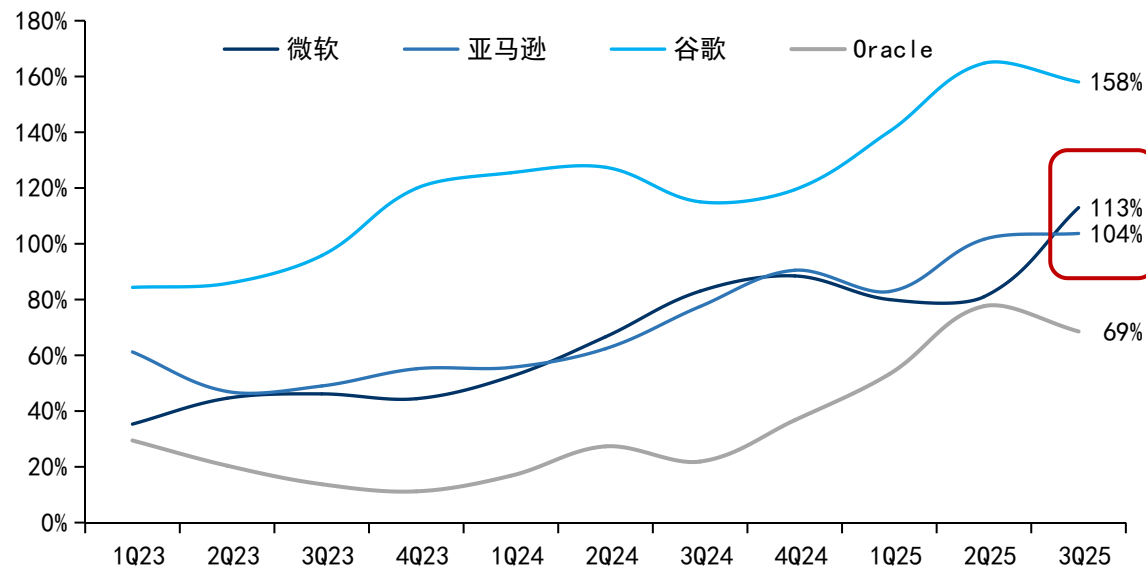
图：国内互联网与云厂Capex年度投资与增速变化情况（百万元、%）



海内外Capex量级：头部云厂投资与云收入相当，小体量云厂比例更高

- **微软云、亚马逊云与阿里云等头部云厂商资本开支金额与全年云收入基本相当。**相比正常发展阶段30-40%云收入比例的投入，目前显著处于加速扩张的阶段。一方显示了AI云更高的“资金壁垒”，通过大规模基建投入抢占赛道话语权，需要“重资产、长周期”投入。另一方面，当前高比例资本开支会压低短期利润率，更加强调短期投入与长期盈利的平衡博弈。
- **谷歌：**投入持续大幅上升，且占比达到了165%。一方面由于自身云业务收入体量小，另一方面除了云业务上需求外，其资本开支也用于TPU芯片研发建设、Gemini模型训练以及自身业务AI需求。
- **Oracle：**此前以传统数据库业务为主，借助AI持续云服务转型，资本支出增长源于云数据中心的大规模建设以及对客户上云的支持。
- **腾讯：**投入比例波动极大，一方面由于自身云业务收入体量小，另外大起大落反映其云业务投入节奏、战略布局存在阶段性调整。

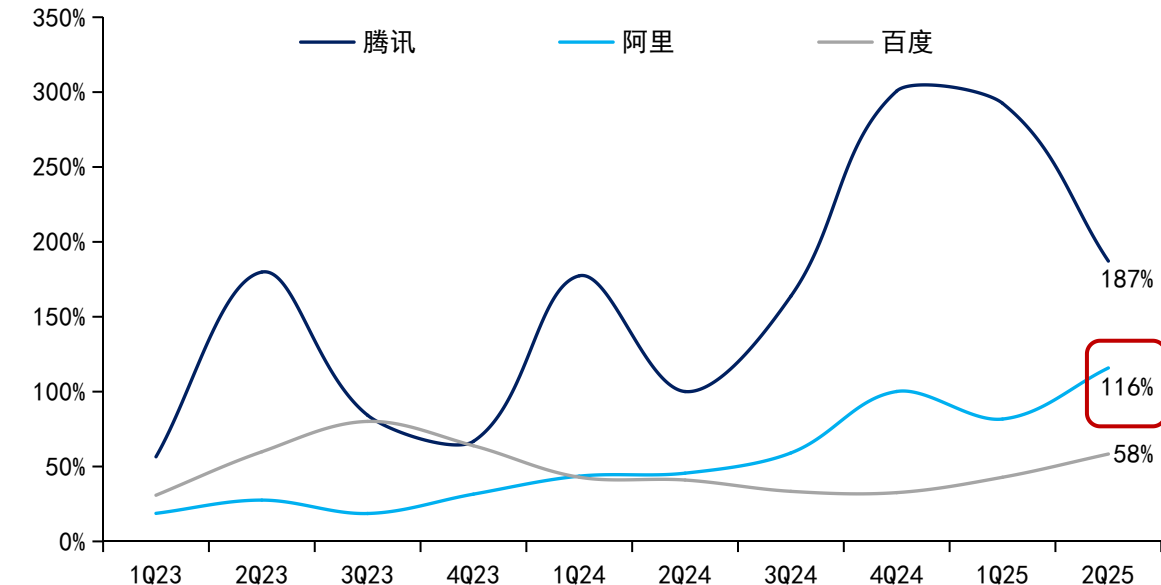
图：海外互联网Capex/云收入季度变化情况（%）



资料来源：公司财报、国信证券经济研究所整理

请务必阅读正文之后的免责声明及其项下所有内容

图：国内互联网Capex/云收入季度变化情况（%）

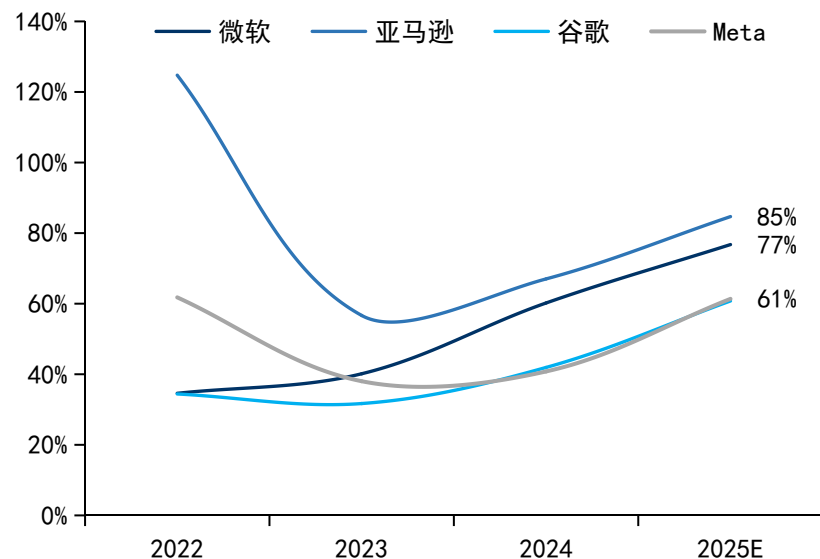


资料来源：公司财报、彭博一致性预期、国信证券经济研究所整理

海内外Capex量级：海外Capex已超过经营性现金流的60%

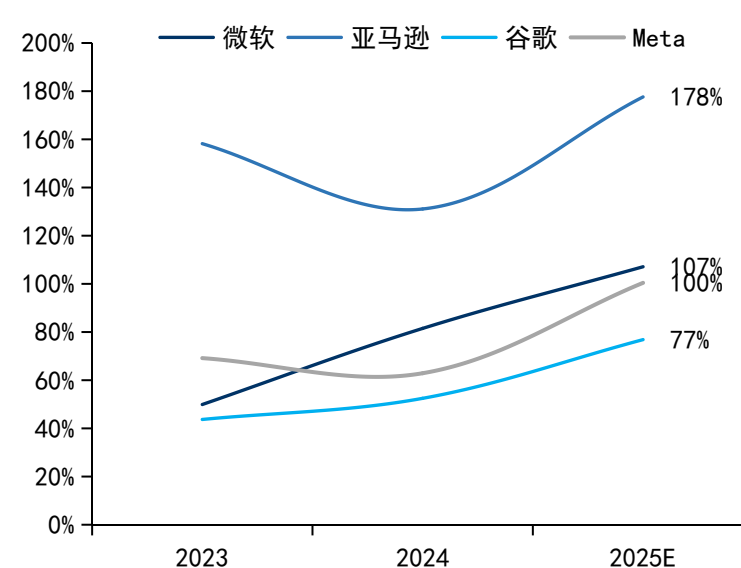
- 当前海外科技厂商已将约100%的净利润、60%以上的经营性现金流进行资本开支投入。过去三年，海外科技厂商的资本开支呈现持续提升趋势，且目前各家资本开支增速指引依然显著高于业务收入/利润增速，考虑到海外科技厂商（典型如微软、META等）通常还会有约30%净利润进行分红回购，因此未来各家现金流与利润将会对资本开支的持续投入形成一定制约。如我们测算按照META当前投资趋势，26年折旧摊销将同比增加约150亿美金（同比+80%），或导致明年净利润与今年净利润（加回159亿美元一次性非现金所得税费用）持平。
- 国内云厂中阿里闪购业务的补贴成本与云基建资本开支形成叠加压力。阿里同时推进“电商闪购业务”和“AI 云基建”的双线投入，计划未来三年投入3800亿元用于云计算和AI基础设施。相较而言，腾讯仍然采取投资与回报平衡策略，2024年“股息+回购”超1340亿元（是资本开支的近两倍）。百度由于主业利润压力投资偏谨慎。

图：海外互联网Capex/经营性现金流变化情况（%）



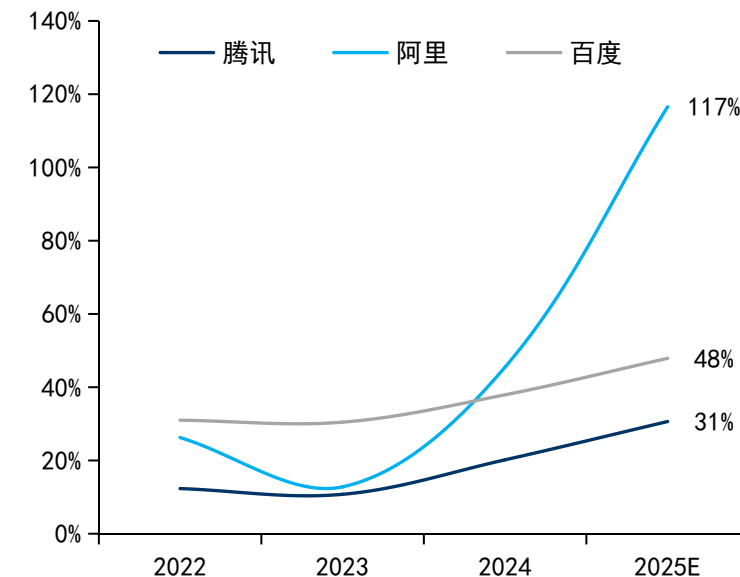
资料来源：公司财报、彭博一致性预期、国信证券经济研究所整理

图：海外互联网Capex/净利润变化情况（%）



资料来源：公司财报、彭博一致性预期、国信证券经济研究所整理

图：国内互联网Capex/经营性现金流变化情况（%）

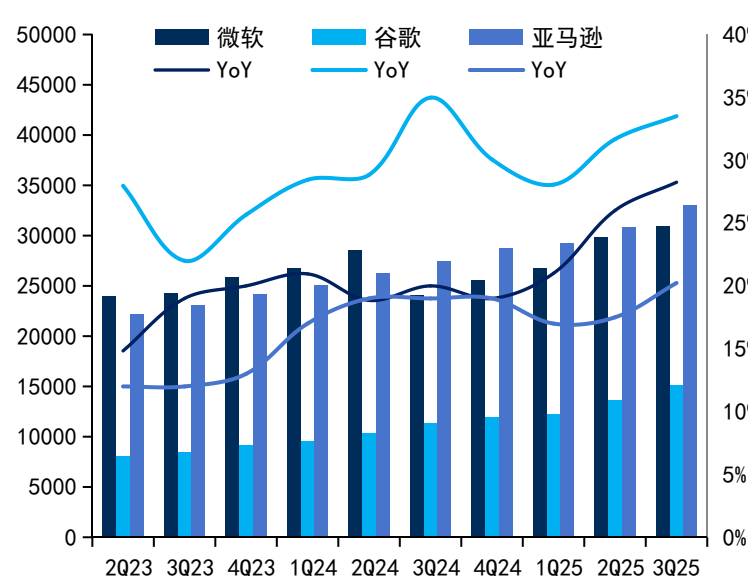


资料来源：公司财报、彭博一致性预期、国信证券经济研究所整理



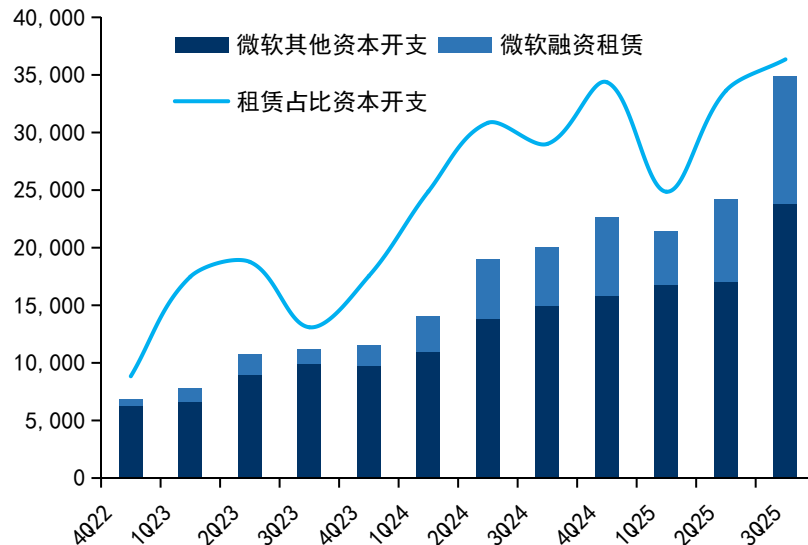
- 在经营性现金流（净利润）约束下，我们同时看到为了维持高资本开支增长需求下的AI基础设施金融化操作以及加大融资租赁举动：
- 2025年10月META通过与Blue Owl合资体发行270亿美元债务：用于Hyperion超级数据中心建设，通过特殊目的实体（SPV）架构实现表外运作——Meta仅持有项目20%股权，**债务不计入自身资产负债表**。
- 微软持续大幅提升融资租赁占比：微软25Q3融资租赁达111亿美元，占同期资本开支（349亿美元）的31.8%，指引后续持续提升。**优点**：融资租赁基本为长期资产投资，资产生命周期长达15年以上，从净利润角度较长的折旧年限减少了当期折旧金额。另外，通过分期支付租金的方式减轻短期现金流压力，租金中本金部分计入“投资活动现金流”、利息部分计入“经营活动现金流”。**缺点**：管理层多次提到“融资租赁交付时间存在不确定性”。

