

公司研究

“港股 GPU 第一股”，国产算力新标杆

——壁仞科技（6082.HK）投资价值分析报告

要点

壁仞科技专注 GPGPU 芯片研发，2026 年 1 月作为“港股 GPU 第一股”成功港股上市。公司自主研发出了智能计算整体解决方案，涵盖 GPGPU 架构与芯片的硬件系统和 BIRENSUPA 计算软件平台。股权方面，创始人张先生及其一致行动人为控股股东（15.35%），IPO 全球发售股份占比 13.4%，IPO 募资净额为 64.2 亿港元、绝大部分将用于产品研发和技术迭代。

中国智能计算芯片市场高速增长，国产芯片有望开启放量。根据灼识咨询，中国智能计算芯片的市场规模从 2020 年的 17 亿美元快速增长至 2024 年的 301 亿美元（CAGR 为 105%），预计 2029 年中国智能计算芯片的市场规模将达到 2012 亿美元，对应 2024-2029 年 CAGR 为 46.3%，市场高速增长主要受 AI 商业化加速、AI 计算需求高增、科技进步和政策利好驱动。当前智能计算芯片市场呈现头部高度集中局面，大部分市场份额主要由海外厂商占据，2024 年第一大厂商占据 76.2% 的份额。随着中国企业技术迭代和竞争力提升，中国厂商将凭借更强竞争力不断提升市场份额，公司预计国产智能计算芯片占据的中国市场份额将从 2024 年的约 20% 提升至 2029 年的约 60%。

公司自研软硬件构建技术壁垒，芯片迭代推动未来销售增长和客户多元化。1) 产品：自研 GPGPU 芯片+BIRENSUPA 软件平台，并配套不同配置的 PCIe 板卡、OAM、GPGPU 服务器或服务器集群，支持解决方案出货。**2) 技术：**拥有 GPGPU 架构、Chiplet、集群互联能力等核心技术，管理团队和研发团队强大。**3) 客户和商业化：**当前客户集中度较高，BR106 芯片为销售主力产品，新品 BR166 和下一代 BR20X 芯片将推动公司芯片销售进一步增长。**4) 供应链：**晶圆制造&封测由第三方提供，实体清单后采购转向国产替代。

营收快速增长，持续研发投入，公司尚处于商业化初期。2023~2025 年公司分别实现收入 62.0/336.8/1034.6 百万元人民币，高速增长。销售产品为主要收入来源，其中智能计算解决方案占据主导。毛利率方面，2023~2025 年毛利率分别为 76.4%/53.2%/53.8%。公司仍处商业化初期，研发投入较高，目前净利润呈亏损状态，主要系实现规模化之前仍需完成芯片设计、客户验证及市场渗透，收入尚未涵盖前期投入，相关规模经济效益仍需逐步释放。

盈利预测、估值与评级：我们预计壁仞科技 2026/2027/2028 年分别实现营收 23.38/70.47/123.03 亿人民币元，同比增长 126%/201%/75%；净利润分别为 -8.69/9.26/28.35 亿人民币元，调整后净利润分别为 -7.22/10.99/30.43 亿人民币元。当前股价对应 2026/2027 年预测 PS 为 46x/15x。目前壁仞科技的 2027 年 PS 估值水平仍低于同行业公司，未来潜在催化：1) 产品，BR166 和 BR20X 等高性能芯片放量；2) 供应链，中国先进制程产能的持续爬升，公司的供应链将在 2026 年底到 2027 年明显改善；3) 客户，期待公司开启向互联网等更多元客户的商业化进程。考虑到国产算力出货和公司商业化开启，看好公司业绩增长和估值抬升。首次覆盖，给予公司“买入”评级。

风险提示：算力需求下行；供应链风险；行业竞争加剧；技术迭代不及预期；商业化进展不及预期；次新股股价波动风险。

公司盈利预测与估值简表

	2024	2025	2026E	2027E	2028E
营业收入（百万元）	337	1,035	2,338	7,047	12,303
营业收入增长率	443.0%	207.2%	126.0%	201.4%	74.6%
净利润（百万元）	(1,538)	(16,493)	(869)	926	2,835
调整后净利润（百万元）	(767)	(874)	(722)	1,099	3,043
EPS（元）	(0.93)	(8.72)	(0.36)	0.38	1.16
P/S	217	81	46	15	9

资料来源：公司公告，wind，光大证券研究所预测，股价时间为 2026-05-06；1 HKD = 0.8717 CN¥；公司 2020~2025 年进行多次上市前融资、2026 年 IPO 融资，2024/2025 年股本为 16.46/18.92 亿股，2026~2028 年股本假设为当前股本 24.39 亿股。

买入（首次）

当前价：51.00 港币

作者

分析师：付天姿，CFA，FRM

执业证书编号：S0930517040002

021-52523692

futz@ebsecn.com

分析师：董馨悦

执业证书编号：S0930525070008

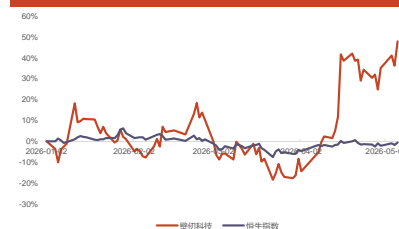
021-52523858

dongxinyue@ebsecn.com

市场数据

总股本(亿股)	24.39
总市值(亿港元)	1,243.82
一年最低/最高(港币)	27.32-52.00
近 3 月换手率(%)	62.9

股价相对走势



资料来源：Wind

收益表现

	%	1M	3M	1Y
相对	68.5	57.6	/	/
绝对	72.9	55.2	/	/

资料来源：Wind

投资聚焦

关键假设

营收预测方面：按照“量*价”的逻辑测算。分别假设 BR166、BR20X、BR30X 等未来产品的出货节奏、下半年确收更强的季节性影响，综合利用公司在手订单、存货金额等预测中短期营收，利用市场规模和公司市占率预测长期营收。

毛利率预测方面：一方面参考可比公司的当前毛利率水平来预计公司未来两三年毛利率水平，另一方面因切换国产供应链且初期放量，我们预计 26-27 年毛利率略有下滑，但长期考虑到供应链（晶圆代工、封装测试）技术迭代和良率提升，叠加出货规模化效应凸显，单颗芯片 Die 的成本增长幅度应低于价格提升幅度。

费用预测方面：由于公司处于商业化早期阶段，营收规模仍有限导致费用率高且不稳定，因此我们对各费用分别假设同比增速，其中销售及营销开支伴随营收规模而高速增长、一般及行政开支同比增速稳健、研发开支因产品和技术迭代持续高速增长。

我们区别于市场的观点

市场认为，国产芯片普遍性能偏弱。我们通过量化数据对比发现，壁仞科技芯片产品在 TPP 性能角度处于国产芯片第一梯队，基本可对标英伟达 H 系列芯片水平，伴随互联网厂商大规模验证国产 AI 芯片，我们预计 2026 年将成为国产芯片在推理侧“好用”元年。

市场认为，国产芯片供应链问题仍不明朗。我们看好先进制程产能释放和良率提升，预计 2026 年底、2027 年后“国产、国造”芯片的供应将明显改善，帮助中小型国产 AI 芯片厂商的产品加速放量和商业化。

股价上涨的催化因素

公司处于商业化初期阶段，出货放量和客户拓展成为关键。关注 1) BR166 和 BR20X 等新芯片的性能表现和验证效果；2) 公司芯片产能和良率的提升，助力芯片出货规模化、稳定化；3) 切入互联网云服务厂商等新客户的商业化突破。

估值与评级

我们预计壁仞科技 2026/2027/2028 年分别实现营收 23.38/70.47/123.03 亿人民币元，同比增长 126%/201%/75%；净利润分别为-8.69/9.26/28.35 亿人民币元，调整后净利润分别为-7.22/10.99/30.43 亿人民币元。当前股价对应 2026/2027 年预测 PS 为 46x/15x。

目前壁仞科技的 2027 年 PS 估值水平仍显著低于同行业公司，叠加 FCFF 法估值得出公司合理价格为 63.67 港币。未来潜在催化：1) 产品迭代方面，公司伴随 BR166 和 BR20X 等高性能芯片放量，带来更高营收贡献；2) 供应链方面，2026 年将成为国产国造芯片的元年，随着中国先进制程产能的持续爬升，我们预计公司的供应链问题将在 2026 年底到 2027 年明显改善；3) 客户拓展方面，期待公司加速向运营商、国家算力场、大型企业等客户出货，并积极进行互联网云服务商的适配和验证通过，开启向互联网等更多元客户的商业化进程。考虑到国产算力出货和公司商业化开启，看好公司业绩增长和估值抬升。首次覆盖，给予公司“买入”评级。