图：海外云厂商云收入变化情况（百万美元/%）



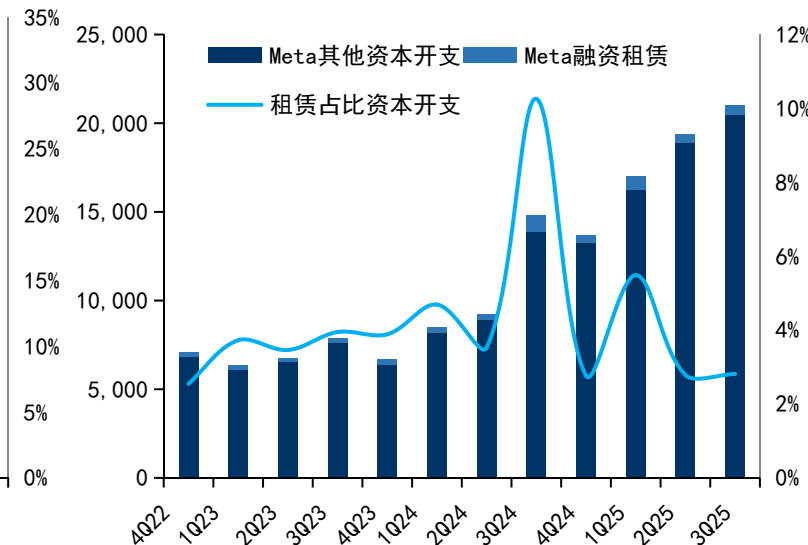
资料来源：公司财报、国信证券经济研究所整理

图：微软融资租赁占比资本开支变化（%）



资料来源：公司财报、彭博一致性预期、国信证券经济研究所整理

图：META融资租赁占比资本开支变化（%）

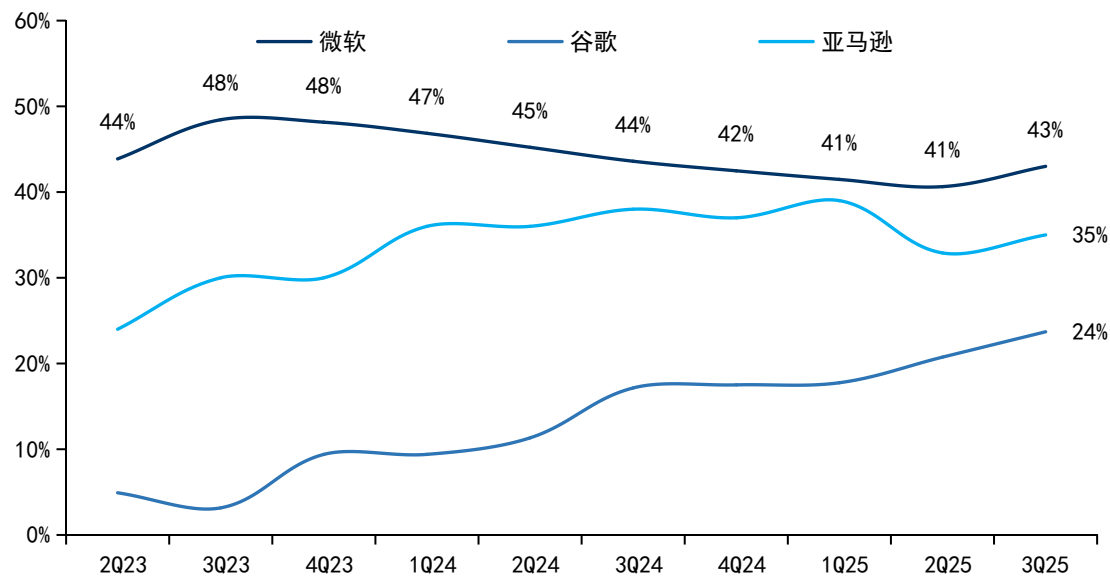


资料来源：公司财报、彭博一致性预期、国信证券经济研究所整理

海外云厂利润率：规模效应与折摊均提升，微软云OPM率先企稳

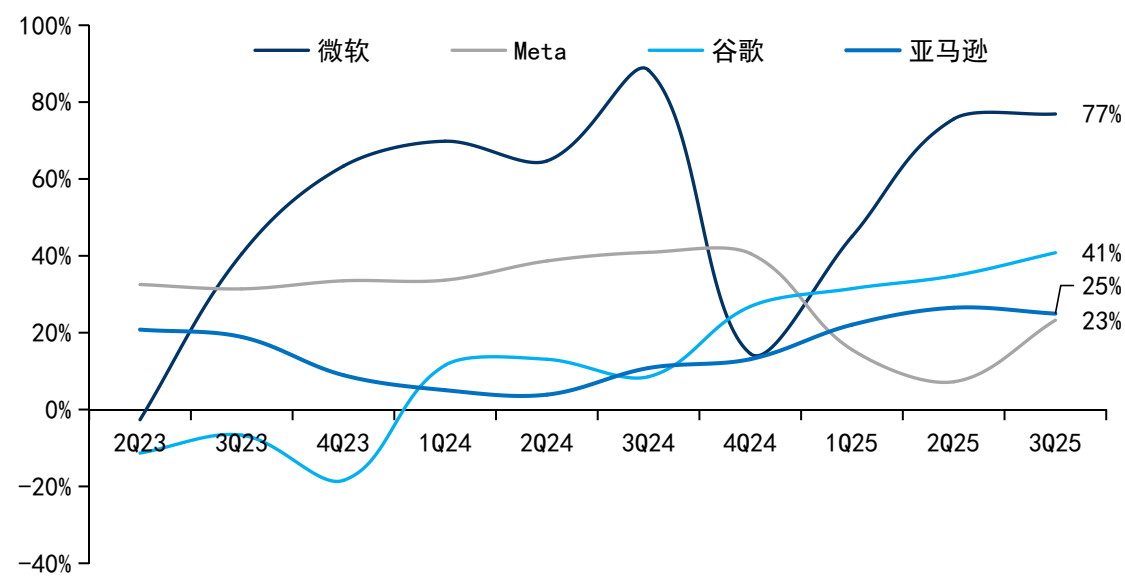
- 微软云：FY26Q1（25Q3）智能云营业利润率43%（同比-1pcts、环比+2pcts）。对比其他海外云厂、微软Capex加速最早最快，也率先迎来收获期。云具备典型的资本前置与规模效应，尽管仍有算力容量限制、但需求强预计可维持高增速，后续利润杠杆有望进一步释放。FY26Q2智能云收入指引同比+25%-26%，其中Azure指引+37%cc、维持高增速。
- AWS亚马逊云：25Q3 OPM 35%（同比-3pcts、环比+2pcts）。波动的核心短期因素是AI相关资产投入带来的折旧影响。自研AI芯片 Trainium2已为几十亿美元业务，25Q3收入环比+150%，且相较于其他GPU具备30%-40%的价格性能优势。
- 谷歌云：25Q3 OPM提升至24%（同比+7pcts、环比+3pcts）。受益于强劲的收入增长和持续的成本效率提升，核心产品和人工智能产品方面的增速远高于云业务的整体收入增长。未完成订单达155亿美元（同比+ 82%），新客户数量同比+34%。

图：各互联网云业务OPM变化情况（%）



资料来源：公司财报、国信证券经济研究所整理

图：各互联网与云厂折旧摊销同比变化情况（%）



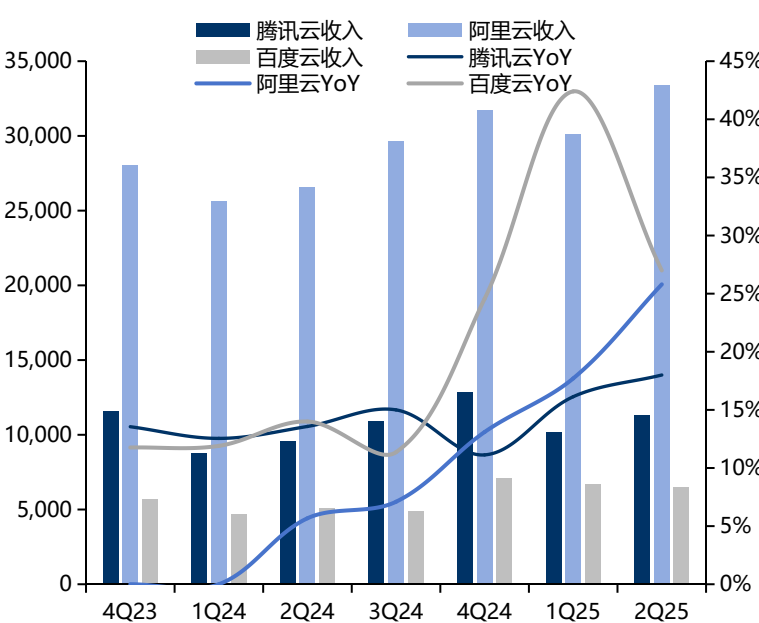
资料来源：公司财报、彭博一致性预期、国信证券经济研究所整理

国内云厂利润率：近期利润率持平，长期AI业务毛利率更高



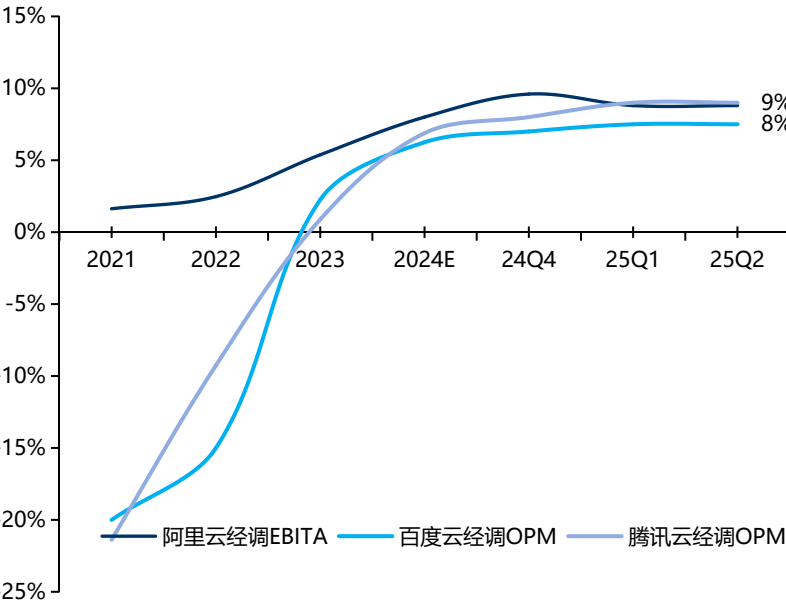
- 近几年互联网云厂利润率持续提升，经调整OPM/EBITA口径约9%。互联网云厂商都在逐渐从“创收”向“盈利”方向上转变，主动放弃非良性项目，提升自身被集成比例，舍弃低利润业务。阿里云近4年来营业利润率已经改善了近18个百分点。24Q3，阿里财报披露阿里云经调整EBITA为9%。
- AI GPU租赁与API调用业务平均毛利率高于传统IaaS，但短期AI相关折旧与研发运营成本提升可能带来一定压力。尽管大模型定价/毛利率水平不高，但整体上调用量会大幅提升，云厂商通过配套的高利润PaaS产品组合将整体合同利润提升，比如：知识数据库插件、联网搜索插件等。

图：国内云厂商季度收入与同比变化预测（百万元/%）



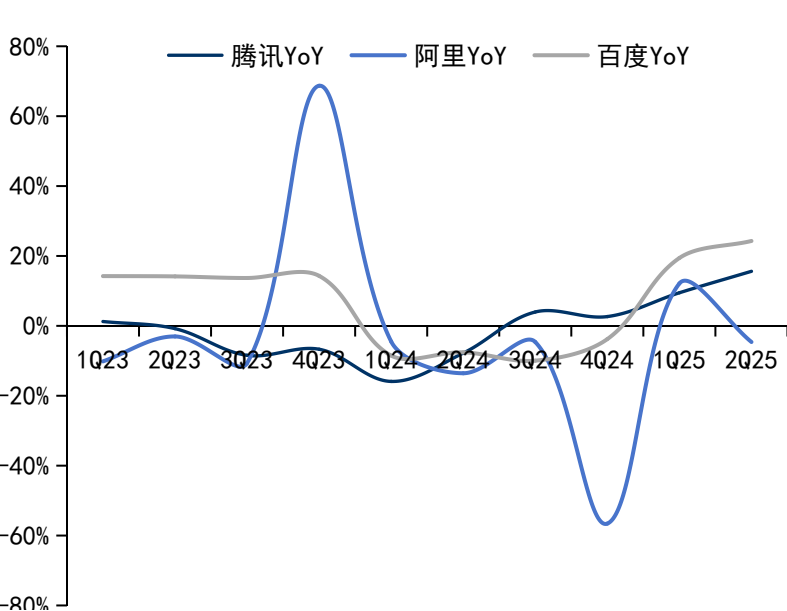
资料来源：公司财报、彭博一致性预期、国信证券经济研究所整理测算

图：国内云厂商经调EBITA与OPM变化(%)



资料来源：公司财报、彭博一致性预期、国信证券经济研究所整理测算

图：国内互联网云厂折旧摊销增速变化(百万元/%)

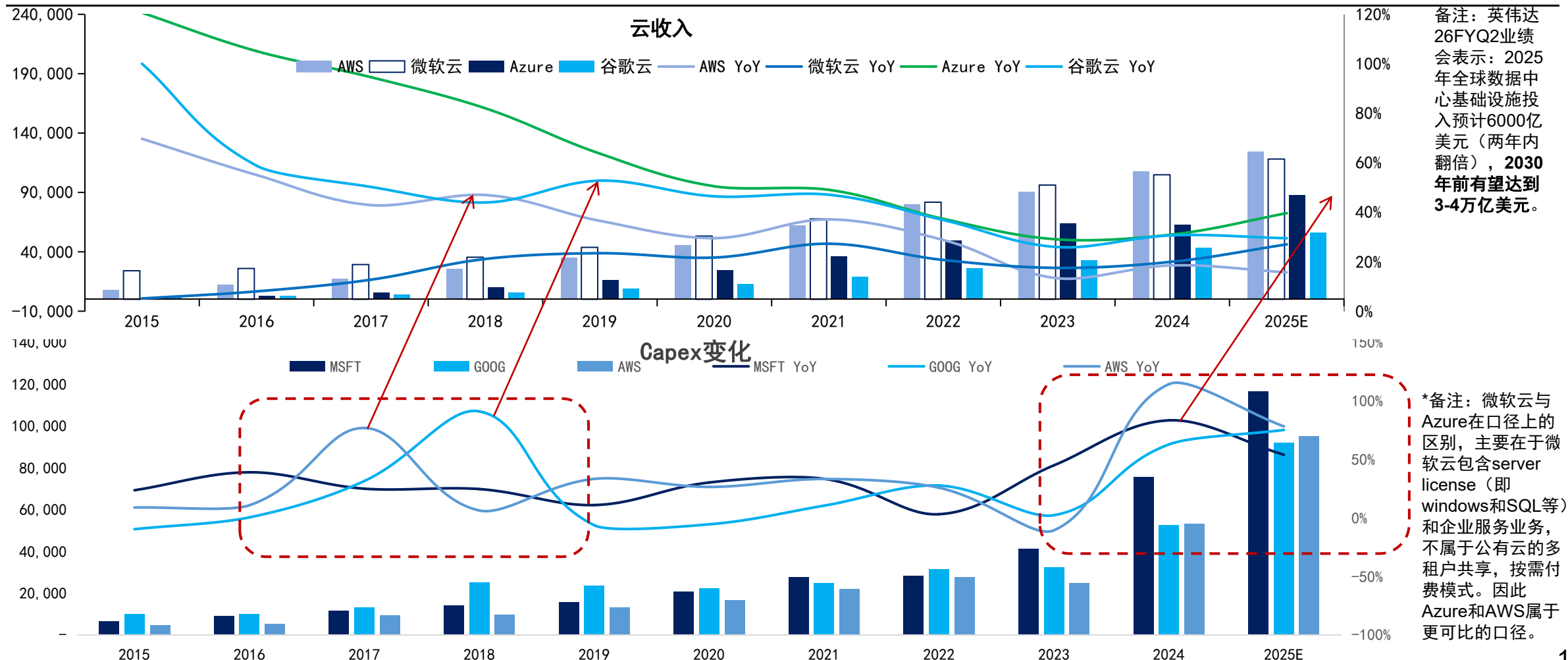


资料来源：公司财报、彭博一致性预期、国信证券经济研究所整理预测

海外云厂商资本开支与收入变化复盘

- 复盘过去10年的云计算市场，传统云扩张投资约持续四年、且投入到收入兑现约一年时间差。但此次AI投资杠杆更高、供需错配度更高、建设周期更长、芯片迭代速度更快、商业模式亦有所变化，导致收入兑现节奏更慢（如微软云投资增速顶点到Azure最高增速至少1.5年），中性预期下此次AI投资可能会持续五到六年扩张周期。

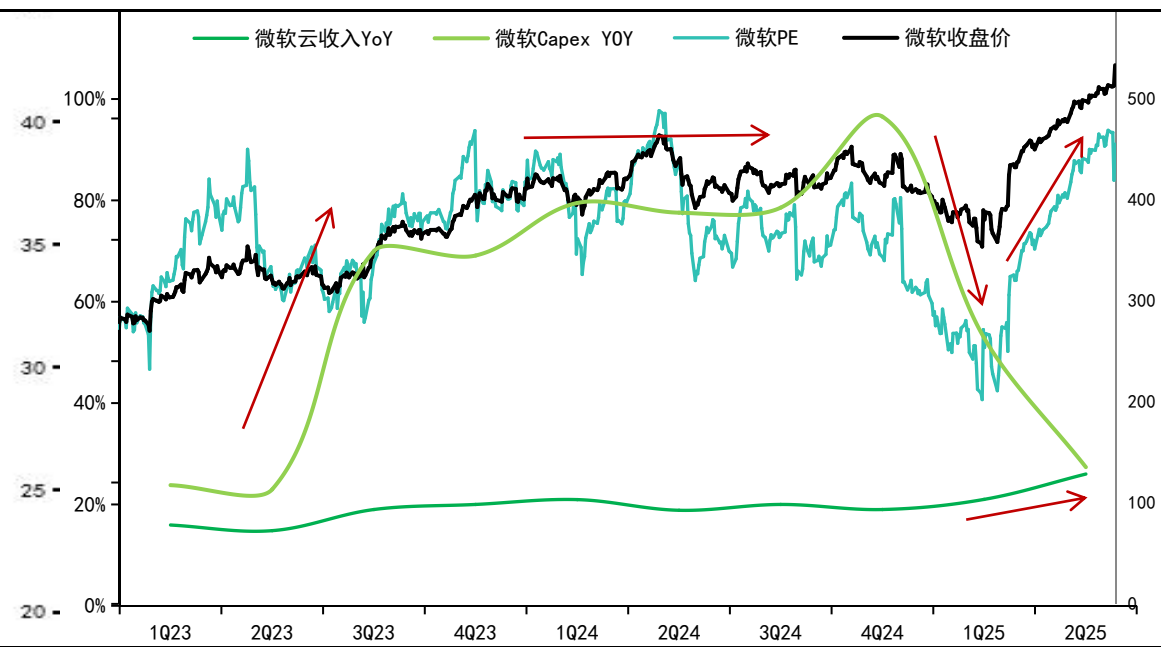
图：海外云厂商收入与资本开支变化(%)



微软云与阿里云：估值与资本开支、云收入增速关系

- 复盘微软云过去两年发展过程，估值变化与资本开支和云收入显著相关。
- ① 一阶段：估值与资本开支增速正相关。前期高投入阶段，资本开支加速推动估值（PE）上升，加大云基础设施、配合云收入增速的稳定性，市场会强化“投入→未来增长加速”的预期，估值上升、进而放大股价涨幅。后期投入放缓阶段，资本开支增速下降带来估值调整，短期可能引发需求不足的担忧，导致估值阶段性调整、股价波动。
- ② 二阶段：资本开支兑现成业绩，云收入加速成为估值核心驱动，将进一步抬升估值。证明云业务已进入“投入效率释放期”（收入增速跑赢投入端、回报率提升、规模效应显现），为估值与股价的后续回升奠定基础。
- 当前我们认为国产云厂商以阿里云亦处于资本开支积极以及云业务加速推动估值继续上升阶段。