目 录

1、公司概况：港股 GPGPU 第一股	6
1.1 发展历程：港股 GPU 第一股，2026 年 1 月正式港股上市	6
1.2 业务：GPGPU 芯片+BIRENSUPA 软件平台，构成智能计算解决方案	6
1.3 融资情况：IPO 以前经历 10 轮融资，最新投后估值约为 209 亿元	7
1.4 股本结构：全球发售 H 股占比 13.4%，创始人张先生及其一致行动人为控股股东	8
1.5 股权结构：创始人张先生及其一致行动人为控股股东，合计持股 15.4%	8
1.6 募资用途：聚焦智能计算解决方案研发迭代，兼顾商业化与营运支持	9
2、行业概况：中国智能计算芯片市场高速增长，国产芯片有望出货放量	10
2.1 行业空间：智能计算芯片市场高速增长，预计中国快于全球增速	10
2.2 需求：国产大模型异军突起，云厂商大举投入，助推国产算力需求高增	11
2.2.1 国内公司积极备战 AI 算力军备竞赛，加码 CAPEX 投入 AI 基础设施建设	11
2.2.2 国内大模型调用量大幅增长，Tokens 出海开启	11
2.3 竞争格局：国产芯片积极产品迭代+性能升级+模型适配，2026 年预计放量加速	12
2.3.1 GPGPU 路径 vs ASIC 路径	12
2.3.2 竞争格局：海外厂商占据主要份额，2025 年壁仞科技市占率低	13
2.3.3 国产厂商：产品迭代、算力升级、模型适配，国产芯片商业化奇点已至	14
3、公司竞争力：高性能芯片集群构建技术壁垒，芯片迭代推动未来销售增长	15
3.1 产品：自研 GPGPU 芯片+软件平台，支持系统级解决方案出货	15
3.2 技术：GPGPU 架构+Chiplet 助力单芯片性能突出，布局 OCS 打造大规模集群互联能力	17
3.2.1 拥有 GPGPU 架构、SoC 设计、Chiplet 等核心技术	17
3.2.2 具备大规模集群互联能力，布局超节点、OCS 和 CPO 等光通信技术	20
3.2.3 公司管理团队和研发团队经验丰富	21
3.3 客户和商业化：客户集中度较高，BR106 芯片为销售主力产品	22
3.3.1 尚处商业化早期阶段，客户集中度较高，前五大客户变动频繁	22
3.3.2 销售：GPGPU BR106 为当前销售主力，公司同时出售整体智能计算解决方案	23
3.3.3 客户合作：当前以运营商、国产算力场客户为主，未来积极向互联网客户拓展	24
3.4 供应链：晶圆制造&封测由第三方提供，实体清单后采购转向国产替代	25
4、财务分析：营收快速增长，持续研发投入，公司尚处于商业化初期	26
5、盈利预测与估值	28
5.1 关键假设及盈利预测	28
5.1.1 营收假设	28
5.1.2 毛利率预测	29
5.1.3 费用预测	30
5.2 相对估值	31
5.3 绝对估值	31
5.3.1 关于基本假设的几点说明：	31
5.4 估值结论与投资评级	32
6、风险分析	33

图目录

图 1: 壁仞科技发展历程.....	6
图 2: 壁仞科技具有 GPGPU 芯片及基于 GPGPU 的智能计算解决方案	7
图 3: 壁仞科技采取「1+1+N+X」平台战略.....	7
图 4: 壁仞科技股本结构（全球发售完成后，考虑发售量调整权和超额配股权后）	8
图 5: 壁仞科技股权结构（全球发售完成后，考虑发售量调整权和超额配股权后）	9
图 6: 2020 年至 2029 年（估计）全球智能芯片行业市场规模	10
图 7: 2020 年至 2029 年（估计）中国智能芯片行业市场规模	10
图 8: 2026 年初，中国大模型每周 Tokens 调用量首次超过美国	12
图 9: OpenRouter 上榜中国大模型 Tokens 调用量趋势	12
图 10: GPGPU 架构(壁仞)和 ASIC 架构(谷歌 TPU)示意图对比	13
图 11: 中国国产智能计算芯片厂商的产品迭代路线图	14
图 12: 壁仞科技产品路线图	15
图 13: 壁仞科技主要芯片产品（BR106 和 BR166）的型号介绍.....	16
图 14: 公司 BIRENSUPA 软件平台采用四层分层架构	17
图 15: 公司智能计算解决方案的开发流程	17
图 16: 壁仞科技 GPGPU 架构设计图	18
图 17: 壁仞科技的单芯片性能在国产芯片里表现优异	19
图 18: 壁仞芯片采取 CoWoS 封装+Die to Die 高速互联方案.....	19
图 19: 壁仞科技发布光跃超节点 128 卡解决方案	20
图 20: 2025H1 公司研发开支结构	22
图 21: 壁仞科技的研发成果和研发投入	22
图 22: 公司各时期客户采购规模及客户集中度变化（左轴单位：百万元）	23
图 23: 壁仞科技合作客户涵盖运营商、智算中心、互联网、行业客户等多领域.....	24
图 24: 公司采购规模及供应商集中度变化（单位：百万元）	25
图 25: BIS 清单事件下关键上游供给影响及应对路径.....	26
图 26: 营收逐年增长，2025 营收同比增速 207.2%（单位：千元）	26
图 27: 销售产品业务持续贡献公司绝大部分营收	26
图 28: 期间费用结构以研发投入为主（单位：千元）	27
图 29: 总毛利率走势与产品销售毛利率保持一致	27
图 30: 公司年内亏损与经调整净亏损对比（单位：千元）	28
图 31: 公司经营活动所用现金净额持续为负（单位：千元）	28

表目录

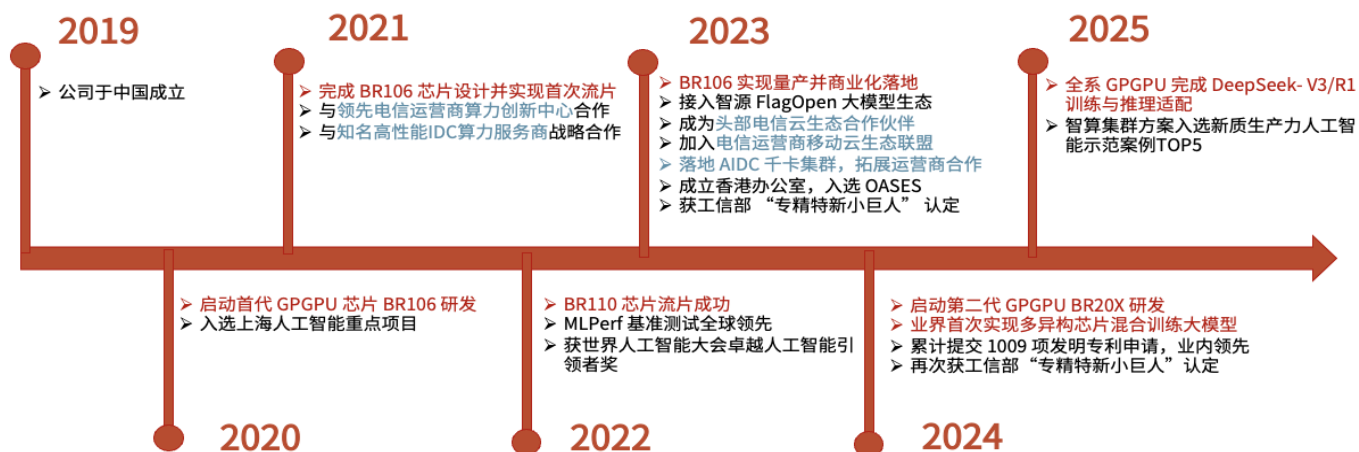
表 1: 壁仞科技融资历程.....	7
表 2: 壁仞科技全球发售的募资所得款项用途.....	9
表 3: 壁仞科技全球发售的募集资金具体计划开支 (单位: 百万港元)	10
表 4: 国内云厂商资本开支远期计划	11
表 5: GPGPU 和 ASIC 架构的性能优劣势对比汇总表.....	12
表 6: 2024 年中国智能计算芯片行业的竞争格局	13
表 7: 2024 年中国 GPGPU 行业的竞争格局	13
表 8: 国产芯片厂商积极适配国产大模型, 软硬结合提升性能竞争力	14
表 9: 壁仞科技芯片产品定位及优势	15
表 10: 壁仞科技 GPGPU 架构与 SoC 设计中的核心技术和创新功能	18
表 11: 壁仞科技管理团队介绍.....	21
表 12: 公司各时期客户数量变化及交易情况.....	22
表 13: GPGPU 芯片销售量 (单位: 颗)	23
表 14: 公司智能计算解决方案的标杆项目「南京 1024-GPU 集群」项目细节	23
表 15: 壁仞科技与是运营商客户重要战略合作的国产芯片厂商	24
表 16: 壁仞科技与运营商客户的具体合作项目情况整理	24
表 17: 壁仞科技与国家训练场的具体合作项目情况整理	25
表 18: 壁仞科技营收预测.....	29
表 19: 壁仞科技与可比公司毛利率对比表	30
表 20: 壁仞科技盈利预测表	31
表 21: 壁仞科技与可比公司的盈利预测和估值比较.....	31
表 22: 绝对估值核心假设表	32
表 23: 现金流折现及估值表	32

1、公司概况：港股 GPGPU 第一股

1.1 发展历程：港股 GPU 第一股，2026 年 1 月正式港股上市

国内领先通用智能计算解决方案提供商，2025 年 12 月通过港交所聆讯，2026 年 1 月正式港股上市。壁仞科技自 2019 年成立以来，专注于开发 GPGPU 芯片及基于 GPGPU 的智能计算解决方案，以提供 AI 所需的基础算力。公司持续研发迭代多款 GPGPU 芯片，并于 2021 年起，公司同电信运营商在算力领域开启战略合作，后续持续进行芯片出货、数据中心部署等商业化落地。2026 年 1 月 2 日，壁仞科技作为“港股 GPU 第一股”成功上市。

图 1：壁仞科技发展历程



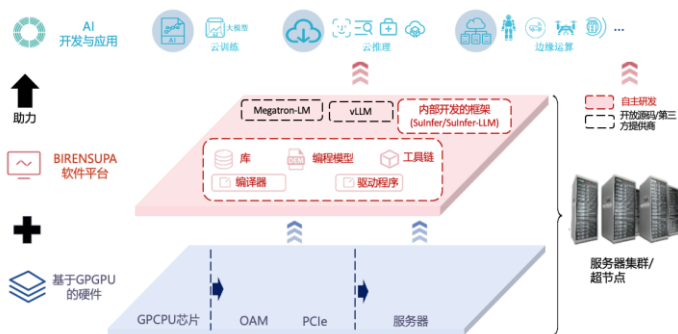
资料来源：公司招股说明书，光大证券研究所整理绘制；标红部分为公司重要技术/产品突破，标蓝部分为公司重要客户/商业化进展。

1.2 业务：GPGPU 芯片+BIRENSUPA 软件平台，构成智能计算解决方案

壁仞科技自主研发出了智能计算整体解决方案，涵盖 1) GPGPU 架构与芯片的硬件系统；2) BIRENSUPA 计算软件平台。解决方案可以通过大规模智能计算集群的形式交付，大量互联 GPGPU 单元在 BIRENSUPA 软件平台的调度下，支持从云端到边缘的广泛应用中 AI 模型的训练及推理。