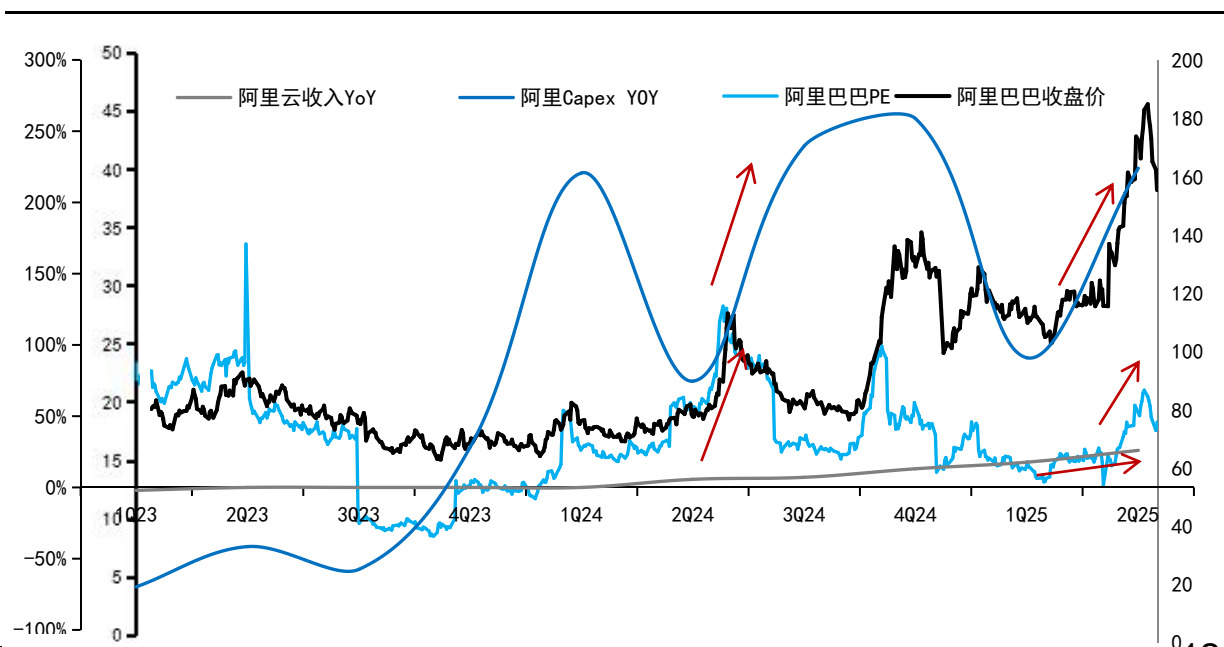
图：微软估值、估值、云增速与资本开支增速变化(%)



资料来源：公司财报、彭博一致性预期、国信证券经济研究所整理测算

请务必阅读正文之后的免责声明及其项下所有内容

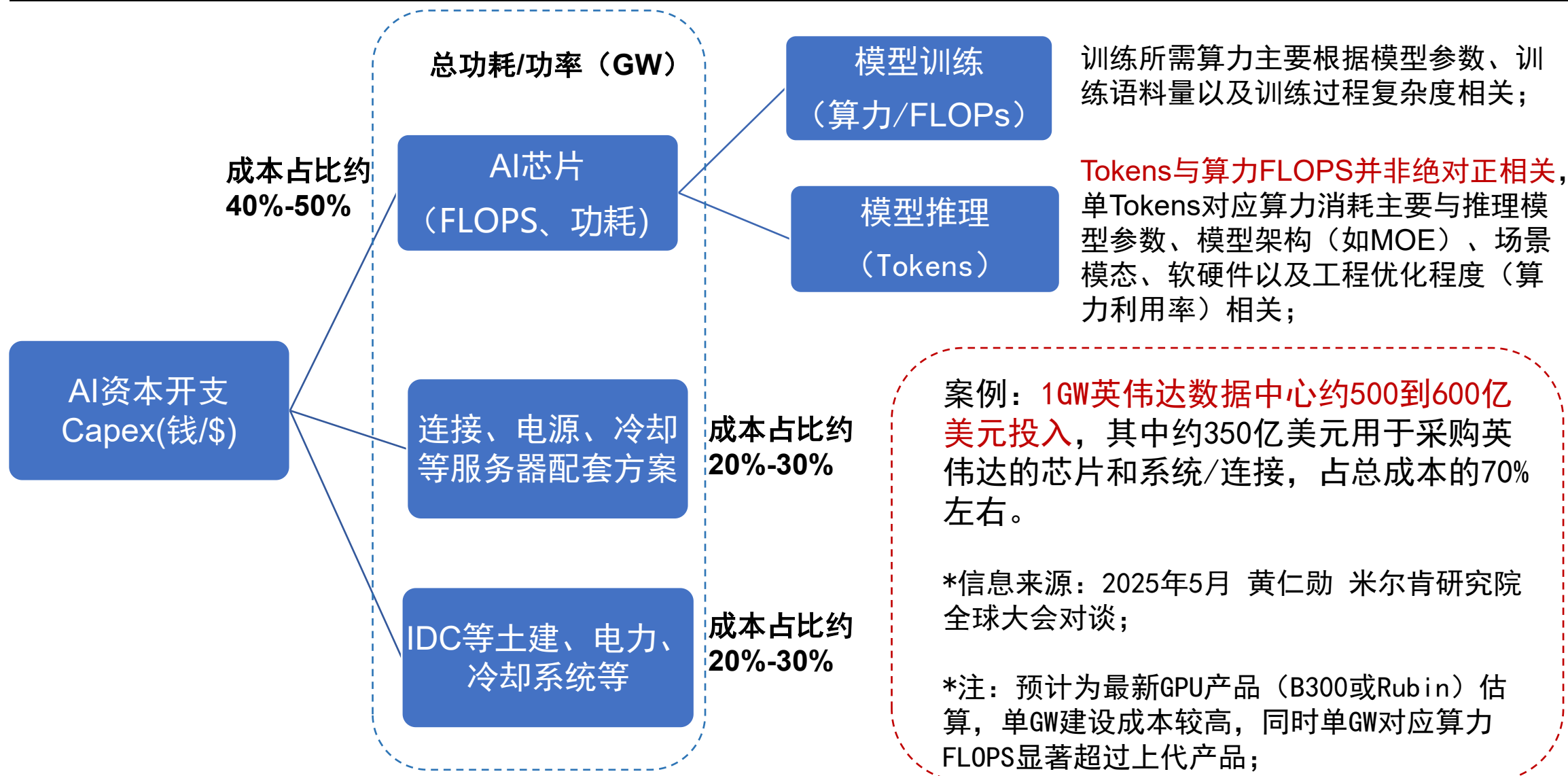
图：阿里估值、估值、云增速与资本开支增速变化(%)



资料来源：公司财报、wind、国信证券经济研究所整理测算

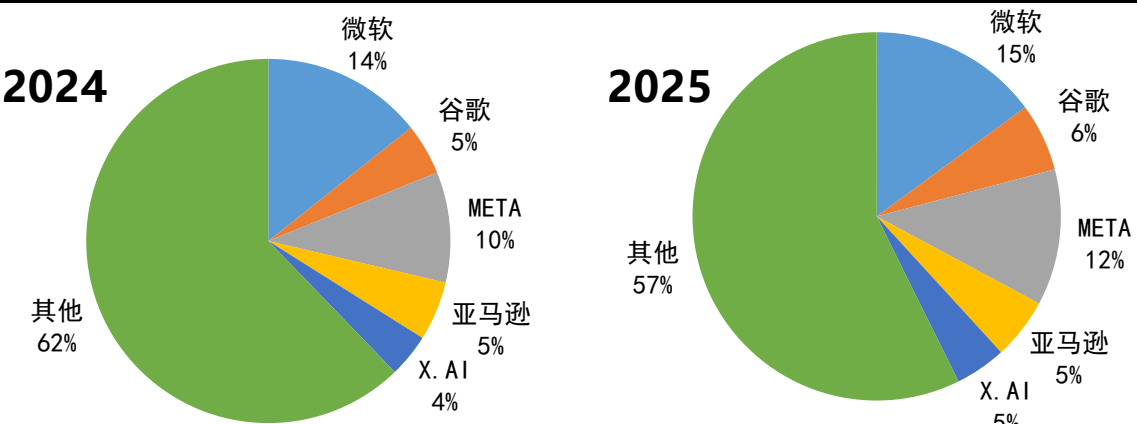
- 【 01 】 海内外云厂商资本开支复盘分析
- 【 02 】 云厂商算力规模与建设计划
- 【 03 】 云厂自研芯片布局与进展
- 【 04 】 云厂AI云收入、AI ROI测算与估值

图：资本开支投资分配



- 英伟达26FYQ2业绩会表示：2025年全球数据中心基础设施投入预计6000亿美元（两年内翻倍），2030年前有望达到3-4万亿美元。增长驱动来自于推理型AI（agenticAI）需求激增、全球主权AI与企业AI建设、物理AI（机器人/自动驾驶）兴起。
- 云服务提供商（CSP）约50%：25FYQ4（截至2025年1月）英伟达业绩会表示云服务商约贡献了50%的数据中心收入，包括 AWS、Azure、Google Cloud、CoreWeave、Oracle等大型云服务商。其需求主要来自于AI 基础设施建设，包括大模型训练、推理等需求。
- 消费互联网公司约20%：（如 Meta、字节跳动、其他社交媒体和互联网平台）约贡献了数据中心收入的20%-25%，且增速显著。25FYQ4英伟达业绩会表示消费互联网领域营收同比增长3倍，包括推荐系统、视觉语言理解、合成数据生成搜索、多模态内容处理和代理AI等。
- 政府与主权AI约10%：潜力巨大的新市场，2025年预计入占比数据中心收入的10%。英伟达26FYQ2业绩会（截至2025年7月）表示主权AI从2024年初的零收入到2025年主权AI收入目标200亿美元（较2024年翻倍）。全球多国正积极建设本土AI能力，为保障数据安全、技术自主性及 AI产业主导权。
- 企业客户与AI初创公司约20%：企业客户（如SAP、ServiceNow）与初创AI公司贡献了数据中心的剩余主要收入。企业端主要用于模型微调和代理 AI工作流，以适配垂直领域需求（如医疗影像诊断、工业质检）。优化数据库查询、数据分析等任务，提升效率。初创AI公司后续预计成为快速增长点，NVIDIA 计划投资OpenAI 高达1000亿美元换取OpenAI将部署至少10 吉瓦的 NVIDIA 系统。

图：2024与2025年海外科技厂贡献英伟达数据中心芯片收入占比测算



资料来源：Omdia、公司官网、彭博一致性预期，国信证券经济研究所整理测算

表：2025年底海外科技厂商算力保有量与数据中心功耗测算

	等效H100 (千片)	总价 (百万美元)	总浮点计算量 (Exaflops/FP16)	对应AIDC功耗 (GW)	说明
微软	3,244	54,325	6,419	3.5	除谷歌TPU外 其他厂商算力 未计算除NV以 外芯片。 TPU系列芯片 功耗水平高于 NV系列。
谷歌	2,756	43,875	5,454	5.3	
META	2,457	39,300	4,854	2.7	
亚马逊	1,137	18,030	2,198	1.2	
X.AI	897	14,600	1,776	1.0	

资料来源：TNP、Omdia、公司官网，国信证券经济研究所整理测算

海内外科技厂商关于资本开支与算力容量的相关表述

表：海外科技厂商关于资本开支与算力容量的相关表述

	资本开支表述	AI 算力表述	数据中心算力表述	数据中心功耗表述	电力供应表述	测算2024/2025 AI 在手容量*(GW)
微软	- FY26（25Q3-26Q2）的资本支出增速将高于FY25（yoy+58%）。 - 微软每年在AI基建上的固定投资超过800亿美元，这是十年规划。	公司计划FY26（25Q3-26Q2）将AI总容量提升超80%。	25Q3业绩会：未来两年内将数据中心总规模扩大约一倍	25Q2业绩会：数据中心超400个，过去一年新增超2GW； - 微软CEO 25年10月发言：微软云全球数据中心总功耗5GW；		1.2/3.5
谷歌	2025年Capex 预期上调至910-930 亿美元（原预期850亿美元），2026年计划进一步大幅增加。					2/5.3
META	上调2025年资本开支至700-720亿美元（前值660-720亿美元），同比+78%-84% 24年底计拥有等效60万 预计2026年的资本支出增长将显著高于2025年。	H100 GPU的计算能力；				0.9/2.7
亚马逊	25年全年CAPEX上修至1250亿美元（上季度后预计为1100-1200亿美元），且 预计26年将进一步增加。		25Q3业绩会：AWS 目标2027年底实现 算力翻倍		25Q3业绩会：过去12 个月AWS新增3.8吉瓦 电力，第四季度预计 再增1吉瓦以上。	0.4/1.2
阿里巴巴	2025年2月宣布，未来三年将投入至少 3800亿元人民币用于AI和云计算基础设施 建设			云栖大会宣布：到2032年， 阿里云全球数据中心的能 耗规模将比2022年（GenAI 元年）提升10倍。		
OpenAI	- 星际之门计划投资超5000亿美元； - OpenAI承诺在未来多年内向微软 Azure云服务采购约2500亿美元的 服务；	英伟达未来供应10GW数 据中心的芯片； AMD未来供应6GW芯片； 博通未来供应10GW芯片；		2025：未来四年内星际之 门Stargate计划容量达到 近10GW；		

资料来源：各公司官网、各公司业绩会、华尔街新闻，国信证券经济研究所整理 *在手容量仅考虑英伟达与TPU芯片算力容量，参考后续ppt具体单页公司测算

请务必阅读正文之后的免责声明及其项下所有内容

微软算力规模测算：过去两年新增超2GW，未来两年数据中心总规模翻倍



- FY26Q1业绩会指引FY26的资本支出增速将高于FY25（Capex 882亿美元、yoy+58%），FY26Q2资本开支将环比增加。
- 公司计划本年度（25Q3-26Q2）将AI总容量提升超 80%，未来两年内将数据中心总规模扩大约一倍。预计明年中旬AI数据中心容量将扩大至6.3GW，未来两年数据中心总容量将达10GW。已部署全球首个 NVIDIA GB300大规模集群，威斯康星州 Fairwater AI 数据中心（明年上线、规模2吉瓦）将成全球最强AI数据中心。

表：微软2026年上线的北美数据中心建设计划

数据中心名称	地址	类型	投资金额 (美元)	GPU 数量	功率总量
Fairwater AI 数据中心	美国威斯康星州	建设	73 亿美元	数十万块 GB200	2GW
威斯康星州第二数据中心	美国威斯康星州	建设	40 亿美元	未公布	未公布
Nebius新泽西数据中心	美国新泽西州文兰	租赁	174-194亿美元（五年期租赁总金额）	100,000 块GB300	未公布
Nscale 德州数据中心	美国德克萨斯州	租赁	240 亿美元 (4 个数据中心总计，另外三个为英国、葡萄牙、挪威)	104,000 块GB300	初240MW，可扩展至 1.2GW

资料来源：公司官网、国信证券经济研究所整理

图：微软拥有的算力规模

年份	型号	GPU (千片)	F16理论峰值算力 (TFLOPS)	等效H100 (千片)	单价 (美元)	总价 (百万美元)	占比当年资本开支	总浮点计算量 (Exaflops/FP16)	单芯片功耗 (W)	芯片总功耗 (MW)	对应AIDC功耗 (MW)	数据中心总容量 (GW)	备注
2022	H系列	75	1,979	75	50,000	3,750		148					假设芯片占服务器功率80%，假设服务器占AIDC功率55%； 算力保有量相比当年资本开支存在一定时间差； 25Q2：数据中心超400个，过去一年新增超2GW 根据芯片官方价格假设单价，后续测算假设一致； 全球数据中心总功耗5GW
2023	H系列	150	1,979	150	25,000	3,750	9%	297	700	105	239		
2024	H系列	485	1,979	485	25,000	12,125	16%	960	700	340	772		
2024	B200	90	4,600	209	30,000	2,700	4%	414	1,000	90	205		
	2024之前总计			920		22,325		1,819			1,215		
2025	B200/300	1,000	4,600	2,324	32,000	32,000	31%	4,600	1,000	1,000	2,273		
	总计			3,244		54,325		6,419		1,535	3,488		

资料来源：TNP、Omdia、公司官网、微软CEO西雅图创新峰会发言，国信证券经济研究所整理测算

请务必阅读正文之后的免责声明及其项下所有内容

微软AI资本开支：十年规划、每年在AI基建投资超800亿美元

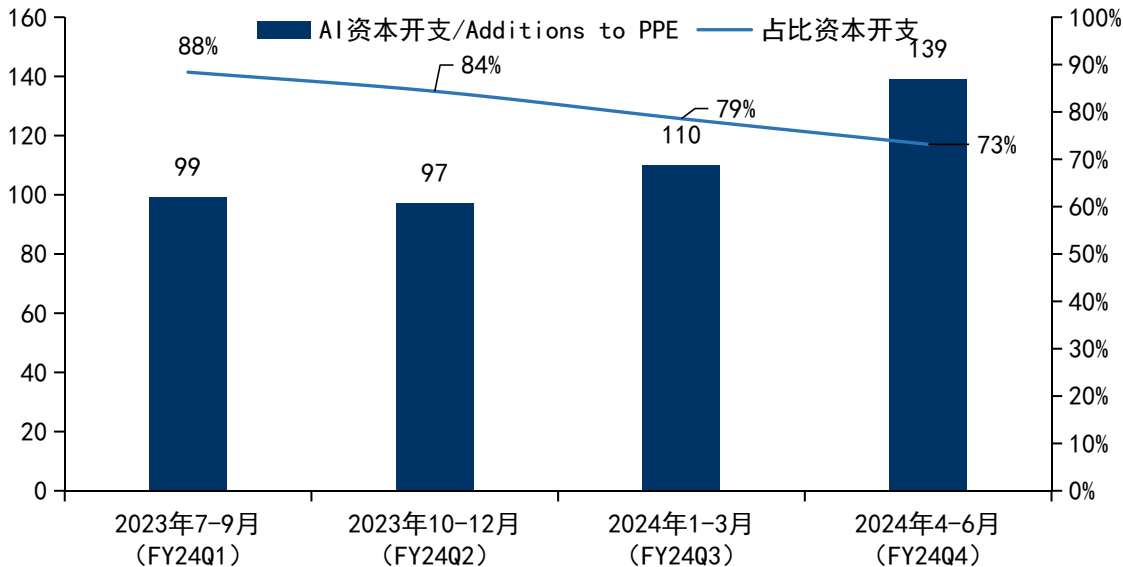
- 25年初微软总裁BradSmith在官方博客中宣布，微软计划在2025财年（截至2025年6月30）投入约800亿美元用于AI 数据中心建设（2025财年实际资本支出达882亿美元），其中超过一半将用于美国本土的数据中心建设。预计AI 资本开支占比25财年总资本开支的90%。
- “微软每年在AI 基建上的固定投资超过800亿美元，这是“十年规划的系统性布局”而非临时投入”**。微软CEO萨蒂亚·纳德拉在2025年10月9日西雅图创新峰会的发言表示，微软全球数据中心总功耗已达5GW。AI 基础设施不是短跑而是马拉松。当别人还在画图纸时，微软已经让算力在全球流动了。

表：微软全球数据中心情况

设施指标	数据
全球数据中心总功耗	5GW（相当于5座大型核电站）
覆盖范围	140 个国家、300 多个区域节点
技术特点	全栈解决方案（芯片适配+液冷技术）
投资回报	算力成本比行业平均水平低 30%

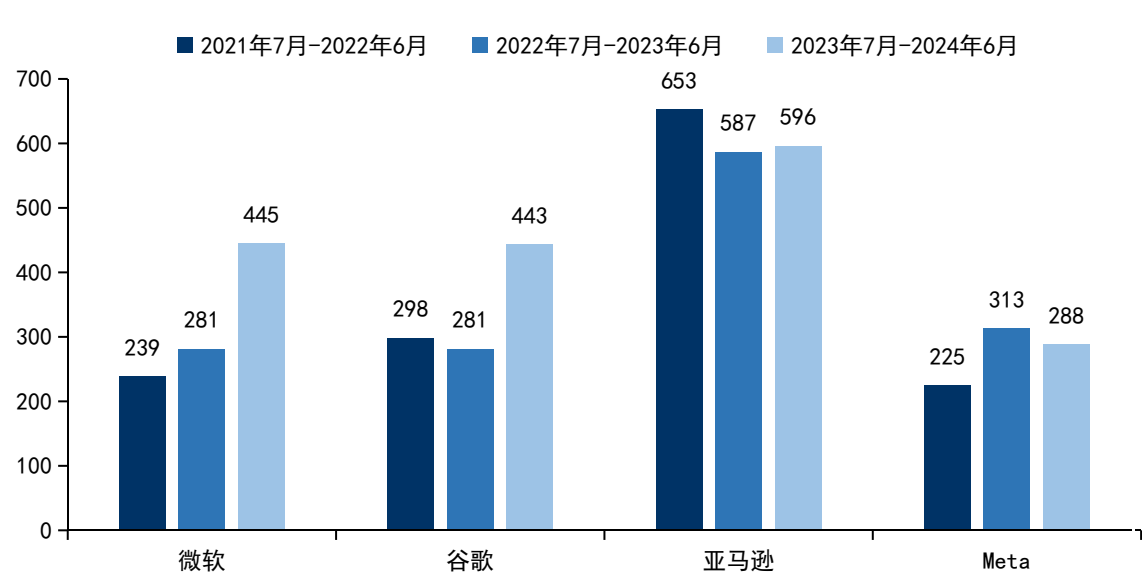
资料来源：微软CEO西雅图创新峰会发言、国信证券经济研究所整理

图：微软AI基建/Additions to PPE变化（亿美元/%）



资料来源：SEC官网，统计口径：Additions to property and equipment），国信证券经济研究所整理
请务必阅读正文之后的免责声明及其项下所有内容

图：海外科技厂商AI基建/Additions to PPE变化（亿美元/%）

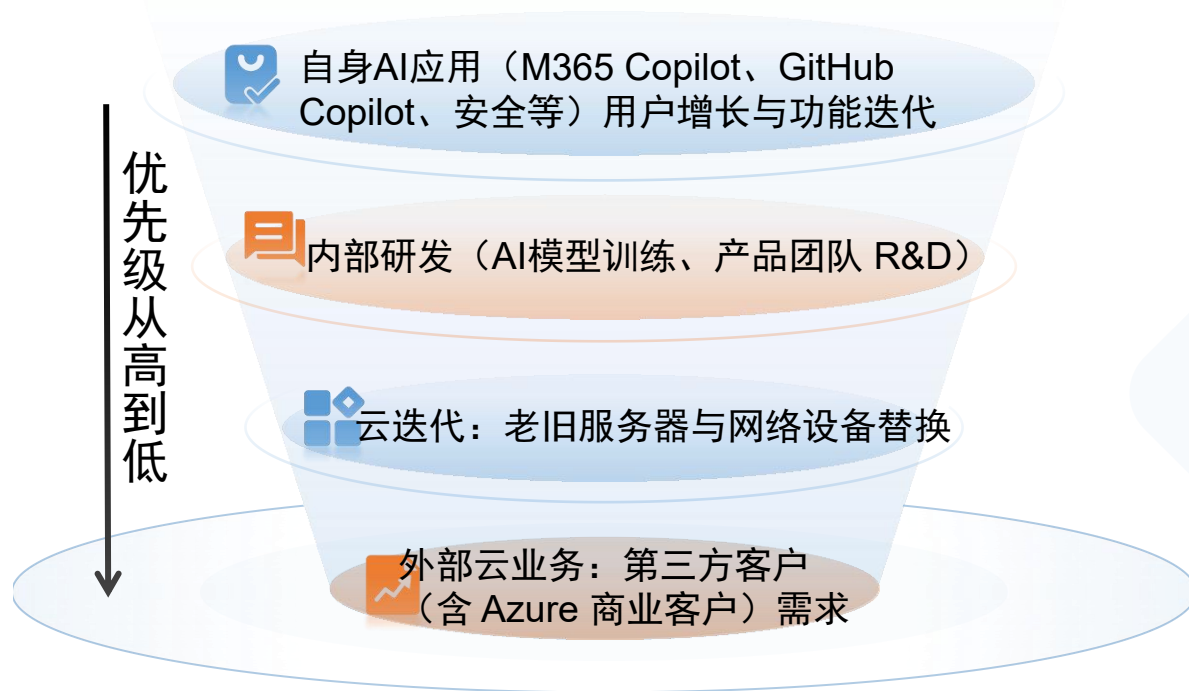


资料来源：SEC官网，统计口径：Additions to property and equipment），国信证券经济研究所整理

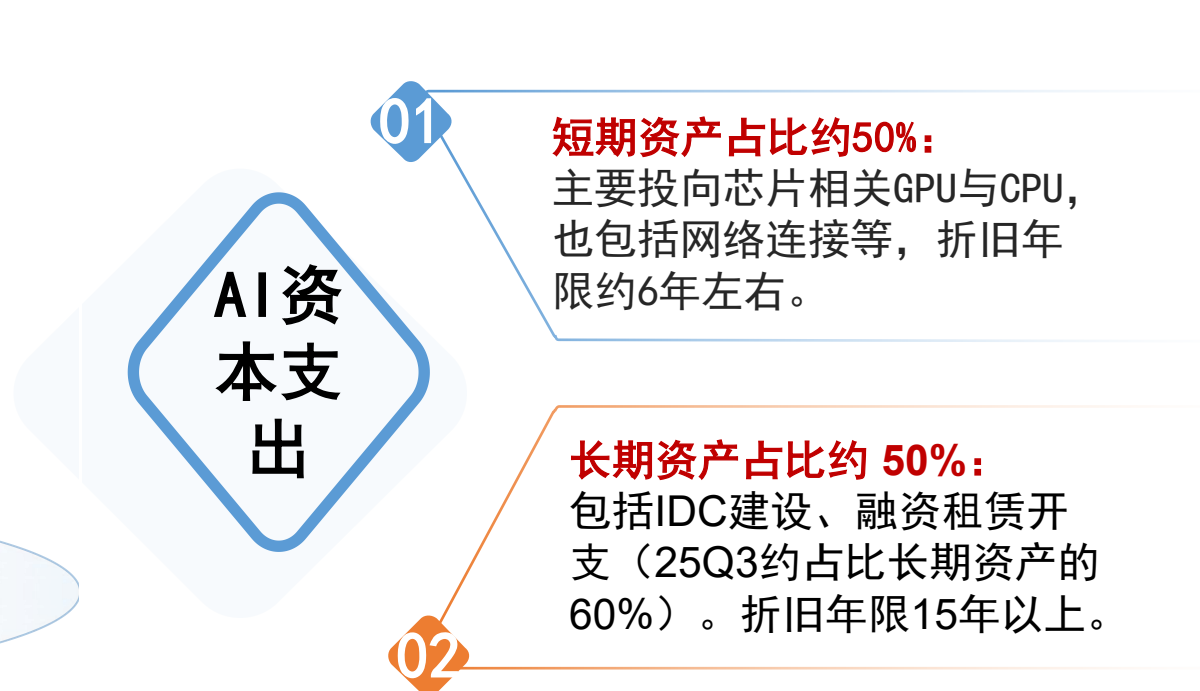
微软资本开支分配：短期资产（芯片）与长期资产（IDC与租赁）各一半

- FY26Q1业绩会表示AI算力将优先供给内部AI应用，其中短期资产占比AI资本开支的一半。Azure算力需求再次超过现有产能，即便加速扩容，仍需按优先级分配算力，优先内部AI Copilot等应用增长与内部研发。
- 2025年10月28日，OpenAI承诺在未来多年内向微软Azure云服务采购约2500亿美元的服务。
- **OpenAI获得权益：**微软不再拥有作为OpenAI计算服务提供商的"优先提供权"，OpenAI获准与第三方开展部分产品合作（API类产品仍由Azure独家承载），架构上完成向"公共利益公司"（PBC）的转型，为未来IPO铺平道路；
- **微软获得权益：**新增2500亿美元Azure服务签约，微软的收入分成、OpenAI API 独家授权（Azure 专属）将延续至2030年（每年能数十亿美元利润），模型与产品IP延长至2032年。微软CEO表示微软对OpenAI投资已获约10倍回报。

图：微软AI算力分配优先级



图：微软AI资本开支分配



谷歌资本投入方向、算力规模测算

- 投入规模：2025年全年Capex预计超900亿美元，本轮大模型发展以来，**2023到2025三年累计投入超1700亿美元，展望2026年继续大幅增加。**
- 结构方面：24Q3财报会表述，AI支出为130亿美元，其中主要用于搭建技术基础设施，其中**60%是服务器（GPU/TPU服务器），40%是数据中心建设（厂房建安、冷却、电气设备配套等）和网络连接设备支出。**而TPU和GPU内部结构方面，预计此前TPU和GPU投入金额、算力规模整体量级接近，后续随着TPU技术、性能和生态成熟度持续提升，**预计2026年TPU投入和算力规模将超越GPU。**

表：谷歌24Q3 Capex构成测算

mn USD	金额	占比	说明/假设
Total Capex 24Q3	13,000	100%	
1.Technical infrastructure	10,000	77%	majority, CAPEX大头, 假设约80%总支出
1.1 Datacenter&networking	4,000	31%	40% of TI, 数据中心建设、连接
1.2 Servers	6,000	46%	60% of TI, 服务器
1.2.1 AI Server	5,500	42%	AI服务器价值, both TPUs and GPUs
i. TPU Server	3,000	23%	
TPU v4/v5e/v5p	2,100	16%	假设TPU占70%, 博通单季ASIC收入20亿美元
ii. GPU Server	2,500	19%	
H100/H200	1,800	14%	假设GPU占服务器价值70%
1.2.2 Non AI Server	500	4%	非AI服务器投入
2.Facilities, the bets&other	3,000	23%	minority, 剩余CAPEX

资料来源：公司财报会，国信证券经济研究所整理测算

图：谷歌拥有的算力规模

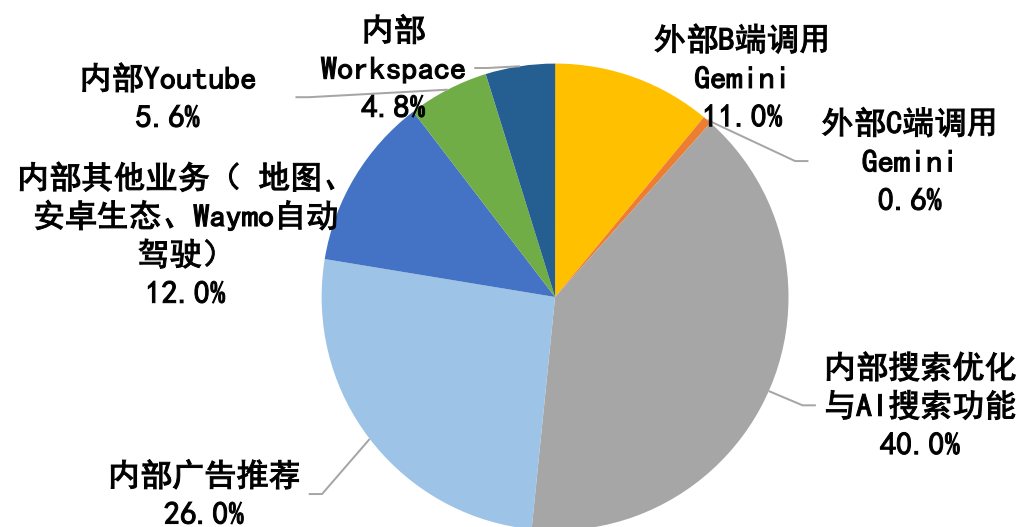
年份	型号	GPU (千张)	F16理论峰值算力 (TFLOPS)	等效H100 (千片)	单价 (美元)	总价 (百万美元)	占比当年资本开支	总浮点计算量 (Exaflops/FP16)	单芯片功耗 (W)	芯片总功耗 (MW)	对应AIDC功耗 (MW)	备注
2023	H系列	50	1979	50	25,000	1250	4%	99	700	35	80	假设芯片占服务器功率80%, 假设服务器占AIDC功率55%; 算力保有量相比当年资本开支存在一定时间差; 根据芯片官方价格假设单价。
2023	TPU系列	1,000	300	152	4000	4,000	12%	300	300	300	682	
2024	H系列	169	1979	169	25,000	4,225	8%	334	700	118	269	
2024	TPU系列	1500	300	227	4000	6,000	11%	450	300	450	1,023	
2024	B200	10	4,600	23	30000	300	1%	46	1,000	10	23	
	2024之前总计	662		621		15775		1229			2,076	
2025	B200/300	400	4,600	930	32000	12,800	16%	1,840	1,000	400	909	
	TPU V5	1200	300	182	4000	4,800	6%	360	300	360	818	
	TPU V6	1500	1350	1,023	7000	10,500	13%	2,025	450	675	1,534	
	总计			2,756		43,875		5,454		2,348	5,337	

资料来源：TNP、Omdia、公司官网，国信证券经济研究所整理测算

- 算力使用方向：用于谷歌内外部所有场景和客户，其中**TPU自用更多**，**GPU租外部更多**，预计采购TPU约70%供谷歌自用、30%租赁给第三方，而采购GPU则约30%自用、70%用于租赁，具体看：

- ① 谷歌内部模型训练：开发Gemini大模型新版本，目前Gemini使用TPU训练，最新模型性能居于第一梯队
- ② 谷歌内部应用推理：AI赋能谷歌原有产品能力，包括部署搜索算法优化与AI搜索功能、广告推荐系统改造、Youtube、Workspace、地图、安卓生态等，提升用户体验、交互效率，同时2025Q3 Gemini独立应用产品也拥有MAU 6.5亿，查询量环比增长2倍，视频Veo3增长迅速（生成超2.3亿个视频），今年以来tokens消耗增长迅速，
- ③ 谷歌云GCP：当前云业务高增长，盈利持续改善，后续或支持Anthropic等大模型客户TPU大规模推训；

图：谷歌月度tokens调用量场景占比



资料来源：国信证券经济研究所整理测算

图：谷歌月度tokens增长



资料来源：公司公告，国信证券经济研究所整理测算

亚马逊算力规模测算：目标2027年实现算力翻倍



- 本轮大模型发展以来，2023到2025三年累计AWS投入超1700亿美元。25Q3业绩会表示过去12个月AWS新增3.8吉瓦电力（供电能力并非AI数据中心算力功耗），第四季度预计再增1吉瓦以上，且新增算力正被快速投入产生收益，未出现明显的产能闲置；同时，AWS目标2027年底实现算力翻倍，持续扩大的算力规模将摊薄固定成本，提升运营效率。
- 投入规模：（根据财报披露各业务PPE结构测算）
 - 2023年：总Capex 481亿美元，AWS Capex预计占50%以上，约250亿美元（剩余部分主要投向电商物流）；
 - 2024年：总Capex 777亿美元，AWS Capex预计占60%以上，约500亿美元；
 - 2025年：前三季度Capex 899亿美元，全年Capex预计约1250亿美元，AWS Capex预计1000亿美元左右；