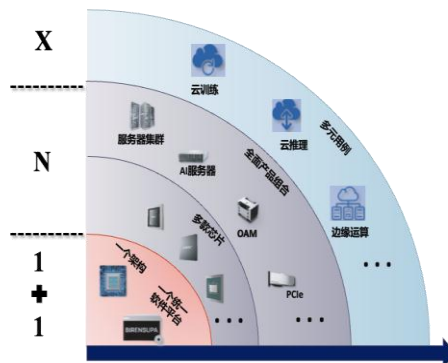
壁仞科技实施高效的「1+1+N+X」平台战略，即「1」个 GPU 架构+「1」个统一软件平台，衍生出「多款」芯片、「全面」产品组合，以赋能「X」个广泛的应用场景。统一的硬件架构及软件平台，为用户提供跨产品的无缝兼容性，降低使用门槛。

图 2: 壁仞科技具有 GPGPU 芯片及基于 GPGPU 的智能计算解决方案



资料来源: 公司招股说明书, 光大证券研究所

图 3: 壁仞科技采取「1+1+N+X」平台战略



资料来源: 公司招股说明书, 光大证券研究所

1.3 融资情况: IPO 以前经历 10 轮融资, 最新投后估值约为 209 亿元

截至 IPO 前, 公司已经历 10 轮融资。2020 年 4 月起、截至 IPO 之前, 公司已经历 10 轮融资。IPO 前的最近一轮一级投资, 于 2025 年 8 月完成全部对价支付和相关工商登记变更, 投资方包括民生通惠资产管理有限公司、图灵安昌等多家机构, 投后估值达 209.15 亿元人民币。

表 1: 壁仞科技融资历程

序号	轮次	认购或收购的 注册资本或股份	每股成本	筹资金额	隐含投资后估值	完成日期	投资方
1	Pre-A 轮	1,872,147 元	10.95 美元	20,500,000 美元	130,000,000 美元	2020/4/30	QM120、杭州独角兽
2	Pre-A+轮	791,476 元	12.63 美元	10,000,000 美元	160,000,000 美元	2020/3/11	Champ Earn
3	Pre-A++轮	1,107,429 元	16.50 美元	18,271,856 美元	243,271,856 美元	2020/3/27	澄丰、宁波梅山星胤峰
4	A 轮	5,188,534 元	152.49 元	791,200,000 元	2,891,200,000 元	2021/1/29	天津宇珩、苏州耀途、苏州华映六期投资合伙企业(有限合伙)、青岛华芯锚点、深圳松禾、珠海格力、珠海大横琴、中通瑞德、苏州聚源
5	Pre-B 轮	4,436,069 元	205.7 元	912,500,000 元	4,812,500,000 元	2021/4/29	Sky9 Alpha、盐城治平、共青城凤球、上海金浦、聚隆景润、高榕康腾、高榕康永、Champion Forest、深圳前海、元启立千及嘉兴誉峰
6	Pre-B+轮	4,201,712 元	259.02 元	1,088,327,500 元	7,148,327,500 元	2021/10/8	湖州景鑫、宁波梅山汇萃、中信证券投资有限公司、BAI GmbH、汇碧二号、碧桂园创投、苏州源启、Matrice Capital、上海国企改革基金
7	B 轮	4,491,717 元	344.24 元	1,546,228,250 元	11,046,228,250 元	2021/10/22	MSA Growth、Cyber Chief、Praise Fortune、Lobelia、南通江海基金、GBA Fund、智汇独角兽、PA GCC、RCIF、嘉芯智造
8	B+轮	827,296 元	398.89 元	330,000,000 元	13,130,000,000 元	2023/4/28	中小企业华映基金、北京裕润、共青城升和
9	战略轮	5,443,125 股	440.51 元	2,397,750,230 元	16,897,750,230 元	2025/6/20	上海临科壁芯私募基金合伙企业(有限合伙)、广州产投重大项目投资专项基金合伙企业、知识城(广州)产业园发展集团有限公司
10	2025 年 8 月	193,309,850 股	495.32 元	1,914,983,800 元	20,914,983,800 元	2025/8/22	民生通惠资产管理有限公司, 图灵安昌

资料来源: 公司招股说明书, 光大证券研究所

1.4 股本结构：全球发售 H 股占比 13.4%，创始人张先生及其一致行动人为控股股东

壁仞科技 2026 年 1 月 2 日上市时，全球发售 2.85 亿股 H 股股份（不含超额配股），12.38 亿股非上市股份和 8.73 亿股转换成 H 股的非上市股份，外加 2026 年 2 月 2 日超额配股 0.43 亿股，总股本达 24.39 亿股：

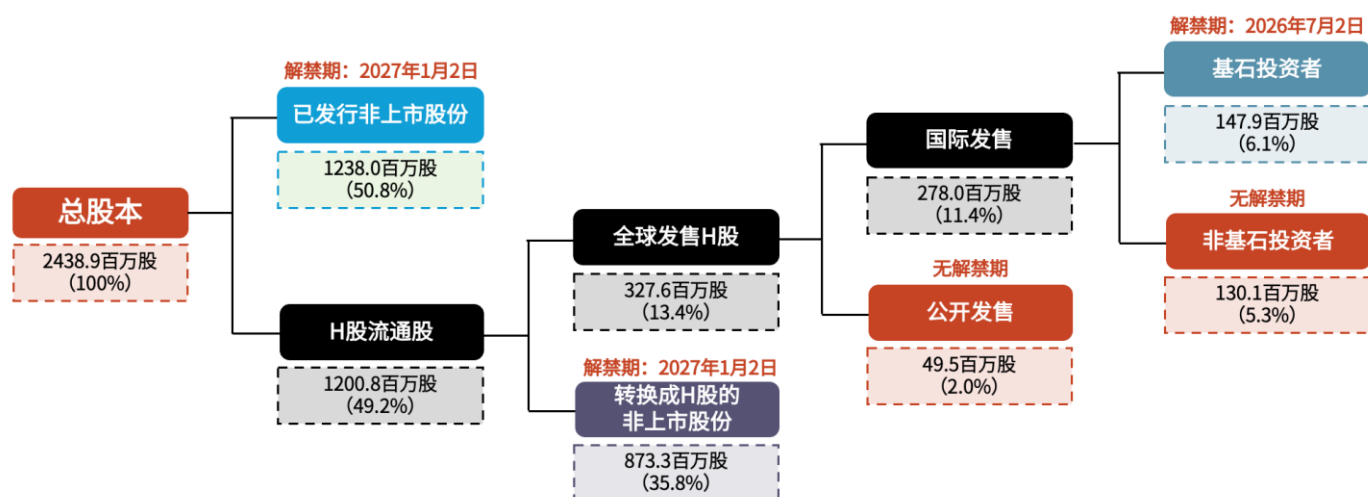
全球发售：其中国际发售（面向机构投资者或专业投资者）2.78 亿股，公开发售（面向个人投资者）0.50 亿股。国际发售中，向基石投资者发售 1.48 亿股，占全球发售的 45.1%。

IPO 发售后，H 股流通股为 12.01 亿股，占总股本的 49.2%，但自由流通 H 股（不受禁售限制）共 1.80 亿股，占总股本的 7.4%。

2026 年 7 月 2 日，全球发售中的基石投资者部分解禁，新增解禁 1.48 亿股，占比 6.1%。该次解禁完成后，自由流通 H 股共 3.28 亿股，占比 13.4%。

2027 年 1 月 2 日，新增解禁 8.73 亿股，占比 35.8%。该次解禁完成后，自由流通 H 股共 12.01 亿股，占比 49.2%（假设期间非上市股份没有新申请转换成 H 股）。

图 4：壁仞科技股本结构（全球发售完成后，考虑发售量调整权和超额配股权后）



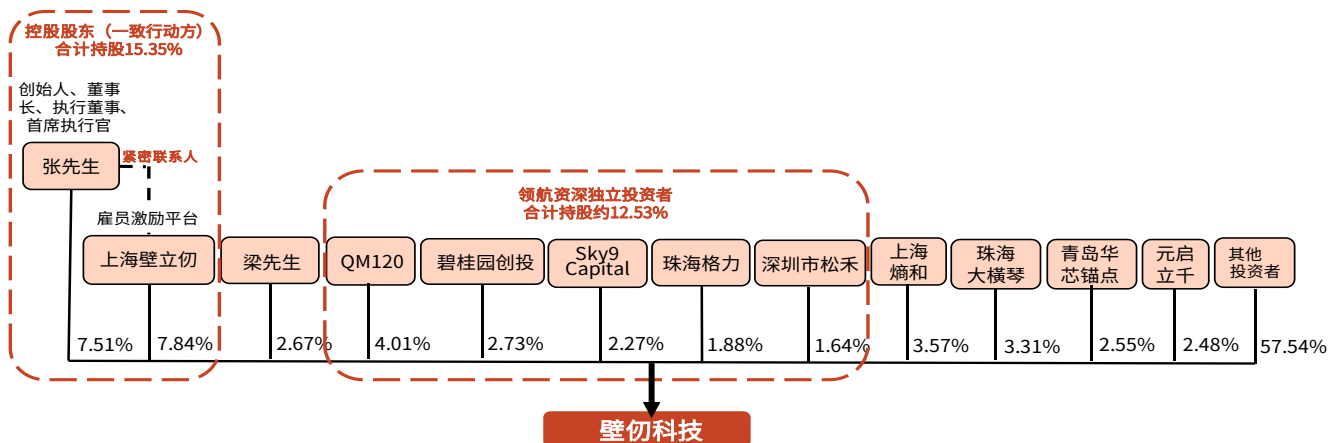
资料来源：公司招股说明书，公司公告，ifind，wind，光大证券研究所整理绘制，股权结构数据截至 2025-12-31；绿色文本框代表非 H 股流通股（虽 27 年 1 月 2 日解禁，但先需向港交所申请转换成 H 股）；红色文本框代表无禁售要求的 H 股自由流通股；浅蓝色文本框代表禁售期 6 个月的 H 股流通股；深蓝色文本框代表禁售期 12 个月的 H 股流通股。