图：亚马逊主要业务板块固定资产新增结构

业务板块（百万美元）	2019	2020	2021	2022	2023	2024
北美电商物流	11,752	29,889	37,397	23,682	17,529	24,348
国际电商物流	3,298	8,072	10,259	6,711	4,144	6,643
AWS	13,058	16,530	22,047	27,755	24,843	53,267
总部	1,910	3,485	2,622	2,688	1,828	1,494
合计净新增固定资产	30,018	57,976	72,325	60,836	48,344	85,752
北美电商物流	39%	52%	52%	39%	36%	28%
国际电商物流	11%	14%	14%	11%	9%	8%
AWS	44%	29%	30%	46%	51%	62%
总部	6%	6%	4%	4%	4%	2%
合计净新增固定资产	100%	100%	100%	100%	100%	100%

资料来源：公司官网，国信证券经济研究所整理测算 注：新增固定资产包括非现金部分，口径与Capex存在部分偏差

- 投入方向：
- AWS Capex除了传统CPU内存采购外（近年每年预计约200亿美元），**主要用于NVDA芯片采购、自研ASIC芯片（Trainium /Inferentia系列）以及数据中心土建配套等**，此前公司大幅押注自研芯片，但实际性能低于预期，同时NV卡采购偏少、数据中心重新调整，造成有效算力扩容偏慢，云业务增长落后竞争对手，AWS的收入和Capex本季度重新开始加速；

➤ 用途上，目前自用训练和推理较少，**主要用于AWS云服务租赁，以及支持投资公司anthropic推理训练；**

➤ **自研AI芯片Trainium2已发展为几十亿美元业务，相较于其他GPU选项具备30%-40%的价格性能优势**。25Q3业绩会公司表示自研AI芯片Trainium2 25Q3收入环比+150%，随着Trainium3（计划2025年底预览、2026年初量产）的推出与客户覆盖扩大，将进一步降低AWS AI服务的单位成本。

图：亚马逊拥有的算力规模（仅包括NV芯片数据中心）测算

年份	型号	GPU (千张)	F16理论峰值算力 (TFLOPS)	等效H100 (千片)	单价 (美元)	总价 (百万美元)	占比AWS当年资本开支	总浮点计算量 (Exaflops/FP16)	单芯片功耗 (W)	芯片总功耗 (MW)	对应AIDC功耗 (MW)	备注
2023	H系列	50	1979	50	25,000	1250	5%	99	700	35	80	25Q3业绩会表示过去12个月AWS新增3.8吉瓦电力，第四季度预计再增1吉瓦以上。该口径为供电能力，而非数据中心运转功耗；
2024	H系列	196	1979	196	25,000	4,900	9%	388	700	137	312	
2024	B200	12	4,600	28	30000	360	1%	55	1,000	12	27	
	2024之前总计			300		6,510		542			419	
2025	B200/300	360	4,600	837	32000	11,520	13%	1,656	1,000	360	818	
	总计			1,137		18,030		2,198			1,237	

- 23年底META表示24年底计划购买35万片英伟达H100，加上其他GPU约等效60万H100 GPU的计算能力。测算2025年底META约拥有250万H100 GPU等效计算能力。
- 2023年：数据中心建设成本约100亿美元，占当年Capex约35%；
- 2024年：数据中心建设成本约230亿美元，占当年Capex约60%；
- 表外融资项目：META与Blue Owl Capital于25年10宣布成立合资公司，共同开发Hyperion数据中心项目。META将与合资公司签订运营租赁协议、租期4年、可延期。总投资约270亿美元，Blue Owl 现金出资约70亿美元，其他通过发行私募债券筹集剩余资金。Blue Owl持有合资公司80%股权，Meta保留20%股权并负责数据中心的建设和运营管理。

图：META 数据中心建设计划

	上线时间	投资金额	算力/GW
德克萨斯州 El Paso 项目	2025年 上线	15 亿美元	达1GW
阿拉巴马州 Montgomery 项目	2026年 底上线	8 亿美元	
俄亥俄州 "Prometheus"	2026 年 上线		超1GW
路易斯安那州 "Hyperion" 超级集群	2030 年 完工	270亿美元， 与Blue Owl合 资，租赁形式	30年提供 2GW，未来可 扩展至5GW

资料来源：公司官网、国信证券经济研究所整理

图：META 拥有的算力规模预测

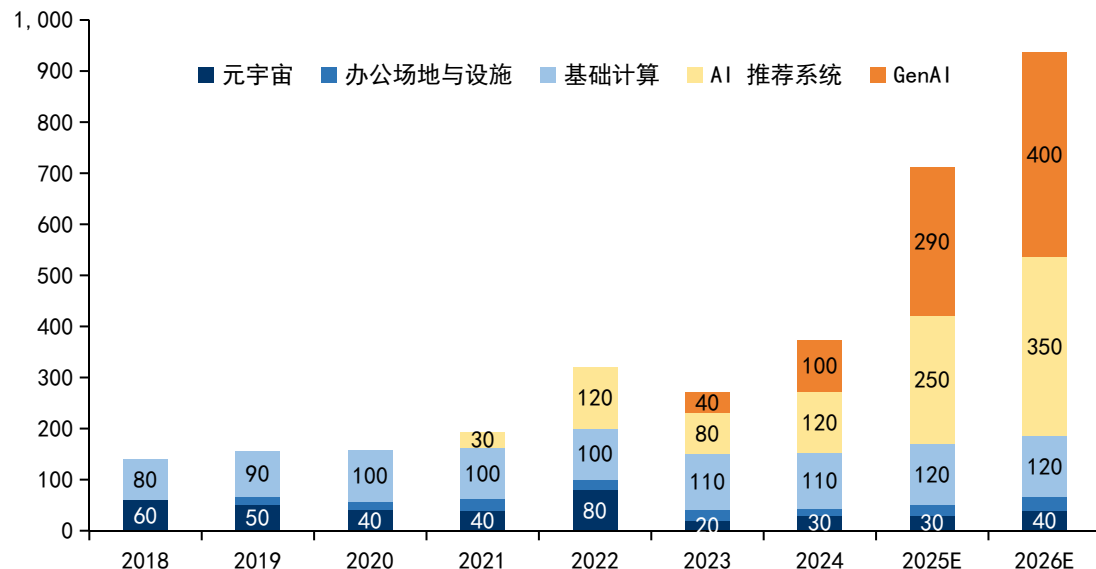
年份	型号	GPU (千张)	F16理论峰值算力 (TFLOPS)	等效H100 (千片)	单价 (美元)	总价 (百万美元)	占比当年资 本开支	总浮点计算量 (Exaflops/FP16)	单芯片功耗 (W)	芯片总功耗 (MW)	对应AIDC功耗 (MW)	备注
2023	H系列	150	1979	150	25,000	3750	14%	297	700	105	239	假设芯片占服务器功率80%， 假设服务器占AIDC功率55% 算力保有量相比当年资本开支存在一定时间差； 根据芯片官方价格假设单价。
2024	H系列	350	1979	350	25,000	8,750	23%	693	700	245	557	
2024	B200	40	4,600	93	30000	1,200	3%	184	1,000	40	91	
	2024之前总计	662		597		13,700		1,174			886	
2025	B200/300	800	4,600	1,860	32000	25,600	37%	3,680	1,000	800	1,818	
	总计			2,457		39,300		4,854			2705	

资料来源：TNP、Meta财报会、国信证券经济研究所整理

META Capex变化与业务分配：AI推荐与GenAI为主要投入环节

- AI推荐系统（AI content ranking） 累计投入Capex约600亿美元。
- “25Q2财报会：目前算力完全聚焦内部需求，如支持内容推荐、模型训练以及未来AI 应用推理需求。核心AI（广告）领域回报强劲，可量化且测量体系成熟。”
- “从商业化角度来看，AI在五大领域广告、内容体验、商业消息传递（whatsApp）、Meta AI（搜索聊天助手）、AI设备（AI眼镜）均有巨大商业化潜力，核心AI（广告）领域回报强劲，GenAI商业化先追求规模（覆盖数十亿用户）和产品质量，需数年再盈利。”
- 通用人工智能计算（GenAI） 累计投入Capex约430亿美元。
- “25Q2财报会：生成式AI处于早期，今明两年不会成为重要收入来源，但中长期来看均有巨大商业化潜力。”

图：META Capex使用分配



- 1) 办公场地与设施：非技术资本支出；
- 2) 元宇宙业务：主要与硬件相关，如VR/AR眼镜；
- 3) 核心应用家族（FoA）业务：
 - ① 基础计算（Core Compute）：主要是维护以托管和运行应用程序，包括像Reels 短视频功能上的视频内容；
 - ② AI推荐系统（AI content ranking）：
 - 内容推荐引擎底层硬件更新；
 - 广告技术平台升级，如 Advantage+新产品；
 - ③通用人工智能计算（GenAI）：如训练 Llama，为Meta AI提供推理算力以及AI Agent等。

资料来源：Meta财报会、Omdia、国信证券经济研究所整理测算

- X. AI融资情况：**2023年以来有报道融资预计超过200亿美元，且仍在募资**，包含初期股权融资约140亿美元， 2024年底股权融资约60亿美元，2025年7月股债组合融资约100亿美元，10月以来传闻进行新一轮200美元级股债组合融资
- **融资投入方向：主要用于算力集群建设和电力设施收购**。X. AI 算力集群建设以 “超高速迭代、规模化部署、垂直整合电力” 为核心特征，核心项目包括Colossus I与Colossus II，主要聚焦 GPU 规模扩张 + 电力设施自建，目标2025年实现30万GB200芯片集群（此目标下仅GPU芯片所需资金就接近100亿美元）
- **用途：**目前Grok在C端和B端应用用户规模目前较小，2025全年营收预计仅几亿美元，**GPU集群主要用于训练**，目前使用10万卡级别集群训练Grok系列，最新Grok4 性能比肩openAI处于第一梯队。

图：X. AI 拥有的算力规模

年份	型号	GPU (千片)	F16理论峰值算 力 (TFLOPS)	等效H100 (千片)	单价 (美元)	总价 (百万美元)	占比当年资 本开支	总浮点计算量 (Exaflops/FP16)	单芯片功耗 (W)	芯片总功耗 (MW)	对应AIDC功 耗 (MW)	备注
2023	H系列	50	1979	50	25,000	1250		99	700	35	80	假设芯片占服务器功率80%， 假设服务器占AIDC功率55%； 算力保有量相比当年资本开支存在一定时间差； 根据芯片官方价格假设单价。
2024	H系列	150	1979	150	25,000	3,750		297	700	105	239	
	2024之前总计			200		5,000		396				
2025	B200/300	300	4,600	697	32000	9,600		1,380	1,000	300	682	
	总计			897		14600		1776			1000	

资料来源：TNP、Omdia、公司官网，国信证券经济研究所整理测算

国内云厂商算力规模：加大国产算力采购与大集群建设

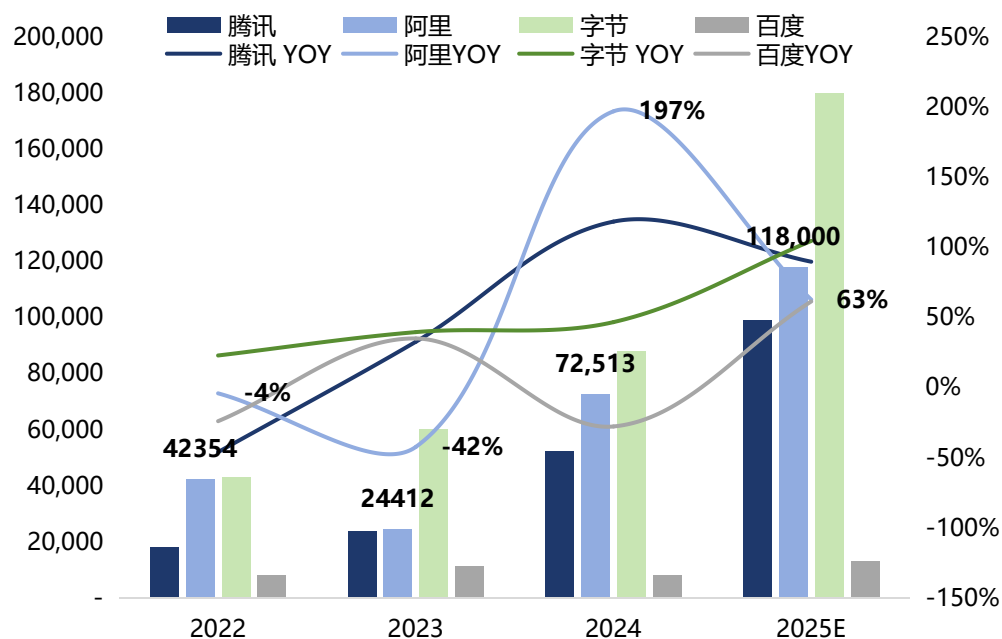
- 云服务商通过自研AI芯片建立从底层硬件到上层应用的全链条自主技术体系。自研芯片华为云以“鲲鹏+昇腾”生态为核心，昇腾系列芯片是其全栈自主的AI芯片产品；阿里平头哥推出含光800自研的AI推理芯片；腾讯投资如燧原科技等AI芯片初创公司，也推出紫霄AI推理芯片并成功流片。
- 国内云厂商在芯片采购商正逐步提升国产算力占比，并与国产芯片供应商紧密合作，云服务商加大投资大型智算集群建设。如海光DCU系列GPU与阿里云异构计算平台兼容，应用于大数据分析和AI训练场景。

表：国内云厂商AI芯片与算力布局

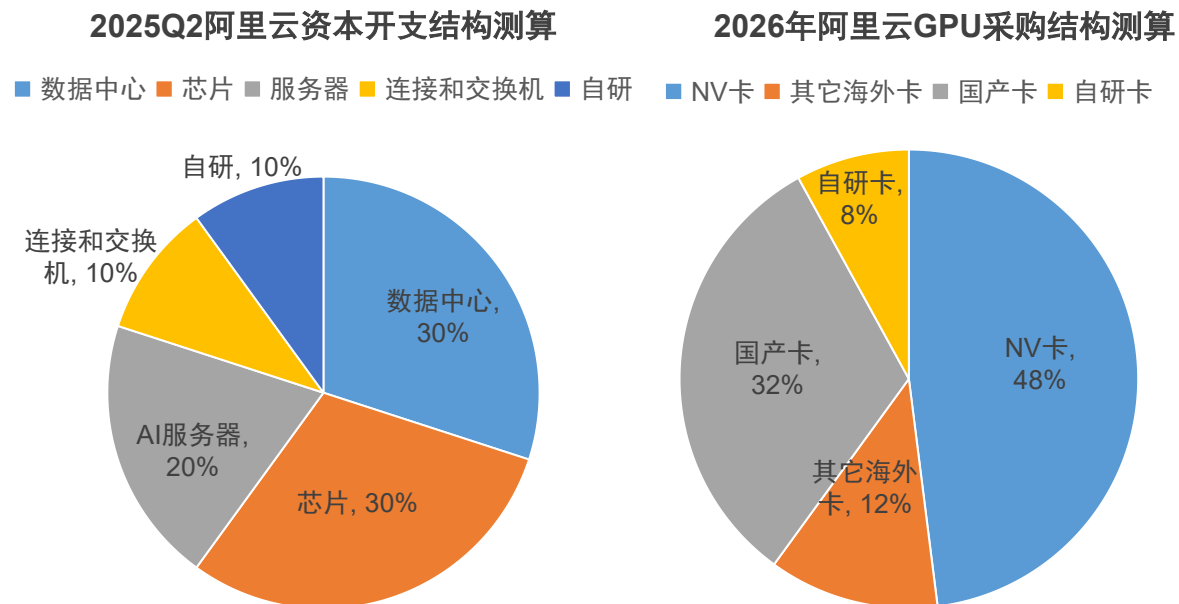
	阿里云	腾讯云	字节云
AI自研芯片	2018 年收购中天微，2019 年推出含光 800 AI 芯片，2023年阿里平头哥发布首个自研RISC-V AI平台。含光 900采用5nm工艺，算力提升至 25TFlops,计划2025年底前实现量产。	2021年披露：AI推理芯片“紫霄”、视频转码芯片“沧海”和智能网卡芯片“玄灵”。2023 年紫霄芯片流片成功并点亮，沧海芯片已量产计划推出紫霄 V2 Pro 芯片，以替代英伟达 L40芯片。	字节正自研两款AI芯片，分别用于模型训练和推理，计划与台积电合作，计划于2026年前实现量产。同时海外与博通开展ASIC芯片研发合作。
预计2024年算力保有量	NV卡30W+，国产卡45W+	NV卡40W+，国产卡1W+	NV卡37W+，国产卡5W+
预计2025年算力保有量	NV卡46W，国产卡20W+	NV卡53W，国产卡20W+	NV卡70W+，国产23W+
AI算力集群布局	阿里云打造灵骏超级智算集群，主要包含四个组件：灵骏计算集群、HPN高性能网络、磐久AI计算服务器以及CPFS高性能存储集群，提供可扩容到10万张GPU卡规模的能力。 25年，阿里云宣布正式启动张北超级智算中心，总建设规模为12 EFLOPS（每秒1200 亿亿次浮点运算）AI 算力， 2025年3月，阿里云宣布启动 "云智算" 计划，未来三年将新增 100万P（PetaFLOPS）算力集群，推动 AI 大模型训练成本降至行业均值的 1/3。	腾讯云仪征数据中心总投资超100亿元，计划部署至少 30万台服务器，算力规模达到 6100PFLOPS，预计 2025 年全面达产。 腾讯长三角人工智能先进计算中心已落地上海松江，预计建成后机柜数量接近 4 万架，服务器数量将达到 80 万台，算力是目前世界排名第一的超算中心的10倍。预计今年下半年投产。	2023年，字节已建立起超过1万张卡的 Ampere架构GPU（A100/A800）集群。 2024年9月，火山引擎内蒙古和林格尔算力中心一期总投资56亿元。 2025年1月，火山云太行算力中心二期总投资 45 亿元，对应功率187MW（测算含基建总功率374MW），规划 12kW 服务器机柜 1.56万台，支撑豆包大模型日均 16.4 万亿 tokens 的调用量。