1.5 股权结构：创始人张先生及其一致行动人为控股股东，合计持股 15.4%

创始人张先生及其一致行动人为控股股东。创始人、董事长、执行董事兼首席执行官张先生持股 7.51%，且与上海壁立仞已订立一致行动协议，因此公司单一最大股东集团包括张先生、上海壁立仞（本公司的雇员激励平台）及上海卓仞（上海壁立仞的普通合伙人），合计持股 15.35%。

公司领航资深独立投资者涵盖 5 家机构。领航资深独立投资者包含 QM120、碧桂园创投、Sky9 Capital、珠海格力、深圳松禾等 5 家机构，考虑 QM120 同时为本次全球发售的基石投资人，目前公司领航资深独立投资者合计持股 12.53%。

图 5: 壁仞科技股权结构 (全球发售完成后, 考虑发售量调整权和超额配股权后)



资料来源: 公司招股说明书, 公司公告, 光大证券研究所整理绘制。QM120 同时作为现有股东和基石投资人, 持股比例已累加计算。

1.6 募资用途: 聚焦智能计算解决方案研发迭代, 兼顾商业化与营运支持

实际全球发售募资额所得款项总额为 55.83 亿港元, 在超额配售后所得款项总额 62.04 亿港元。2026 年 1 月 2 日, 公司实际按照 19.6 港元作为最终发售价发行, 发售 2.85 亿 H 股, 所得款项总额为 55.83 亿港元。2026 年 2 月 2 日, 行使超额配股权发行 0.43 亿普通股, 扣除全球发售应付的包销费用及佣金以及开支, 公司全球发售所得款项净额为 62.04 亿港元 (约当 56.31 亿人民币)。

壁仞科技拟定全球发售的募资额将 85%用于研发和迭代公司的智能计算解决方案。根据招股说明书, 公司拟将本次全球发售所得款项净额约 62.04 亿港元主要用于智能计算解决方案研发, 包括 GPGPU 芯片及相关硬件的持续迭代、以及软件平台的开发与升级, 同时适度投入于解决方案商业化及营运资金, 以支持市场拓展、生态建设及公司日常经营。

表 2: 壁仞科技全球发售的募资所得款项用途

序号	一级用途	二级用途	三级用途	具体用途
1	研发智能计算解决方案 (约 85.0%)	智能计算硬件发展 (约 45.0%)	开发及升级 GPGPU 芯片 (约 32.0%)	研发及工程人才队伍建设 (约 17.0%) 工程设计、验证和流片服务 (约 15.0%)
			开发 GPGPU 硬件平台 (约 13.0%)	购买耗材 (约 10.0%) 研发及工程人才队伍建设 (约 3.0%)
		软件平台开发及升级 (约 40.0%)	扩展智能硬件及 BIRENSUPA 的模型训练与推理支持 (约 10.0%)	研发及工程人才队伍建设 (约 7.0%) 租赁及购买训练或推理服务器 (约 3.0%)
			优化软件平台 (约 20.0%)	研发及工程人才队伍建设 (约 14.0%) 租赁及购买训练或推理服务器 (约 6.0%)
			建设自有软件开发基础设施 (约 10.0%)	研发及工程人才队伍建设 (约 7.0%) 租赁及购买训练或推理服务器 (约 3.0%)
2		智能计算解决方案的商业化 (约 5.0%)		
3		营运资金及一般公司用途 (约 10.0%)		

资料来源: 公司招股说明书, 光大证券研究所

表 3：壁仞科技全球发售的募集资金具体计划开支（单位：百万港元）

项目	2026 年	2027 年	2028 年	2029 年	总计
智能计算解决方案研发	778.8	793.6	1076.8	1048.9	3698.0
—— 智能计算硬件的发展	430.7	445.5	554.7	526.9	1957.8
—— 软件平台的开发及升级	348.0	348.0	522.1	522.1	1740.2
智能计算解决方案的商业化	42.4	54.4	65.3	55.5	217.5
营运资金及其他一般公司用途	87.0	87.0	130.5	130.5	435.1
总计	908.2	934.9	1272.6	1234.9	4350.6

资料来源：公司招股说明书，光大证券研究所。注：此募集资金具体计划开支是招股说明书基于 18.3 港元/股且发售量调整权和超额配股权均未行使的情况下假设的，因此计算的全球发售所得款项净额为 43.51 亿港元元，公司后续计划开支节奏表未更新。

2、行业概况：中国智能计算芯片市场高速增长，国产芯片有望出货放量

2.1 行业空间：智能计算芯片市场高速增长，预计中国快于全球增速

全球智能芯片：根据灼识咨询的数据，全球智能计算芯片的市场规模从 2020 年的 66 亿美元快速增长至 2024 年的 1190 亿美元，对应 2020-2024 年复合年增长率（CAGR）为 106.0%。预计未来五年全球市场将保持快速增长，灼识咨询预计 2029 年全球智能计算芯片的市场规模将达到 5857 亿美元，对应 2024-2029 年 CAGR 为 37.5%。灼识咨询预计至 2030 年，AI 将推动全球国内生产总值的 15% 以上的增长。

中国智能芯片：根据灼识咨询的数据，中国智能计算芯片的市场规模从 2020 年的 17 亿美元快速增长至 2024 年的 301 亿美元，对应 2020-2024 年复合年增长率（CAGR）为 105.0%。灼识咨询预计 2029 年中国智能计算芯片的市场规模将达到 2012 亿美元，对应 2024-2029 年 CAGR 为 46.3%。未来 5 年，中国智能计算芯片的市场规模增速预计将快于全球整体市场规模增速。

图 6：2020 年至 2029 年（估计）全球智能芯片行业市场规模

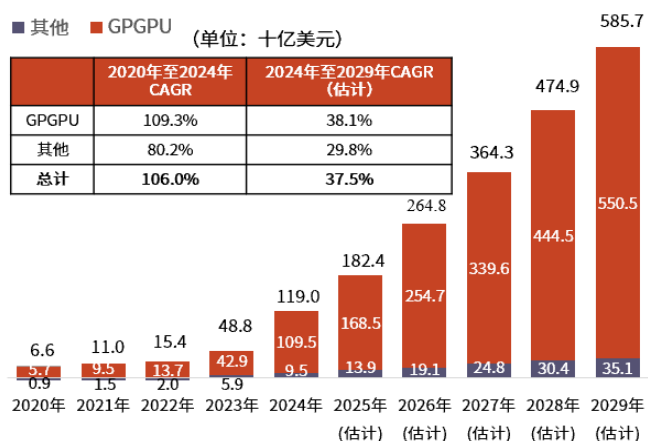
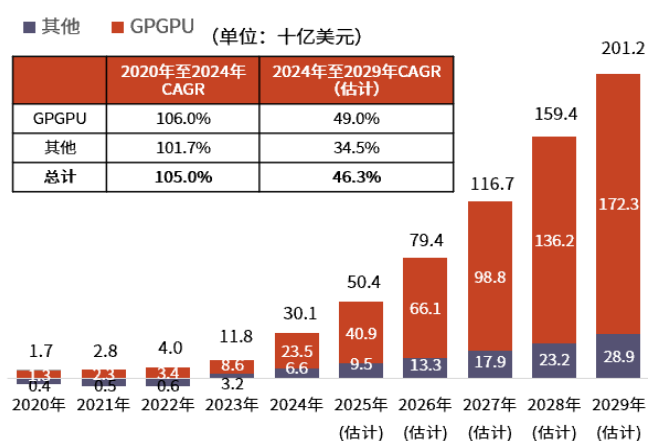


图 7：2020 年至 2029 年（估计）中国智能芯片行业市场规模



2.2 需求：国产大模型异军突起，云厂商大举投入，助推国产算力需求高增

2.2.1 国内公司积极备战 AI 算力军备竞赛，加码 CAPEX 投入 AI 基础设施建设

OpenClaw 的发布点燃 AI Agent 浪潮，推动科技革命进而带动新一轮 AI 投入。OpenClaw 全时在线、插件叠加、高频交互的特性，有望持续推高算力与云资源算力需求，带动云厂商 Token 消耗和涨价浪潮，云服务厂商在量增+价涨+结构升级的三重红利下，有望进一步加大 AI 基础设施的资本支出：

- 1) 7×24 小时稳定运行需求凸显，云端部署成为主流选择：OpenClaw 作为有状态的守护进程，需持续监听飞书、钉钉等多平台消息接口，本地部署受限于设备稳定性、性能负载，且关机即中断服务，云端部署成为个人开发者与企业主流选择；同时，OpenClaw 部署相对复杂，各大云厂商推出一键镜像/上门安装等服务，降低企业与开发者部署门槛，加速上云工作负载落地。
- 2) 插件叠加+多任务并行，驱动 IDC 订单提升和云服务升级：OpenClaw 完成单一任务需与底层大模型进行数十次至数百次 API 交互，Token 消耗与算力需求较传统对话式 AI 显著提升。这种全天候、高并发的算力消耗模式，对数据中心的稳定性和规模提出了更高要求。
- 3) OpenClaw 新增 Token 消耗需求，带动云服务厂商提价潮：2026 年全球云服务市场打破长期降价传统，海外方面，亚马逊 AWS 于 1 月上调 EC2 机器学习实例价格约 15%，谷歌云宣布 5 月起全球数据传输服务涨价，北美地区费率翻倍。