- 阿里巴巴于2025年2月24日正式宣布，**未来三年将投入至少3800亿元人民币用于AI和云计算基础设施建设**：
 - 规模超前：超过阿里过去十年在云和AI基础设施上的投入总和(约3260亿元)；
 - 战略核心：CEO吴泳铭称AI是“一代人一次的机会”，AGI(人工通用智能)是公司主要长期目标；
 - 营收驱动：云业务成为阿里AI领域最清晰的收入驱动，**收入连续四季度加速增长，AI相关收入已连续六个季度实现三位数增长**；
- 测算目前阿里资本开支80%以上用于AI数据中心建设，包括AI服务器、数据中心基础设施及连接等其他业务板块。后续AI GPU采购将显著提升国产芯片与自研卡采购比例。

图：国内互联网与云厂Capex年度投资与增速变化情况（百万元、%）



图：阿里资本开支内部结构分布测算



资料来源：公司财报、彭博一致性预期、国信证券经济研究所整理

资料来源：公司官网、国信证券经济研究所测算整理测算

请务必阅读正文之后的免责声明及其项下所有内容

阿里云全球数据中心能耗十年提升十倍

- 到2032年，阿里云全球数据中心的能耗规模将比2022年 (GenAI 元年) 提升10倍，以支撑超级人工智能 (ASI) 时代的算力需求。阿里巴巴集团CEO吴泳铭于2025年9月24日在云栖大会上正式宣布。
- 根据测算，阿里云能耗预计从2022年的1.5GW至2032年的15GW、十年总资本开支规模累计约1.5万亿元。

表：阿里资本开支结构分布

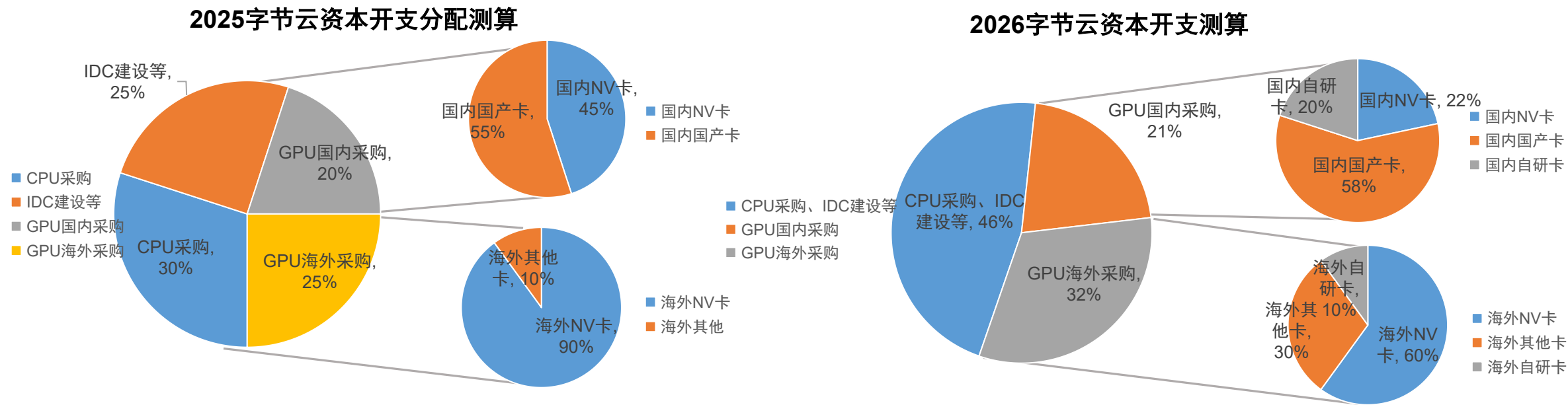
能耗规模测算：从2022年的1.5GW至2032年的15GW						
	服务器台数（万）	平均服务器功率（W）	能耗利用率	总能耗（GW）	能耗增长CAGR	说明
2022	300	400	1.2	1.44		参考杭州临平新城数据中心 服务器10W台与功耗47.4MW
2025	500	400	1.2	2.4	19%	
2032	3000	400	1.2	14.4	29%	
投资金额测算：从2022年至2032年资本开支规模累计约1.5万亿元						
	资本开支（亿元）	数据中心投入占比	新投资功率（GW）	平均功率投资成本	2026-2032平均年投资（亿元）	投资增长CAGR
2025	1300	80%	1.1	945		
2023-2032	15316	80%	13	945	1864	10%
算力规模测算：算力规模增长至少1643 ExaFLOPS(约1643*10 ⁶ TFLOPS)						
	总能耗（GW）	服务器功耗占比数据中心	服务器功耗（KW）	服务器卡数	单卡INT8算力（TOPS）	总算力规模(ExaFLOPS)
2032	14.4	50%	350	128	624	1643
说明：考虑算力的平均能耗显著下降，如Rubin NVL144 机柜的单位算力能耗相比H20服务器下降约一个数量级。以及阿里磐久超节点服务器峰值功耗350KW，实际算力规模增长有望高于下限数据。						

资料来源：公司官网、国信证券经济研究所测算整理测算
请务必阅读正文之后的免责声明及其项下所有内容

字节资本开支变化与分配结构

- 字节跳动AI资本开支呈现爆发式增长，2024年投入约800多亿元，2025年总资本开支预计将达1800亿元（同比增长100%+）。相比其他云厂商，字节资本开支具备以下特点：
 - 资本开支金额：投入量级遥遥领先，25年规模已接近百度、阿里、腾讯三家AI资本开支总和；
 - 资本开支分配：字节资本开支基本用于AI数据中心建设，其中AI算力采购与IDC 基建约各占50%；
 - 数据中心分布：相比国内云厂商，字节有较多海外业务，其海外数据中心建设投入金额超过国内；
 - 芯片采购策略：国内采购与国产芯片厂商有较好的合作关系（如寒武纪、华为昇腾等），海外采购则主要通过“算力飞地”（马来西亚、泰国等）获取高端GPU（如Blackwell系列），同时与Oracle有一定算力租赁合作，与博通开展ASIC自研芯片开发；

图：字节资本开支结构分布测算



- [01] 海内外云厂商资本开支复盘分析
- [02] 云厂商算力规模与建设计划
- [03] 云厂自研芯片布局与进展
- [04] 云厂AI云收入、AI ROI测算与估值

- **ASIC发展**：此前英伟达供应不足，同时部分计算任务场景固定化，主要云、模型企业对芯片供应链自主可控和降本诉求强烈，普遍布局自研ASIC，**此前投入研发的ASIC产品预计明年进入将密集落地期**。具体管线看，已较成熟的谷歌TPU将推出TPU v7代，而Meta、AWS推出下一代芯片，OpenAI落地首款博通合作芯片，微软等也有望后续推出自研芯片。此外老牌AI芯片厂商AMD也将推出新一代MI400系列。
- **NV卡和ASIC格局**：目前英伟达软硬件生态、性能功耗领先较大，依旧主导，从实际算力角度预计当前占据80%以上份额，尤其是在训练场景。ASIC等其他玩家持续迭代跟进，**TPU在谷歌内部已能实现复杂训练和推理，综合性能接近NVDA**，近期斩获Anthropic大单，其他ASIC仍处于迭代和适配阶段，预计先替换部分公司内部固定计算业务推理需求。

表：英伟达与海外科技厂自研ASIC芯片量产落地节奏

公司	24Q1	24Q3	24Q4	25Q1	25Q2	25Q4	26Q2	26Q3	26Q4	27Q3	2028及之后
英伟达Nvidia			B200/300					Rubin		Rubin Ultra	Feynman
超威AMD			MI325X		MI350/355X			MI400系列			新AMD芯片
谷歌Google		TPU v6e			TPU v6p		TPU v7			TPU v8	
亚马逊AWS			Trainium2			Trainium3					
Meta	MTIA 2			MTIA 2.5			MTIA 3				新MTIA芯片
微软Microsoft											Maia 200
苹果											Apple AI云芯片
OpenAI									OpenAI芯片		

注：预计批量供货时间。近期AMD和openAI合作显示MI450或于2026年下半年开始部署；而微软、苹果、openAI目均已启动AI芯片自研计划，但实际回片进展和量产节点不明确，时间轴仅供参考，OpenAI进展相对积极

资料来源：各公司官网、semianalysis、国信证券经济研究所测算整理测算

请务必阅读正文之后的免责声明及其项下所有内容

ASIC自研芯片开发的核心环节有哪些？

- 需求方：主要云厂/模型厂，项目的最终出资方和需求方，负责根据自身计算需求定义和设计核心单元架构，协调整体项目
- 协作设计/流片：博通等IC厂，负责IC设计配套、IP提供、物理实现、性能优化以及与台积电实现流片封装量产
- 软件开发生态：云厂为主，负责开发适配ASIC的软件栈，确保芯片性能释放，TPU为谷歌开发XLA，Meta芯片需适配 PyTorch
- 机柜/AIDC方案：CLS等，负责芯片在数据中心部署（如机柜电力分配、散热设计、网络互联），实现芯片集群化、规模化应用

各家ASIC合作与技术路线的选择侧重、进展如何？

- 谷歌：长期与博通合作，TPU围绕深度学习需求开发，已到第6代，软硬件全栈进展领先，目前TPU训推性能内部使用评价已接近NVDA，并开始逐步外供（比如通过GCP外供以及近期anthropic单独采购），明年新一代V7部分芯片或引入联发科合作；
- Meta：入局较晚，选择与已有TPU成功经验的博通合作，将推出MTIA 3，短期产品预计以内部使用为主，替代部分社交网络推荐算法需求；
- 亚马逊：自研能力较强，入局较早，此前与技术较弱但溢价较低的Marvell/Alchip合作，但本轮AI发展对芯片性能和集群要求显著提升，AWS AI云份额丢失，关注后续Trainium3以及与自研芯片Anthropic需求适配情况；
- OpenAI：同样入局较晚选择博通，旨在替代部分英伟达推理和混合训练需求，此前已与博通公告10GW合作项目，关注年末回片测试情况；
- 其他：此外微软、XAI、苹果等公司均有官宣或报道与博通等公司合作开发ASIC自研芯片，预计目前进度略慢于上文公司，关注后续进展；

表：海外科技厂自研ASIC芯片设计与代工产业链

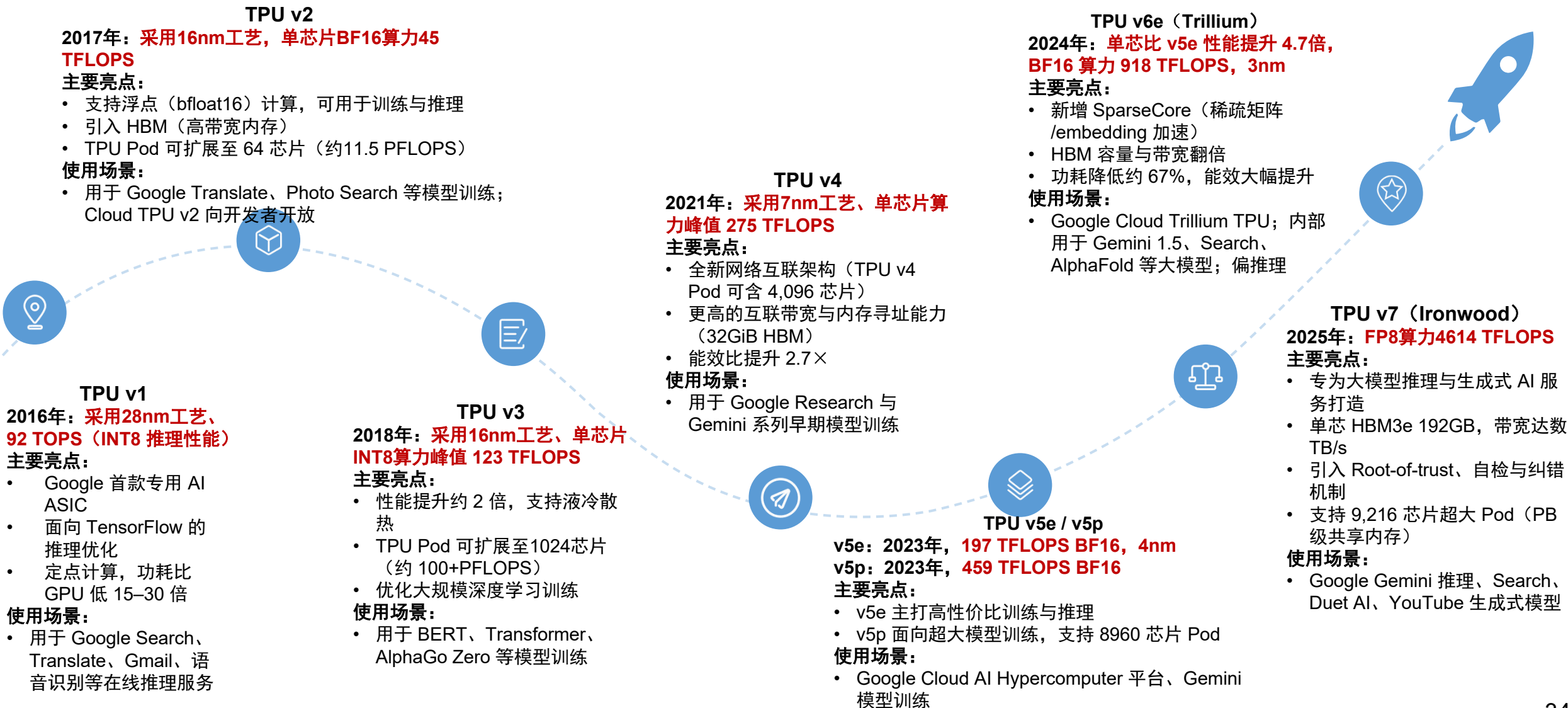
芯片名称	博通Broadcom	迈威尔Marvell	协助设计 联发科MTK	世芯电子Aichip	创意电子Unichip	代工 台积电TSMC	机柜方案 天弘CLS
谷歌	TPU v5p/5e/6p/6e/7p/7e		TPU v6e/7e(部分)			√	√
Meta	MTIA 2/3					√	√
亚马逊		Trainium2/Trainium3 (备用)		Trainium1/Inferentia2/Inferentia3/Trainium3(主要)		√	
微软					Maia 100/200	√	
OpenAI	名称待定					√	√
XAI/苹果/TKTK	名称待定						

资料来源：各公司官网、semianalysis、国信证券经济研究所测算整理测算

请务必阅读正文之后的免责声明及其项下所有内容

谷歌TPU发展历程与进展：针对深度学习和自身业务需求，持续迭代

图：谷歌TPU发展历程



ASIC（谷歌TPU为主）使用场景与英伟达GPU场景的区别

- **谷歌TPU**：设计目标**针对AI任务（尤其是深度学习训练场景）**，硬件架构（如脉动阵列、专用矩阵乘法单元）和软件系统（XLA）均针对神经网络优化，几乎不支持非AI场景（如图形渲染、科学计算）。
- **英伟达GPU**：**偏通用并行计算**，支持AI任务，也兼容科学计算、图形渲染、密码学等多场景，硬件（CUDA Core+Tensor Core）和软件（CUDA）兼顾灵活性与性能。

图：大模型背景下GPU与ASIC实际主要场景



大规模训练

主要为大参数大数据量LLM预训练场景，**对算力和集群互联性能要求高**，GPU和TPU（p系列）是目前**唯二两家实现成功训练的芯片集群产品**，其中**英伟达占据主导地位**，除谷歌之外几乎所有其他大模型厂商均以英伟达H/B系列训练大模型。谷歌TPU也已成功完成Gemini训练，性能处于第一梯队，目前主要为谷歌内部使用，生态广度不如英伟达，此外目前AWS trainium芯片在探索。另外，如果后续训练技术路线发生较大变化，则第三方生态更强、更通用的英伟达芯片有望占据新训练需求。



小规模训练

包括中/后训练，以提升模型垂类能力、finetuning为主，整体数据和计算量低于预训练，对算力和集群要求相对较低，目前**全行业也仍以英伟达为主**，**但谷歌内部主要使用TPU**，同时未来其他ASIC也有望实现小规模训练能力。



大模型推理

ASIC芯片后续重要目标使用场景，当前TPU（e系列）已支持谷歌模型大规模推理任务（如搜索AI Overview、Gemini推理等），且性价比比肩同代GPU，近期openAI与博通合作开发芯片也重点定位后续大模型推理需求。当然英伟达仍有性能和生态优势，预计头部大厂以外的第三方中小企业的推训需求仍将以英伟达为主，而且英伟达架构上也在朝AI能力持续优化。



其他推理

如Meta推荐系统，任务相对固定且对算力要求较低，ASIC预计将持续渗透，Meta MTIA目前芯片主要针对现有推荐系统，管理层表示后续将逐步替代英伟达芯片。

英伟达与谷歌TPU对比：性价比、能效比、综合生态对比

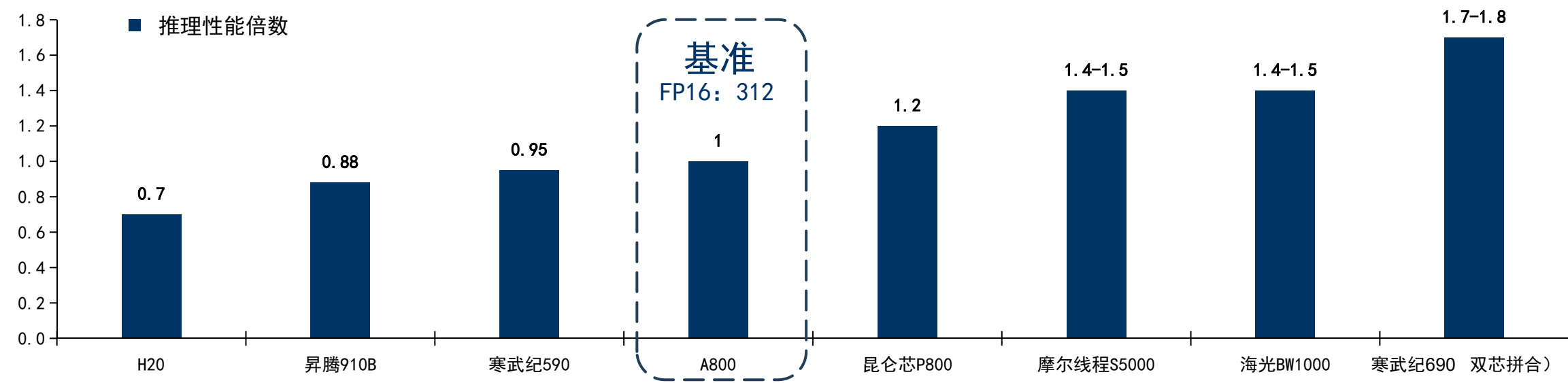
- 性价比：GB200和TPU v6e为目前业内最领先芯片，且均强调其推理性价比出色，我们认为**两者性价比基本可比**。从理论值测算看，**单位算力所需资金投入（TOPS/美元）Blackwell与TPU v6e接近**；实际业务效果看，TPU可针对谷歌业务需求做针对性和全栈优化，谷歌内部对TPU v6e评价积极，从GCP实时租赁价格看，v6e亦可与Blackwell性价比可比。
- 能效比：本代产品NVDA仍有优势。单位算力所需能耗（TOPS/W）理论指标看，Blackwell一定程度领先TPU v6e，而后续Rubin方案进一步强化推理性能，在目前北美电力缺口明显、各家云厂在手电力额度有限的背景下，NVDA功耗优势有望成为销售亮点。
- 生态对比：两者都完成全栈技术积累，但**英伟达生态更完善**，而TPU还处于从谷歌自用逐步转向行业外部阶段。两者在芯片设计、机柜组网方案、软件开发生态、大模型训练和应用推理方面均已完成技术积累，但英伟达面向全行业、生态完善，而TPU此前主要为谷歌内部，随着Anthropic等其他客户下单逐步完善生态和技术。

表：英伟达和谷歌TPU主要产品主要性能对比

1GW AIDC测算	Grace Blackwell	Vera Rubin	TPU v6e
功耗：			
单机柜理论功率：kW	130	225	
机柜功率/AIDC功率：%	55%	55%	
单机柜对应AIDC功率：kW	236	409	
芯片数：			
单柜芯片数：	72	144	
单芯片功率：kW	1.00	0.90	0.25
单芯片机柜下平均功率（包含柜内CPU冷却等）：kW	1.81	1.56	
单芯片AIDC下平均功率（包含AIDC配套等）：kW	3.28	2.84	0.75
1GW AIDC对应芯片数：	304,615	352,000	1,333,333
1GW AIDC对应机柜数：	4,231	2,444	
算力：FP8			
机柜理论峰值算力：EFLOPS 10^18次	0.72	3.60	
单芯片理论峰值算力：PFLOPS 10^15次	10.0	25.0	1.8
1GW 芯片理论峰值算力:EFLOPS 10^18次	3,046	8,800	2,400
Capex:			
芯片单价：千美元	30.0	45.0	4.5
芯片总价：百万美元	9,138	15,840	6,000
芯片/AIDC价值：%	45%	45%	45%
AIDC Capex：百万美元	20,308	35,200	13,333
单位经济：			
TFLOPS/W	3.0	8.8	2.4
美元/W	20	35	13
TFLOPS/美元	0.15	0.25	0.18
算力租赁价格：			
GCP小时租金：美元	9.3		1.3
小时FLOPs	36,000		6,480
FLOPs/HOUR/GPU	3,888		5,063

数据来源：英伟达官网、发布会、GCP、国信证券经济研究所整理 注：GCP租金为2025Q3数据

图：国内外芯片推理性能对比（以FP16精度下推理性能、A800为基准）

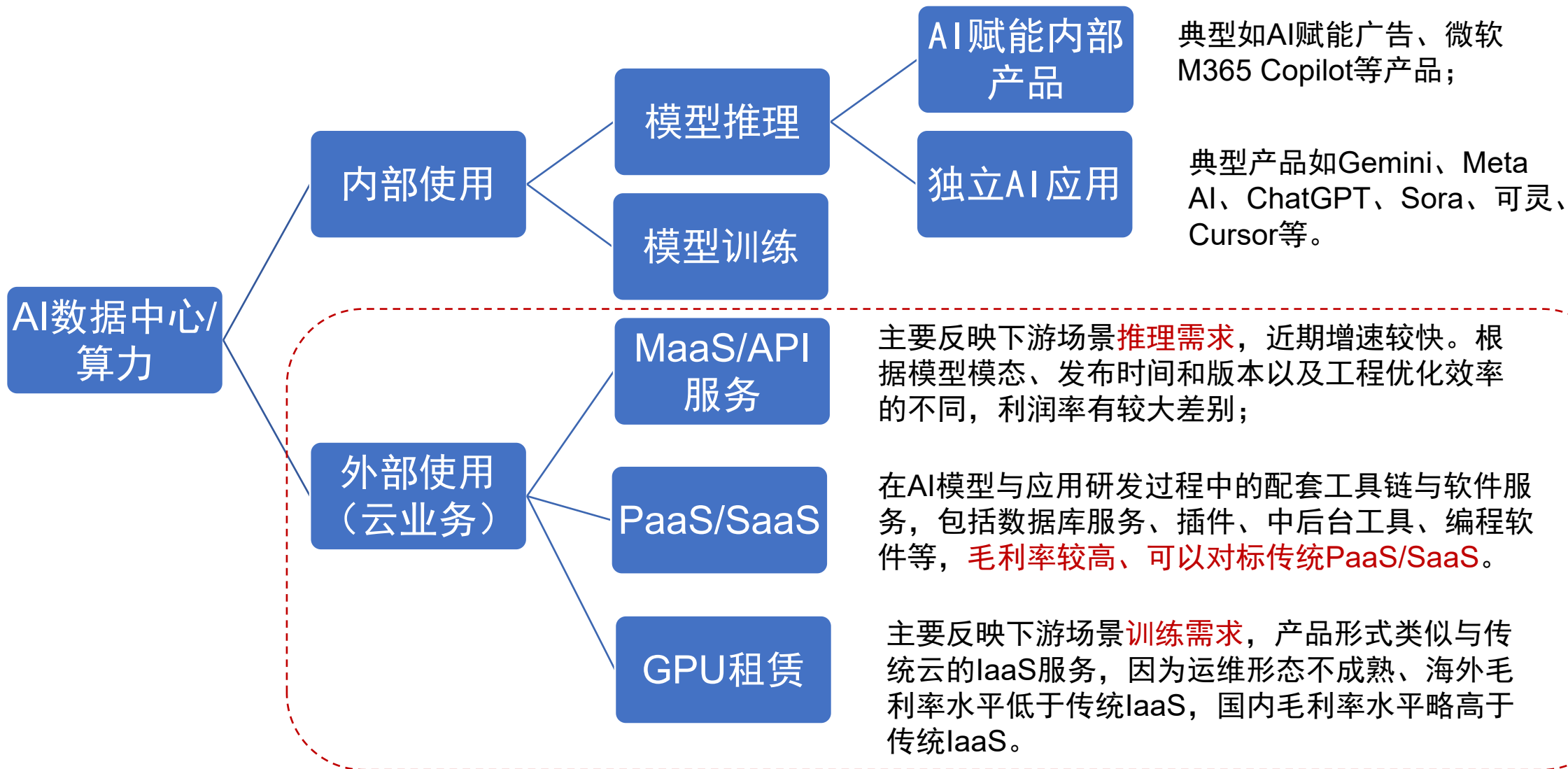


芯片型号	FP16 算力 (TFLOPS)	INT8 算力 (TOPS)	显存容量	显存带宽 (GB/s)	互联带宽	架构	价格 (万元)
英伟达 A800	312	624	80GB HBM2e	1935/2039s	PCIe4.0(64GB/s)、 NVLink(400GB/s)	Ampere 架构	-
英伟达 H20	148	296	96GB HBM3	4000	PCIe5.0(128GB/s)、 NVLink(900GB/s)	Hopper 架构	8
昆仑芯 P800	345	820	96GB HBM3	1800	PCIe5.0(700GB/s)	XPU-P 架构	8-10
寒武纪 690	512	800+	80GB HBM2e	2000	MLU-Link(512GB/s)	训推一体双 die 设计	-
寒武纪 590	314	512	80GB HBM2e	2000	PCIe4.0/5.0	MLUv02 扩展架构	-
昇腾 910B	320	640	64GB HBM2e	400	卡间互联 (392GB/s)	达芬奇架构	9-10
摩尔线程 S5000	400+	-	48GB	768	MTLink(240GB/s)、 PCIe5.0	平湖 架构	-
海光 BW1000	400+	-		896	PCIe5.0	自研架构	-

数据来源：各公司官网、腾讯新闻、国信证券经济研究所测算整理
请务必阅读正文之后的免责声明及其项下所有内容

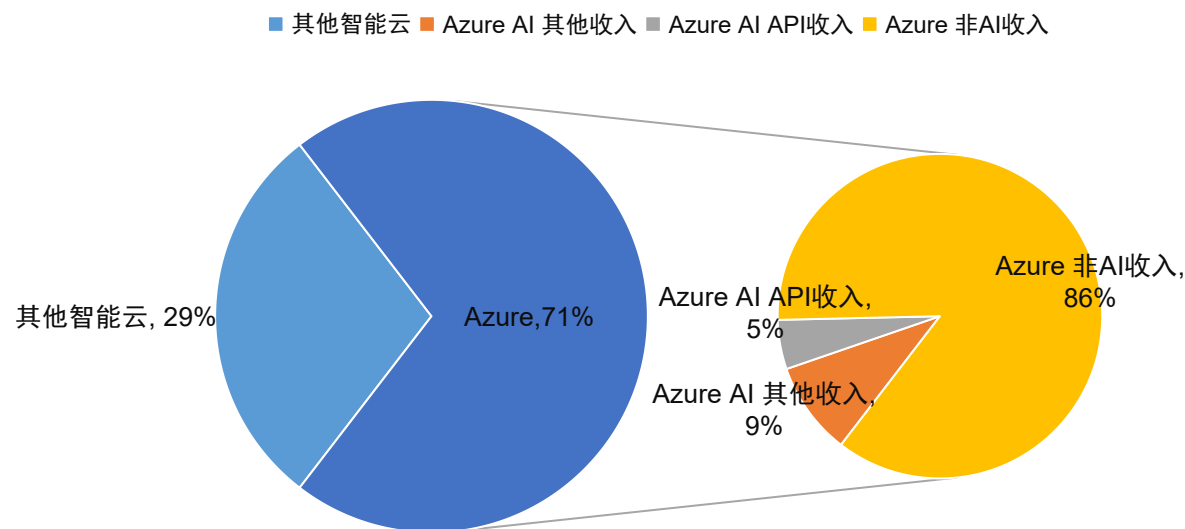
- 【 01 】 海内外云厂商资本开支复盘分析
- 【 02 】 云厂商算力规模与建设计划
- 【 03 】 云厂自研芯片布局与进展
- 【 04 】 云厂AI云收入、AI ROI测算与估值

图：AI数据中心算力分配与商业化变现路径

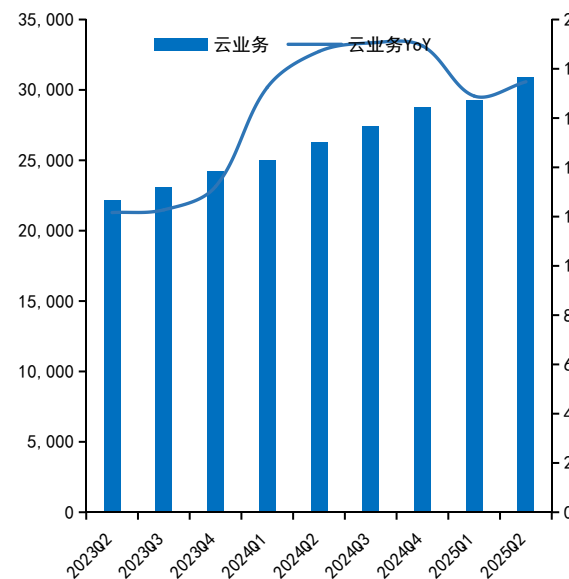


- 微软云：25Q2业绩会披露Azure年收入超过750亿美元，本季度未披露AI贡献Azure占比，上季度AI贡献Azure的16个点。2025 财年全年，Azure及其他云服务收入超过750亿美元，同比增长34%。25Q2披露M365 Copilot已有超1亿月活用户。
- AWS (Amazon)：25Q2业绩会表示AI相关收入继续保持三位数增速，25Q1 AI的年化收入达到数十亿美元。AWS目前仍处于供应不足。原因最大的是电力限制，另外芯片和组件数量不足、芯片交付节奏延迟、服务器良率不达预期等。
- Google Cloud：24Q4业绩会表示云AI基础设施和Gen AI解决方案年化数十亿美元的收入。AI基础设施主要是Vertex AI全托管式AI开发平台与定制化TPU芯片相关业务，为企业提供从模型训练到推理的全链路服务；而Gen AI解决方案则是以大模型为核心，包括Gemini、文生图Imagen和文生视频Veo等大模型调用的API和解决方案。

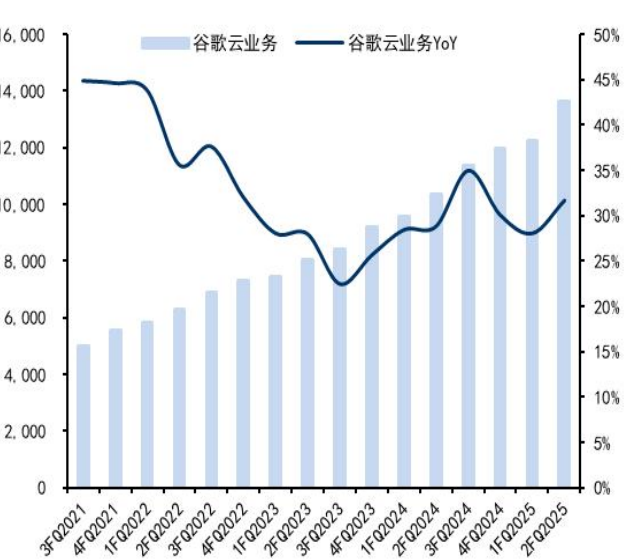
图：FY25（CY24Q3-CY25Q2）微软智能云收入结构测算



图：AWS收入与增速变化（百万美元，%）



图：谷歌云收入与增速变化（百万美元，%）

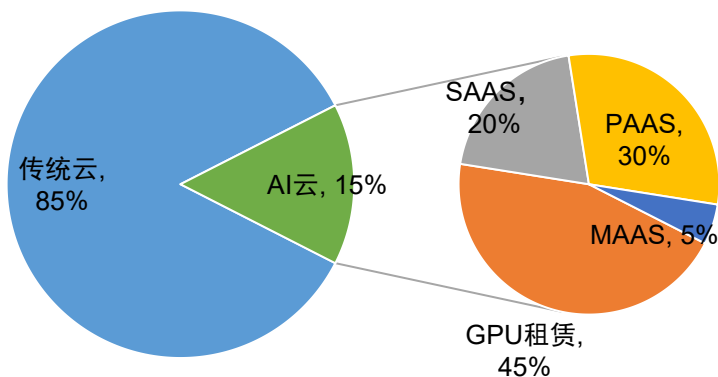


- 从AI云结构来看：GPU租赁是AI云收入主要来源，其次是配套的PaaS（数据库、插件等）与SaaS（代码助手、聊天助手等）服务，MaaS占比较小。但25Q3季度我们观察到MAAS层显著增长、GPU租赁端利润率小幅弱化。这种变化本质反应的是AI从训练转向推理需求为主，模型性能迭代到相当的智能化水平、从而推动下游AI应用开始跑通放量，因此不管是MAAS付费客户数量还是整体调用量都增长快速。其中字节调用量与增速遥遥领先，阿里云调用量次之，然后是百度。阿里云主要因客群结构差异，后者中大型客户较多、落地与放量周期更长。
- 从AI云利润率来看：GPU租赁毛利率高于国内传统IaaS服务，而MAAS即使不考虑训练成本、整体综合毛利也较低。除了模型层本身差距不明显外，各家更多把MaaS服务定位成AI生态的引流/获客产品，并提供较多的免费用量。根据不同的模型模态、版本以及工程优化效率，MaaS毛利率也有较大区别。