表 4：国内云厂商资本开支远期计划

公司	时间	远期计划指标
阿里巴巴	2026 年	Capex 计划约 1500 亿元，其中 AI 基础设施投资超 1000 亿元（海外 60%，国内 40%）。
	未来三年	计划投入 3800 亿元用于云与 AI 基础设施建设，超过过去十年总投资。
	未来五年	吴泳铭在 2025 云栖大会主旨演讲称，“最近一年全球 AI 行业投资总额已超 4000 亿美元，未来 5 年全球 AI 累计投入将超 4 万亿美元”，同时提及“AI 的行业渗透速度将超过历史上所有技术”。
	2032 年远期	远期规划到 2032 年阿里云全球数据中心能耗规模较 2022 年提升 10 倍。
百度	2026 年	Capex 计划约 400-500 亿元，重点投向昆仑芯、大模型训练、智算中心及自动驾驶算力基础设施。
华为	2030 年远期	预计 2030 年中国 AI Token 带来的网络流量将达约 500TB/日，显著高于当前约 90TB/日，推动网络与算力基础设施升级。
腾讯	2026 年	2026 年预计还会增加资本开支。
字节跳动	2026 年	Capex 计划约 1600 亿元，其中 850 亿元用于 AI 芯片采购、750 亿元用于数据中心建设，重点投向全球算力集群、豆包大模型及海外算力节点扩张。

资料来源：华尔街见闻，新浪财经，证券时报网，东方财富网，澎湃新闻，财联社，界面新闻，光大证券研究所整理。

2.2.2 国内大模型调用量大幅增长，Tokens 出海开启

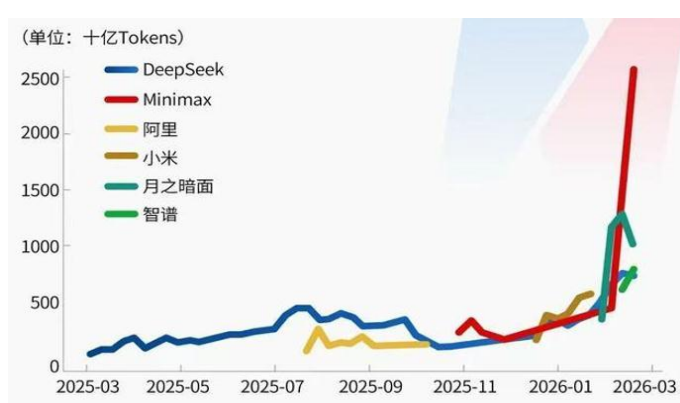
国产 AI 大模型异军突起，2026 年 2 月~3 月 Tokens 调用量和增速均大超美国。根据全球最大的 AI 模型 API 聚合平台 OpenRouter 显示，2025 年，美国模型 Tokens 调用量是市场增长主要动力，但进入 2026 年，美国大模型调用量增速显露疲态，而中国大模型调用量大幅赶超，2 月 9 日至 15 日当周，中国模型便以 4.12 万亿 Token 的惊人调用量，正式超越同期美国模型的 2.94 万亿 Token，实现历史性赶超；2 月 16 日~22 日当周，中国模型的周调用量更是冲高至 5.16 万亿 Token，三周时间调用量增长 127%，将领先优势进一步扩大。2026 年 2 月 16 日至 22 日的周榜单显示，平台调用量排名前五的模型中，有四款来自中国厂商，分别为 MiniMax 的 M2.5、月之暗面的 Kimi K2.5、智谱的 GLM-5 以及 DeepSeek 的 V3.2。

图 8: 2026 年初, 中国大模型每周 Tokens 调用量首次超过美国



资料来源: 每日经济新闻; 数据截至 2026-02-22

图 9: OpenRouter 上榜中国大模型 Tokens 调用量趋势



资料来源: 每日经济新闻; 数据截至 2026-02-22

2.3 竞争格局: 国产芯片积极产品迭代+性能升级+模型适配, 2026 年预计放量加速

2.3.1 GPGPU 路径 vs ASIC 路径

GPGPU (General-Purpose Graphics Processing Unit, 通用图形处理器) 是指将 GPU 用于通用计算的技术和方法。GPGPU 基于非定制架构, 多用 GPGPU 处理计算密集型工作负载, 比如大型模型训练和推理, 可有效满足不同 AI 模型架构和复杂通用任务, 具备高通用性和灵活性。目前 GPGPU 已成为智能计算芯片的主流选择, 公司预计将保持领先地位。

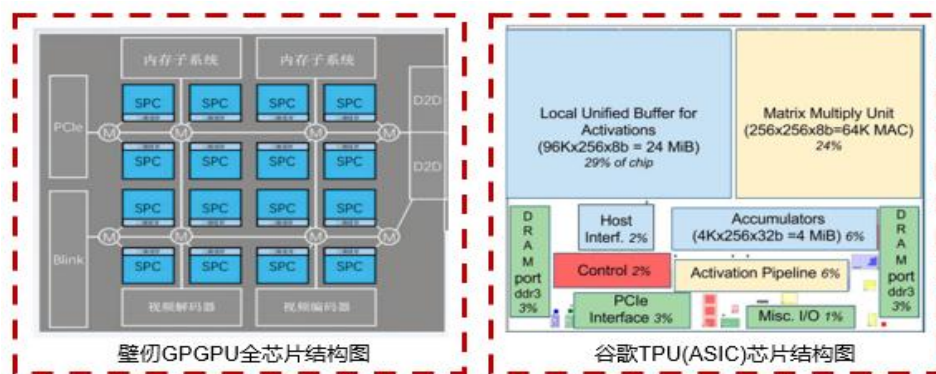
ASIC (Application-Specific Integrated Circuit, 专用集成电路) 是一种专门为特定应用或特定功能而设计和定制集成电路。ASIC 针对特定任务, 在成本、能效方面得到有针对性的优化。

表 5: GPGPU 和 ASIC 架构的性能优劣势对比汇总表

	GPGPU	ASIC
架构	建立在 GPU 架构之上, 通过软件重新利用这些并行核心, 强调可编程性。	ASIC 具有简化的架构, 仅保留其预期任务的基本要素。
性能	通用各种任务, 适用高度并行任务	为特定任务提供全面优化的性能
成本	中等	高开发成本, 量产边际成本会大幅下降
寿命	较长	开发周期长且容易过时
灵活性	高, 拥有完善的软件生态系统	低, 应用不灵活
效率	中等	高
能耗	高	低
应用场景	模型训练推理、金融建模等	边缘 AI 推理等
代表公司	沐曦、燧原、壁仞、摩尔线程、英伟达	华为、寒武纪

资料来源: 公司招股说明书, AI-stack, 光大证券研究所整理

图 10: GPGPU 架构(壁仞)和 ASIC 架构(谷歌 TPU)示意图对比



资料来源: 公司招股说明书, 51CTO, 光大证券研究所

2.3.2 竞争格局: 海外厂商占据主要份额, 2025 年壁仞科技市占率低

中国智能计算芯片和中国 GPGPU 芯片市场均呈现头部高度集中的竞争格局。

中国智能计算芯片市场竞争格局: 根据灼识咨询的数据, 按收入口径统计, 中国智能计算芯片的前两大厂商合计占据 94.4% 的市场份额, 其中第一大厂商为美国厂商, 2024 年市场份额为 76.2%; 其余厂商市场份额相对分散, 超过 15 家规模化参与者均未占据超过 1.0% 的市场份额。

中国 GPGPU 芯片市场竞争格局: 根据灼识咨询的数据, 按收入口径统计, 中国 GPGPU 市场共有不足十家规模化参与者, 前两大参与者合计占据 98% 的市场份额, 其中龙头厂商同时是中国智能计算芯片的第一大参与厂商, 占中国 GPGPU 市场份额的 97.6%。

此前中国智能计算芯片和中国 GPGPU 芯片的市场份额主要由海外厂商占据, 中国厂商将凭借更强竞争力不断提升市场份额。

壁仞 2024 年和 2025 年市场份额较低。根据灼识咨询的数据, 壁仞于 2024 年在中国智能计算芯片和中国 GPGPU 市场分别占据 0.16% 和 0.20% 的市场份额; 公司预计 2025 年在中国智能计算芯片和中国 GPGPU 市场分别占据 0.19% 和 0.23% 的市场份额。

随着中国企业技术迭代和竞争力提升, 公司预计国产智能计算芯片占据的中国市场份额将从 2024 年的约 20% 提升至 2029 年的约 60%。

表 6: 2024 年中国智能计算芯片行业的竞争格局

排名	公司	所在国家	成立时间	上市情况	智能计算芯片主要类型	市场份额 (以收入计)
1	公司 A	美国	1993 年	美股纳斯达克上市	GPGPU	76.2%
2	公司 B	中国	2004 年	非上市	ASIC	18.2%

资料来源: 公司招股说明书援引灼识咨询, 光大证券研究所

表 7: 2024 年中国 GPGPU 行业的竞争格局

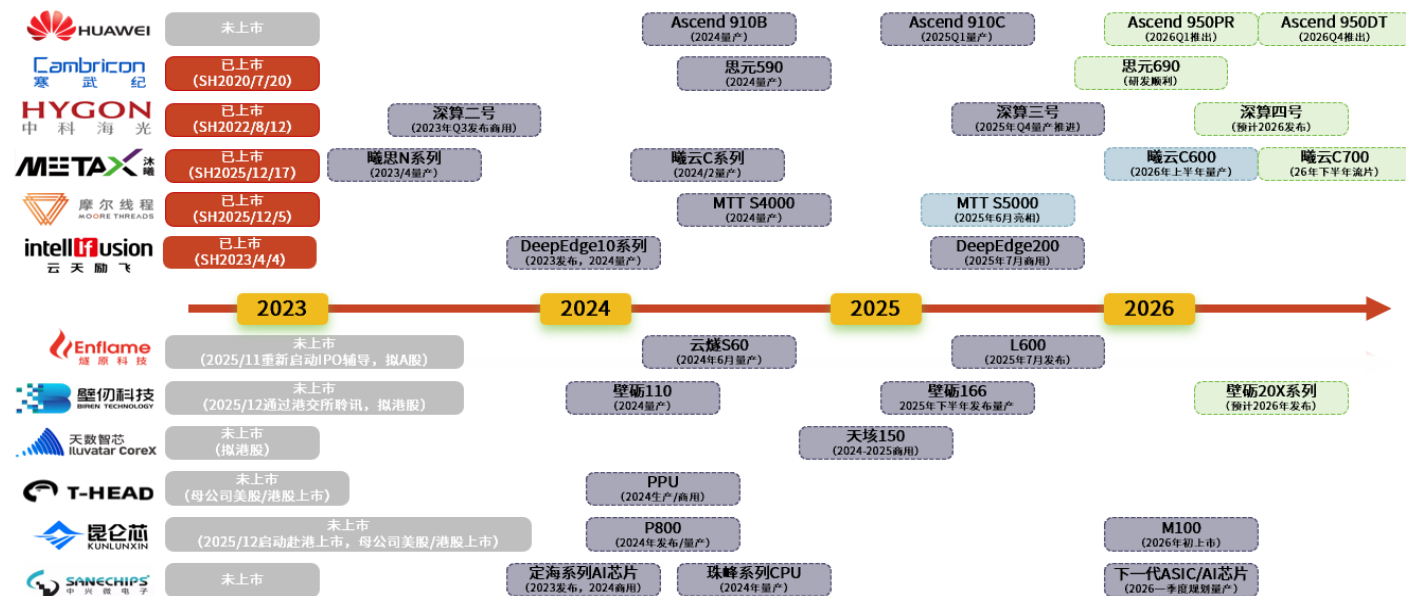
排名	公司	所在国家	成立时间	上市情况	市场份额 (以收入计)
1	公司 A	美国	1993 年	美股纳斯达克上市	97.6%
2	公司 C	美国	1969 年	美股纳斯达克上市	0.4%