图：阿里、字节、腾讯AI云占比与内部结构分布测算

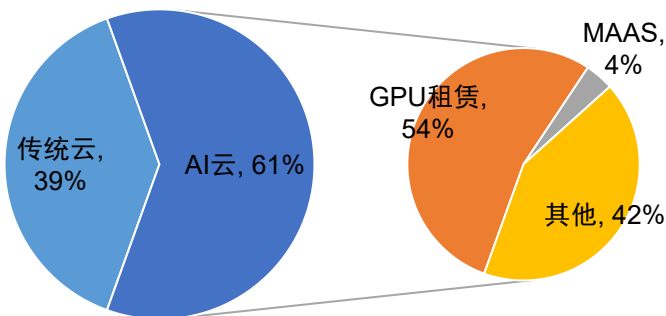
阿里25H1 AI云结构测算

■ 传统云 ■ GPU租赁 ■ SAAS ■ PAAS ■ MAAS



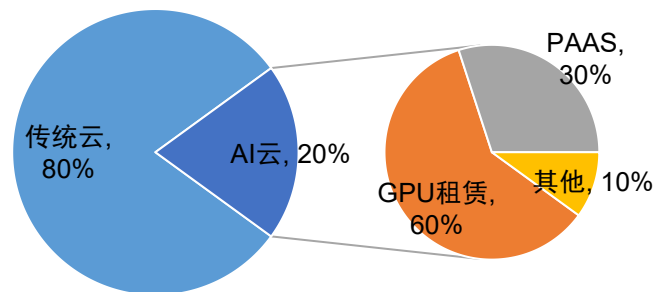
字节云25H1 AI云结构测算

■ 传统云 ■ GPU租赁 ■ MAAS ■ 其他



腾讯云25H1 AI云结构测算

■ 传统云 ■ GPU租赁 ■ PAAS ■ 其他



微软AI云业务投资与回报 ROI测算



- 前期高折旧摊销压力下AI云OPM低于当前智能云OPM水平（25Q3为43%），现金流端ROI小于1，后伴随着AI云规模增长与算力成本下降、利润率提升，2030年现金流贡献转正。资本回报端持续提升，2030年ROIC可达到17%，即六年可收回投资周期。

表：微软云AI投资与ROI测算

百万美元/%	2023	2024	2025E	2026E	2027E	2028E	2029E	2030E	备注
Capex (含融资租赁)	41,200	75,600	116,500	157,000	170,000	178,500	187,425	196,796	假设依据26FY增速高于25FY，当前微软净利润体量1000亿美金
Capex YOY	46%	83%	54%	35%	8%	5%	5%	5%	
Capex中AI投资占比	40%	72%	90%	90%	90%	90%	90%	90%	
Capex AI增量	16,630	54,080	104,850	141,300	153,000	160,650	168,683	177,117	假设依据：25Q3短期资产占50%，约6年折旧周期。长期资产占比50%、约15年折旧周期。
年中GenAI投入总量	8,360	45,130	124,595	247,670	394,820	551,645	716,311	889,211	
折旧（综合折旧年限为10年）	873	4,633	12,460	24,767	39,482	55,165	71,631	88,921	
运维成本	0	4,189	10,123	21,210	35,194	48,816	61,474	73,120	假设资本开支投出一年后开始使用，每年运维成本为总投入capex的12% 其余内部使用与OpenAI消耗
总成本	873	8,822	22,582	45,977	74,676	103,980	133,105	162,041	
云业务（外部）所用算力占比	40%	55%	55%	60%	60%	65%	65%	65%	
微软智能云收入	96,213	104,859	134,220	179,854	233,810	296,939	365,235	441,935	假设依据：IDC预测28年全球公有云约1.7W亿美元，假设微软份额约18%（3000亿美元）、当前13%，假设Azure占比智能云90%（当前75%）约2700亿美元。
YOY	17%	20%	28%	34%	30%	27%	23%	21%	
Azure收入	63,789	72,468	100,731	145,052	203,073	274,148	356,393	463,311	
其中Azure YOY	28%	31%	39%	44%	40%	35%	30%	30%	假设依据算力成本下降、规模效应提升。当前云业务OPM为43%
AI贡献Azure收入	3%	10%	19%	28%	35%	40%	45%	47%	
AI云收入	4,441	11,318	19,139	40,615	71,076	109,659	160,377	217,756	
AI云YoY		155%	69%	112%	75%	54%	46%	36%	假设依据算力成本下降、规模效应提升。当前云业务OPM为43%
AI云毛利率测算		57%	35%	32%	37%	38%	46%	52%	
AI云OPM		45%	20%	18%	24%	26%	35%	42%	
云非AI收入		93,541	115,081	139,240	162,735	187,280	204,858	224,179	假设依据算力成本下降、规模效应提升。当前云业务OPM为43%
YoY			23%	21%	17%	15%	9%	9%	
(AI云营业利润+折旧) /当年AI投入			19%	26%	44%	62%	94%	129%	
ROIC (AI云营业利润/累计AI投入)			6%	5%	8%	9%	13%	17%	大于100%则当年现金流正贡献 资本回报率：17%则6年收回投资

资料来源：公司财报会、彭博、semianalysis，国信证券经济研究所测算整理测算 *口径并未包括对传统云业务、内部其他业务自用以及对OpenAI投资回报

请务必阅读正文之后的免责声明及其项下所有内容

谷歌AI云业务投资与回报 ROI测算

- 目前AI云OPM略低于当前谷歌云OPM水平（25Q3为26%），现金流端ROI小于1，后伴随着AI云规模增长与算力成本下降、利润率提升，**2030年现金流贡献转正**。资本回报端持续提升，**2030年ROIC可达到17%，即六年可收回投资周期**。

表：谷歌云AI投资与ROI测算

百万美元/%	2023	2024	2025E	2026E	2027E	2028E	2029E	2030E	备注
Capex (含融资租赁)	32,251	52,535	92,096	120,000	144,000	158,400	166,320	174,636	
Capex YOY	2%	63%	75%	30%	20%	10%	5%	5%	假设依据当前公司资本开支指引
Capex中AI投资占比	39%	74%	90%	90%	90%	90%	90%	90%	
Capex AI增量	12,685	38,735	82,886	108,000	129,600	142,560	149,688	157,172	24Q3:技术基础设施投资里60%是服务器(GPU+TPU服务器)，40%是数据中心建设(厂房建安、冷却、电气设备配套等)和网络连接设备支出。
年中GenAI投入总量	5,564	34,346	95,157	190,600	309,400	445,480	591,604	745,035	
折旧 (综合折旧年限为9年)	714	3,889	10,573	21,178	34,378	49,498	65,734	82,782	
运维成本	0	3,084	7,265	15,943	26,362	37,788	48,956	59,030	假设资本开支投出一年后开始使用，包含电力、租金、人力、软件摊销和授权、带宽等，数据中心每年运维成本为总投入capex的12%
总成本	714	6,973	17,839	37,121	60,740	87,286	114,690	141,812	
云业务 (外部) 所用算力占比	21%	25%	33%	35%	38%	40%	41%	42%	谷歌数据中心中GPU与TPU算力比例约为6:4。其中GPU 70%对外
谷歌云收入	1,182	8,587							
YOY	33,088	43,229	57,441	80,000	115,000	152,950	198,835	248,544	假设依据：IDC预测28年全球公有云约1.7W亿美元，假设谷歌份额约9%(1500亿美元)、当前6%。
AI云收入占比	26%	31%	33%	39%	44%	33%	30%	25%	
AI云收入	2%	8%	15%	25%	35%	40%	45%	48%	
AI云收入YoY	423	2,436	8,616	20,000	40,250	61,180	89,476	119,301	
AI云毛利率测算		476%	254%	132%	101%	52%	46%	33%	考虑TPU成本优势，假设依据算力成本下降、规模效应提升。
AI云OPM		28%	32%	35%	43%	43%	47%	50%	
云非AI收入	25%	26%	17%	21%	30%	31%	36%	40%	Anthropic 将合作接入多达 100 万个谷歌的 TPU，价值数百亿
YoY	32,665	40,793	48,825	60,000	74,750	91,770	109,359	129,243	
		25%	20%	23%	25%	23%	19%	18%	
(AI云营业利润+折旧) /当年AI投入			18%	31%	51%	68%	97%	125%	大于100%则当年现金流正贡献
ROIC (AI云营业利润/累计AI投入)			5%	7%	11%	12%	15%	17%	

资料来源：公司财报会、彭博、semianalysis，国信证券经济研究所测算整理测算 *口径并未包括对传统云业务、内部其他业务自用测算

阿里AI云业务投资与回报 ROI测算

- 测算近两年AI云OPM与阿里云OPM水平相近，现金流端ROI小于1，后伴随着AI云规模增长、利润率提升，**2030年现金流贡献转正**。资本回报端持续提升，**2030年ROIC可达到11%，即10年可收回投资周期**。

表：阿里云AI投资与ROI测算

百万元/%	2023	2024	2025E	2026E	2027E	2028E	2029E	2030E	备注
Capex (含融资租赁)	24,412	72,513	123,272	133,134	139,791	153,770	169,147	186,061	未来三年将投入至少3800亿元人民币用于AI和云计算基础设施建设，到2032年阿里云全球数据中心的能耗规模将比2022年(GenAI元年)提升10倍
Capex YOY	2%	197%	70%	8%	5%	10%	10%	10%	
Capex中AI投资占比	23%	67%	85%	85%	88%	90%	90%	90%	
Capex AI增量	5,613	48,908	104,781	113,164	123,016	138,393	152,232	167,455	资本开支中约30%用于IDC建设，其余用于AI服务器与配套。
年中GenAI投入总量	1,556	24,568	101,413	210,385	328,475	459,179	604,491	764,335	
折旧 (综合折旧年限为7年)	324	4,102	14,488	30,055	46,925	65,597	86,356	109,191	
运维成本	0	1,542	9,159	24,509	38,638	51,573	63,948	75,147	包含电力、租金、人力、软件摊销和授权、带宽等，假设每年运维成本为总投入capex的17%
总成本	324	5,644	23,647	54,564	85,563	117,170	150,304	184,338	GPU 70%对外
云业务 (外部) 所用算力占比	50%	60%	70%	73%	75%	78%	80%	80%	
阿里云收入	99,419	113,496	145,275	196,121	258,880	336,544	430,776	542,778	
YOY	26%	14%	28%	35%	32%	30%	28%	26%	假设依据算力成本下降、国产芯片与自研芯片比例提升、规模效应提升。
AI云收入占比	4%	9%	15%	25%	35%	40%	43%	45%	
AI云收入	3,540	9,856	21,791	49,030	90,608	134,618	185,234	244,250	
AI云收入YoY		178%	121%	125%	85%	49%	38%	32%	
AI云毛利率测算			24%	19%	29%	32%	35%	40%	
AI云OPM			4%	1%	11%	15%	20%	25%	
云非AI收入	95,879	103,640	123,484	147,091	168,272	201,926	245,542	298,528	
YoY		8%	19%	19%	14%	20%	22%	22%	
(AI云营业利润+折旧) /当年AI投入			15%	27%	49%	66%	87%	110%	
ROIC (AI云营业利润/累计AI投入)			1%	0%	5%	7%	9%	11%	大于100%则当年现金流正贡献 资本回报率：11%则10年收回投资

资料来源：公司财报会、彭博、semianalysis，国信证券经济研究所测算整理测算 *口径并未包括对传统云业务、内部其他业务自用测算

44

海外云业务估值：

- ① 微软一般为PE估值，彭博一致预期2025年31x PE、11xPS。云业务增速更高、利润率与总业务持平，给予更高估值约40xPE，约14xPS。
- ② 亚马逊通常采用EV/EBITDA估值或PE估值，一致预期下2025年28xPE，分布估值法下给AWS 25年25xPE、约8xPS。
- ③ 谷歌云一般采用一般为PE估值，彭博一致预期2025年24x PE。其中云业务增速为广告的三倍，25年OPM提升至21%、利润低率约为广告的1/2，给与10x PS；

表：互联网云厂商云业务增速与估值比较

云业务 (百万)	2023 收入	2024收入	2025E收入	2025 YOY	2025 OPM	2025年估值倍数 (PS)	市值贡献	占比当前市值
微软云	\$96,213	\$104,859	\$120,026	26%	42%	14	1,680,364	44%
AWS	\$90,757	\$107,556	\$127,857	19%	35%	8	1,022,856	39%
谷歌云	\$33,088	\$43,229	\$57,441	33%	21%	10	574,410	17%
阿里云	¥ 105,521	¥ 113,496	¥ 140,735	24%				
百度云	¥ 18,748	¥ 21,805	¥ 27,474	26%				
腾讯云	¥ 39,198	¥ 42,126	¥ 50,551	20%				

资料来源：各公司财报、彭博一致性预期、国信证券经济研究所测算整理 注：采用10月31日收盘价，阿里云25Q2财报披露经调整EBITA 8.8%

第一，宏观经济波动。若宏观经济波动，公司业务、产业变革及新技术的落地节奏或将受到影响。

第二，下游需求不及预期。若下游AI需求不及预期，相关的AI研发投入增长或慢于预期，致使行业增长不及预期。

第三，核心技术水平升级不及预期的风险。AI大模型研发进度落后，AIGC相关产业技术壁垒较高，核心技术难以突破，影响整体进度。

第四，AI快速迭代、平权化下竞争加剧，影响云业务利润率。

免责声明



国信证券投资评级			
投资评级标准	类别	级别	说明
报告中投资建议所涉及的评级（如有）分为股票评级和行业评级（另有说明的除外）。评级标准为报告发布日后6到12个月内的相对市场表现，也即报告发布日后的6到12个月内公司股价（或行业指数）相对同期相关证券市场代表性指数的涨跌幅作为基准。A股市场以沪深300指数（000300.SH）作为基准；新三板市场以三板成指（899001.CSI）为基准；香港市场以恒生指数（HSI.HI）作为基准；美国市场以标普500指数（SPX.GI）或纳斯达克指数（IXIC.GI）为基准。	股票投资评级	优于大市	股价表现优于市场代表性指数10%以上
		中性	股价表现介于市场代表性指数±10%之间
		弱于大市	股价表现弱于市场代表性指数10%以上
		无评级	股价与市场代表性指数相比无明确观点
	行业投资评级	优于大市	行业指数表现优于市场代表性指数10%以上
		中性	行业指数表现介于市场代表性指数±10%之间
		弱于大市	行业指数表现弱于市场代表性指数10%以上

分析师承诺

作者保证报告所采用的数据均来自合规渠道；分析逻辑基于作者的职业理解，通过合理判断并得出结论，力求独立、客观、公正，结论不受任何第三方的授意或影响；作者在过去、现在或未来未就其研究报告所提供的具体建议或所表述的意见直接或间接收取任何报酬，特此声明。

重要声明

本报告由国信证券股份有限公司（已具备中国证监会许可的证券投资咨询业务资格）制作；报告版权归国信证券股份有限公司（以下简称“我公司”）所有。本报告仅供我公司客户使用，本公司不会因接收人收到本报告而视其为客户。未经书面许可，任何机构和个人不得以任何形式使用、复制或传播。任何有关本报告的摘要或节选都不代表本报告正式完整的观点，一切须以我公司向客户发布的本报告完整版本为准。

本报告基于已公开的资料或信息撰写，但我公司不保证该资料及信息的完整性、准确性。本报告所载的信息、资料、建议及推测仅反映我公司于本报告公开发布当日的判断，在不同时期，我公司可能撰写并发布与本报告所载资料、建议及推测不一致的报告。我公司不保证本报告所含信息及资料处于最新状态；我公司可能随时补充、更新和修订有关信息及资料，投资者应当自行关注相关更新和修订内容。我公司或关联机构可能会持有本报告中所提到的公司所发行的证券并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问或金融产品等相关服务。本公司的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告意见或建议不一致的投资决策。

本报告仅供参考之用，不构成出售或购买证券或其他投资标的的要约或邀请。在任何情况下，本报告中的信息和意见均不构成对任何个人的投资建议。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。投资者应结合自己的投资目标和财务状况自行判断是否采用本报告所载内容和信息并自行承担风险，我公司及雇员对投资者使用本报告及其内容而造成的一切后果不承担任何法律责任。

证券投资咨询业务的说明

本公司具备中国证监会核准的证券投资咨询业务资格。证券投资咨询，是指从事证券投资咨询业务的机构及其投资咨询人员以下列形式为证券投资人或者客户提供证券投资分析、预测或者建议等直接或者间接有偿咨询服务的活动：接受投资人或者客户委托，提供证券投资咨询服务；举办有关证券投资咨询的讲座、报告会、分析会等；在报刊上发表证券投资咨询的文章、评论、报告，以及通过电台、电视台等公众传播媒体提供证券投资咨询服务；通过电话、传真、电脑网络等电信设备系统，提供证券投资咨询服务；中国证监会认定的其他形式。

发布证券研究报告是证券投资咨询业务的一种基本形式，指证券公司、证券投资咨询机构对证券及证券相关产品的价值、市场走势或者相关影响因素进行分析，形成证券估值、投资评级等投资分析意见，制作证券研究报告，并向客户发布的行为。



国信证券
GUOSEN SECURITIES

国信证券经济研究所

深圳

深圳市福田区福华一路125号国信金融大厦36层

邮编：518046 总机：0755-82130833

上海

上海浦东民生路1199弄证大五道口广场1号楼12楼

邮编：200135

北京

北京西城区金融大街兴盛街6号国信证券9层

邮编：100032