资料来源: 公司招股说明书援引灼识咨询, 光大证券研究所

2.3.3 国产厂商：产品迭代、算力升级、模型适配，国产芯片商业化奇点已至

中国 AI 芯片厂商积极推进上市流程，并加快产品迭代进度和商业化进展。2026 年，期待华为昇腾 950、寒武纪思元 690、海光深算三号/四号等一些列新芯片发布和量产。

图 11：中国国产智能计算芯片厂商的产品迭代路线图



资料来源：华为官网，海光信息官网，沐曦官网，摩尔线程官网，同花顺财经，证券时报网，财联社，路透社，36 氪，芯语，财华智库网，证券日报网，财经头条，科创板日报，南都视频，香港经济日报，时代周报，光大证券研究所整理绘制

国产芯片厂商通过适配大模型，软硬结合增强性能表现。国产 AI 模型公司与国产 AI 芯片企业正紧密合作。以华为昇腾、寒武纪为代表的国产芯片厂商，将其硬件设计与国内主流大模型的特性深度绑定，耦合发展、不断提升计算效率。壁仞科技已实现对智谱 GLM5 大模型的适配工作。

表 8：国产芯片厂商积极适配国产大模型，软硬结合提升性能竞争力

厂商	大模型	发布时间	适配的国产 AI 芯片厂商	适配情况说明
DeepSeek	DeepSeek-V4	2026 年 3 月 1 日	昇腾、寒武纪	深度支持国产芯片，国产化算力利用率 85%
	DeepSeek-V3.2-Exp	2025 年 9 月 29 日	华为昇腾	day0 完成适配。基于 vLLM-Ascend / SGLang 推理框架，使用 CANN 软件栈，开源推理代码与算子实现
			寒武纪	day0 完成适配。发布 vLLM-MLU 推理引擎，结合模型的 Sparse Attention 优化。
智谱	GLM-5	2026 年 2 月 11 日	华为昇腾、摩尔线程、海光、寒武纪、昆仑芯、沐曦以及燧原	多款国产芯片的 day 0 全栈适配
			壁仞科技	依托壁仞™166 系列产品的高通用性与全自研 BIRENSUPA™ 软件栈的成熟度，壁仞科技极速完成对 GLM-5 的适配
	GLM-4.6	2025 年 9 月 30 日	寒武纪	首个“模型-芯片一体化 FP8+Int4 方案”，在保持模型精度不变的情况下显著降低推理成本
阿里通义千问	Qwen 3.5	2026 年 2 月 16 日	昇腾、摩尔线程、沐曦、海光信息	day0 模型发布时宣布支持
			太初元基	在其 T100 加速卡上完成 Qwen 3.5-397B-A17B 的深度适配

资料来源：壁仞科技官网，华为计算，寒武纪开发者，海光信息官方微信公众号，沐曦股份官方微信公众号，摩尔线程开发者，GLM-5 官方论文，Deepseek 官方文档，量子位，观察者网，光大证券研究所整理

3、公司竞争力：高性能芯片集群构建技术壁垒，芯片迭代推动未来销售增长

3.1 产品：自研 GPGPU 芯片+软件平台，支持系统级解决方案出货

基于 GPGPU 的硬件系统：公司围绕自研 GPGPU 芯片作为核心搭建产品体系。

- 1) 芯片：目前有 BR106、BR110 与 BR166 三款量产芯片，同时正在开发下一代 BR20X 系列，以及规划后续 BR30X 与 BR31X 系列。
- 2) 其他硬件：以 GPGPU 芯片为核心，配套其他关键硬件组件，最终主要落地为 PCIe 板卡与 OAM 等。

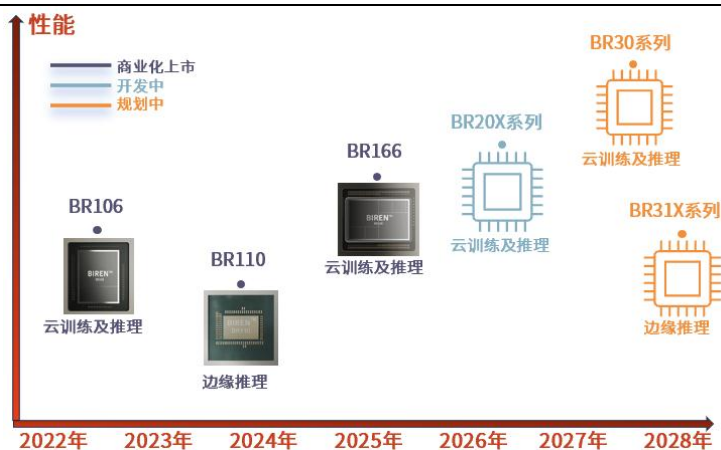
公司根据客户要求，提供具有不同配置的 PCIe 板卡、OAM、GPGPU 服务器或服务器集群。硬件系统支持风冷与液冷解决方案，能够有效降低数据中心的整体能耗。

表 9：壁仞科技芯片产品定位及优势

定位	芯片	优势
云训练及推理	BR106	专为大规模计算而设计的 GPGPU 芯片，为 AI 训练和推理应用提供灵活、高效、安全及多样化的计算解决方案。
	BR166	在峰值算力、内存、视频编解码、互连等方面性能是 BR106 的两倍。
	BR20X(开发中)	预计比第一代提供更强的单卡运算能力，配备更大更快的内存、更高速的互连带宽及超节点系统设计。
	BR30X(规划中)	预计提供更强算力、大内存容量、更强生态系统适应性、更好的可扩展性及更低的 TCO。预计 2028 年实现商业化上市。
边缘推理	BR110	第一代边缘及云推理芯片。它是一种经济高效、节能的解决方案，主要针对边缘推理工作负载而设计。
	BR31X(规划中)	与 BR30X 相同

资料来源：公司招股说明书，光大证券研究所

图 12：壁仞科技产品路线图



资料来源：公司招股说明书，光大证券研究所

BR106 系列芯片：用于训练和推理的 GPGPU 芯片。具体出货型号方面，①形态规格分成 OAM 和 PCIe 板卡两种，②OAM 冷却方面可分成风冷和液冷，因此壁仞科技在 BR106 方面提供四种具体的产品型号 (106L、106M、106B 和 106C)。

BR166 系列芯片：将两个 BR106 裸晶和四个 DRAM 集成到一个封装中，组成高性能 BR166，在峰值算力、内存、视频编解码、互连等方面性能是 BR106 的两倍。BR166 在①形态规格上分成 OAM 和 PCIe 板卡两种，②OAM 冷却分成风冷和液冷，共提供三种具体产品型号 (166L、166M、166C)。

图 13：壁仞科技主要芯片产品 (BR106 和 BR166) 的型号介绍

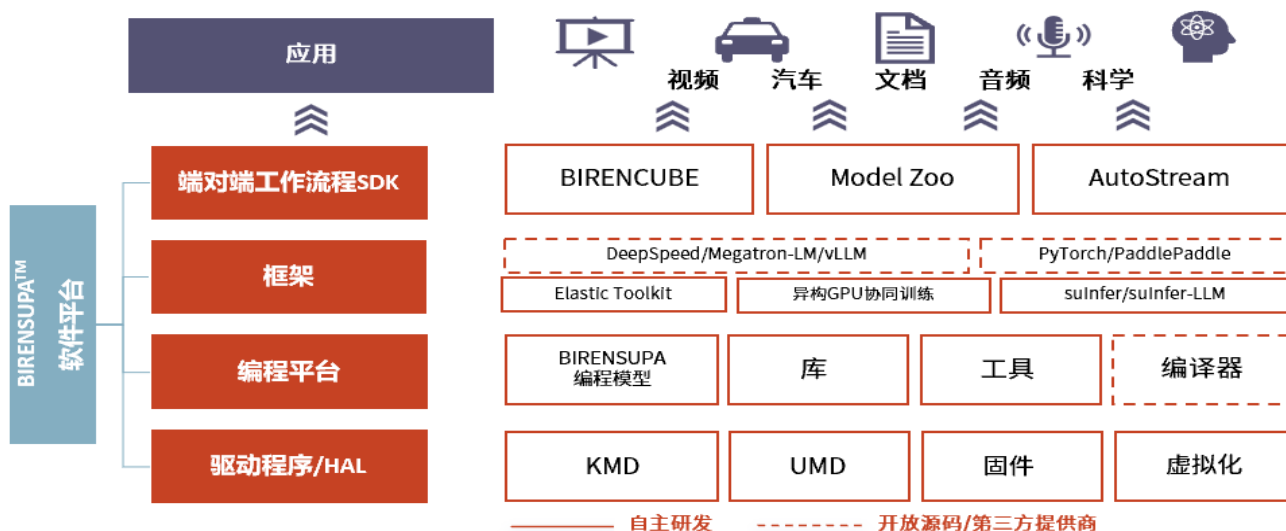
基于 GPGPU 的硬件	形态规格	编码器/解码器	双向互联宽带	散热设计功率	主要目标市场	当前状态	
BR106	壁砺 m 106L	OAM (液冷)	32通道 /256通道	256GB/s	400W	云训练及推理，尤其适用于大模型	商业化上市
	壁砺 m 106M	OAM (风冷)					
	壁砺 m 106B	PCIe板卡		192GB/s	300W	云训练及推理	商业化上市
	壁砺 m 106C	PCIe板卡		128GB/s	150W	云推理	商业化上市
BR166	壁砺 m 166L	OAM (液冷)	64通道 /512通道	576GB/s	600W	云训练及推理，尤其适用于大模型	商业化上市
	壁砺 m 166M	OAM (风冷)		576GB/s	550W		
	壁砺 m 166C	PCIe板卡		512GB/s	450W	云推理	商业化上市

资料来源：公司招股说明书，光大证券研究所

软件系统：BIRENSUPA 作为搭建于 GPGPU 之上的软件栈，是基于 GPGPU 的硬件运行所使用的自主开发软件平台。BIRENSUPA 采用四层分层架构：

- 1) 驱动程序：**能够促进高效数据传输以及多任务处理，同时确保与不同的 GPGPU 型号和操作系统的兼容性；
- 2) 编程平台：**BIRENSUPA 编程模型、编译器、库、开发者工具；
- 3) 框架：**①自研：Elastic Toolkit 用于监控大规模集群，sulInfer 和 sulInfer-LLM 提供高性能的推理引擎及服务；②第三方兼容：支持大部分主流的开源深度学习框架和开源大语言模型框架，包括 PyTorch、TensorFlow 和飞桨；
- 4) 端对端工作流程：**BIRENCUBE 支持在数千个 GPGPU 设备上任务、Model Zoo 允许客户部署预训练模型或开发自身模型。

图 14: 公司 BIRENSUPA 软件平台采用四层分层架构

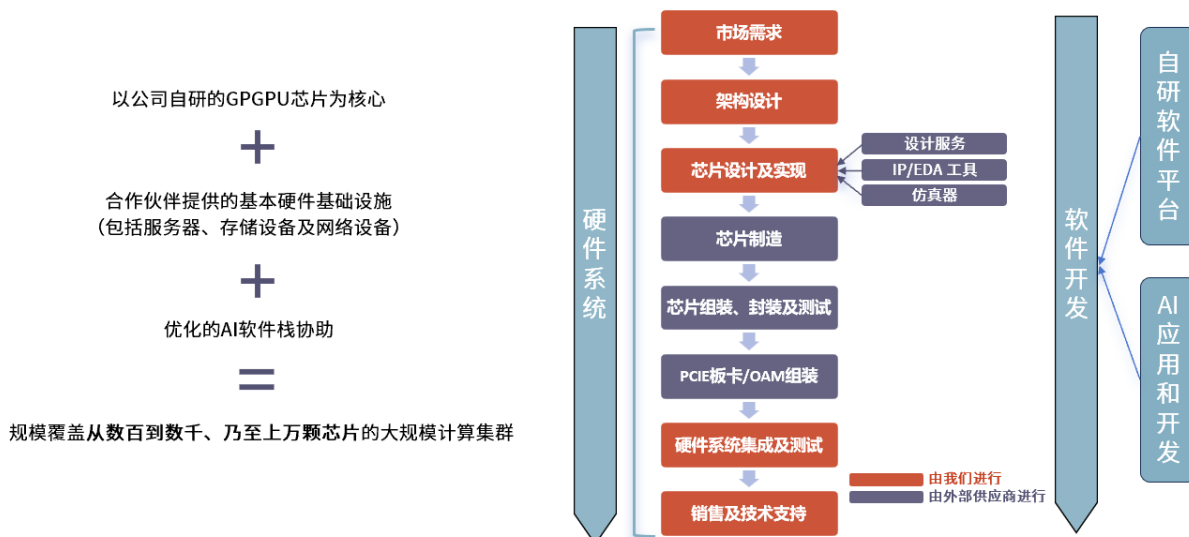


资料来源: 公司招股说明书, 光大证券研究所

公司智能计算整体解决方案主要包括:

- ①基于 GPGPU 的硬件系统, ②计算软件平台 BIRENSUPA, 支持以“大规模智能计算集群” GPU 集群/超节点 形式交付, GPU 集群/超节点由大量互联的 GPU 服务器及/或芯片组成。

图 15: 公司智能计算解决方案的开发流程



资料来源: 公司招股说明书, 光大证券研究所

3.2 技术: GPGPU 架构+Chiplet 助力单芯片性能突出, 布局 OCS 打造大规模集群互联能力

3.2.1 拥有 GPGPU 架构、SoC 设计、Chiplet 等核心技术

公司的核心技术主要覆盖 GPGPU 架构、SoC 设计、硬件系统设计与软件技术。

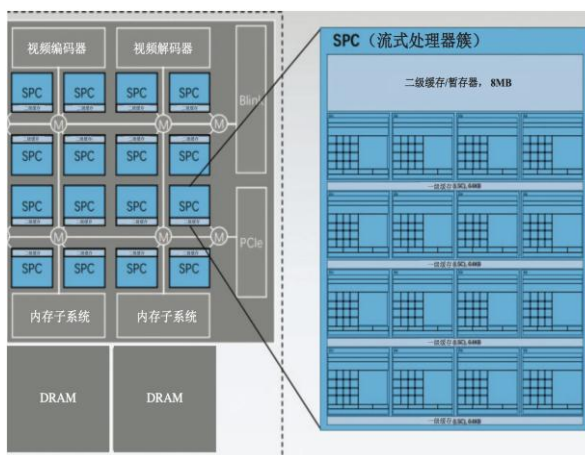
- 1) **GPGPU 架构**: 统一的 GPGPU 架构覆盖训练与推理, 实现高性能、高能效与可扩展性。
- 2) **SoC 设计**: SoC 全流程设计能力保障先进封装与超大规模芯片成功商业化落地。
- 3) **硬件系统设计**: 系统级硬件设计能力覆盖 PCIe、OAM 及服务器形态, 支持风冷与液冷部署, 面向数据中心级算力需求。
- 4) **软件技术**: BIRENSUPA 软件平台实现硬件与 AI 应用协同, 支持大规模 GPGPU 集群管理, 并兼容第三方平台以降低迁移成本。

表 10: 壁仞科技 GPGPU 架构与 SoC 设计中的核心技术和创新功能

层级	技术点	技术作用
GPGPU 架构	SIMT 架构	兼顾通用计算与 AI 加速, 适配并行计算负载
	T-core 张量引擎	支持数据循环, 减少 DRAM 访问, 提升能效
	组播异步传输	一次读取, 多核复用, 降低带宽与能耗
	近内存计算	融合 NUMA、UMA 与 L2 Reduction, 提升数据效率
SoC 设计	SoC 架构	异构 SPC 核配置, 支持云端与边缘计算
	内存系统	多通道访问、调度与预取, 提升有效带宽
	多 GPU 互连	SerDes + BLink, 实现 GPU 间高速互连
	SoC 测试	DFT 贯穿流程, 支持良率分析与降级产品
	封装设计	芯粒技术+ 先进封装, 提升集成度

资料来源: 公司招股说明书, 光大证券研究所

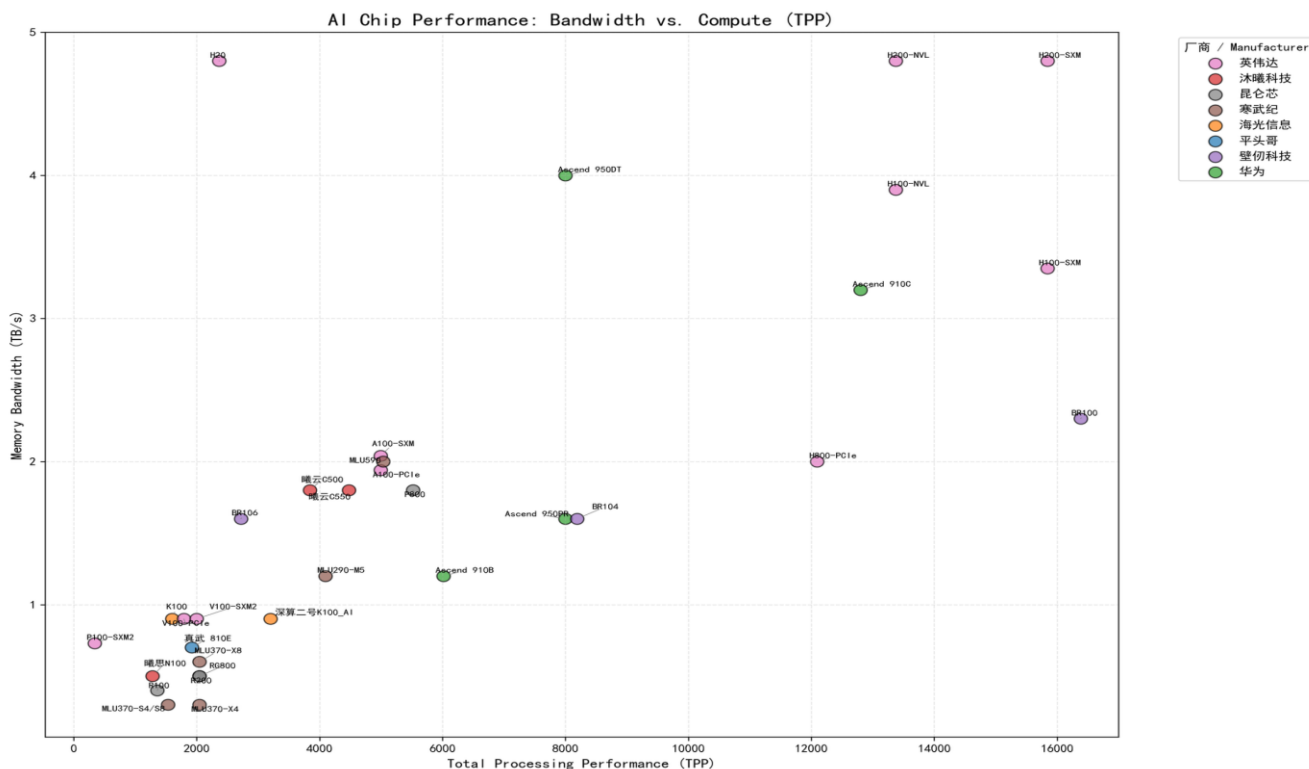
图 16: 壁仞科技 GPGPU 架构设计图



资料来源: 公司招股说明书, 光大证券研究所

壁仞定位高端 AI 芯片, 综合性能在国产芯片中表现优异。我们整理了英伟达、壁仞科技、华为、寒武纪等众多芯片厂商的近几年 AI 芯片产品, 壁仞科技的芯片在 Total Processing Performance (TPP) 方面表现优异, 位列国产芯片第一梯队, 且后续产品线将持续迭代性能。

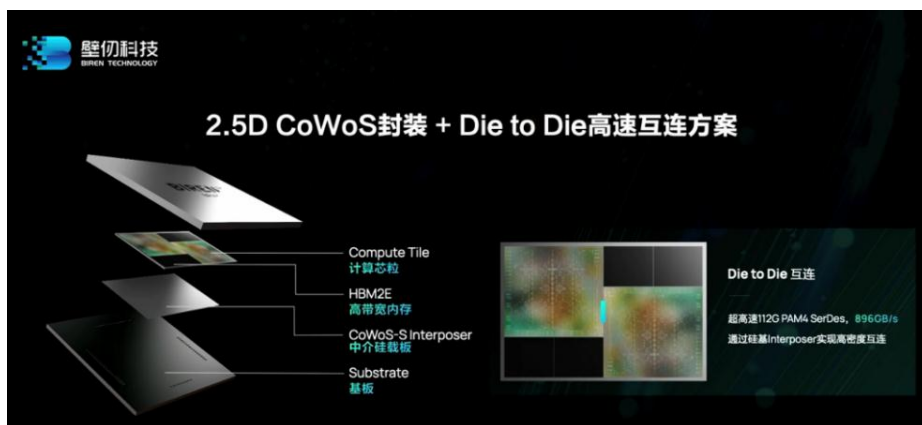
图 17: 壁仞科技的单芯片性能在国产芯片里表现优异



资料来源: 公司招股说明书, 平头哥官网, 沐曦官网, 华为技术, 华为全联接大会, 新浪财经, 界面新闻, 半导体综研, biggo, theregister, 光大证券研究所整理并绘制

领先的 Chiplet 与先进封装技术商业化落地。壁仞科技是国内最早在高性能 GPGPU 领域实现 Chiplet 技术商业化与实际量产的厂商之一。Chiplet 技术通过将多个计算裸晶 (Die) 与高速显存等组件进行 2.5D 高密度封装组合 (即“多 Die 拼接”), 能够有效突破单一芯片物理面积 (光罩尺寸) 的极限, 在大幅提升芯片良率、降低制造成本的同时, 实现算力规模的跨越式增长。早在 BR100 系列上, 壁仞科技便率先在国内验证了该技术的工程可行性; 而在其成功量产的 BR166 芯片中, 公司进一步通过芯粒技术将两颗 BR106 计算裸晶集成于同一封装内, 实现了核心性能参数的成倍跃升。

图 18: 壁仞芯片采取 CoWoS 封装+Die to Die 高速互联方案



资料来源: 壁仞科技「智·未来」新品发布会

3.2.2 具备大规模集群互联能力，布局超节点、OCS 和 CPO 等光通信技术

壁仞科技多级互联架构

1) **单机柜**：公司当前主流单机柜配置为 64-72 卡，主要依托其自研的高速互连技术（如 BR10X 使用 BLink 协议），通过 Scale-up（纵向扩展）方式实现机柜内部的高效数据交互与扩展。

2) **超节点**：面向更庞大的大模型算力需求，BR20X 系列将推出超节点方案，其基于自研互连协议 Blink2.0 进一步演进，单一超节点内部的 Scale-up 物理扩展上限可跃升至约 1024 卡，形成一个局部最优的紧耦合算力池。

该超节点架构突破了传统 8 卡 AI 服务器形态，构建起了一个大规模、紧耦合的全局互连域。面对当前业内流行的大规模 MOE（混合专家）模型，这种超大跨度的高速互连网络能够为其推理与训练提供强大的“专家并行”支持，有效消除节点间的通信瓶颈，从而大幅提升系统的整体吞吐率与并行效率。

3) **集群**：在超节点之上，多个超节点会进一步通过系统级互联，并结合光互连与分布式光交换（dOCS）等方案，Scale-out（横向扩展）成更大规模集群；具体项目方面，公司已完成 2048 卡光互联光交换 GPU 超节点智算集群在国家级算力平台的部署落地，还在中国电信实现了千卡集群的商业化落地，完成了南京电信 1024-GPU 集群交付。

技术布局：光互联技术——OCS 和 CPO

1) **OCS 应用落地**：面对未来智算集群“低延迟、高带宽、低能耗”的极致需求，壁仞科技于 2025 年实现基于光交换机（OCS）在特定场景下的应用落地：与曦智科技、中兴通讯联合打造了国内首个分布式光交换（dOCS）全光互连芯片及超节点解决方案，而后于 2026 年 3 月 12 日，正式发布光跃超节点 128 卡商用版（LightSphere 128）。仅用半年多时间即实现从概念验证到实际商用的跨越。

图 19：壁仞科技发布光跃超节点 128 卡解决方案



资料来源：壁仞科技官网

2) **探索 CPO 技术**：同步推进 CPO（Co-packaged optics，共封装光学）技术的研发与测试，计划在未来逐步取代超节点内的光模块。CPO 是将指把硅光

模块和 CMOS 芯片用高级封装的形式（例如 2.5D 或者 3D 封装）集成在一起，从而在成本、功耗和尺寸上都进一步提升数据中心应用中的光互联技术。

3) GPU 全光互连：最终技术愿景是实现从 GPU 到 GPU 的全光互连连接。此举将规避传统电交换过程中耗时费能的光电转换损耗，彻底打破电互连的物理带宽瓶颈，进一步大幅降低大规模集群的互联成本与运行功耗。

3.2.3 公司管理团队和研发团队经验丰富

公司董事会及管理团队的知识及技能组合均衡，由顶尖科学家和资深商业人士组成。

1) 技术专家：成员均在半导体、AI、GPU 设计、工程等领域数十年工作经验，分别拥有商汤集团、英伟达、华为、AMD、三星电子等多家硬件龙头厂商重要核心岗位的工作经历，且公司技术人员兼具技术创新与商业化实践经验。

2) 管理团队：成员具备战略运营、风险管理、运营销售等多领域丰富经验，为公司战略规划、运营执行提供有力支撑。

表 11：壁仞科技管理团队介绍

人员	职位	教育背景	工作经历
Wen ZHANG 先生	创始人、董事长、 执行董事兼首席 执行官	美国哥伦比亚大学工商管理硕士 美国哈佛大学法学博士	<ul style="list-style-type: none"> • 美国纽约州执业律师，早年于 Kirkland & Ellis LLP 担任律师。 • 在 AI 和硬件技术领域拥有超过 10 年管理经验，长期深耕集成电路、人工智能等新一代信息科技行业。 • 曾任商汤集团总裁、上海鼎域恒睿董事长兼 CEO、映瑞光电总裁兼 CEO。 • 多次获得行业及政府权威认可，包括上海“白玉兰荣誉奖”“上海 IT 行业杰出企业家”“年度创业家”等，并为香港大学客座副教授。
Zhou HONG 先生	执行董事兼首席 技术官	中国北京大学理学学士 清华大学工程硕士 美国纽约州立大学水牛城分校理学硕士	<ul style="list-style-type: none"> • 在 GPU 设计及工程领域拥有近 30 年经验，曾领导国内外多家领先半导体公司的研发团队。 • 曾任美国 S3 公司工程总监、Nvidia 主架构师、华为美国研究中心首席架构师及 S3 Graphics 硬件架构副总裁。 • 在 GPU 领域拥有超过 70 项专利，并为公司 GPGPU 架构的定义和设计做出贡献。
Linglan ZHANG 先生	执行董事兼首席 运营官	中国浙江大学电气工程学士 美国南加州大学电机工程硕士 美国加州大学伯克利分校工商管理硕士	<ul style="list-style-type: none"> • 在半导体行业拥有逾 23 年经验，曾任 Higon Austin 研发中心深度运算副总裁、三星电子美国研发中心高级研发经理，并于 AMD 担任 GPU SoC 架构师。
肖冰先生	执行董事兼集团 总经理	中国清华大学电子工程学士	<ul style="list-style-type: none"> • 在战略运营及销售领域拥有逾 20 年经验，曾任商汤集团业务发展副总裁、Petuum 中国区总经理，并历任 Oracle、IBM 及 Teradata 中国区核心业务负责人。 • 现任集团若干附属公司（珠海壁仞、北京壁仞、杭州壁仞及上海遨岩等）的董事。
Luting PAN 先生	执行董事、董事 会秘书及财务负 责人	美国康涅狄格大学电子工程学士 美国康涅狄格大学电子工程硕士 美国哥伦比亚大学工商管理硕士	<ul style="list-style-type: none"> • 在企业战略运营、风险管理及运营领域经验丰富，曾任法国巴黎银行经理、纽约联合爱迪生商品预测部门高级规划管理人员，并于 Petuum 担任亚洲区总监及战略发展主管。

资料来源：公司招股说明书，光大证券研究所

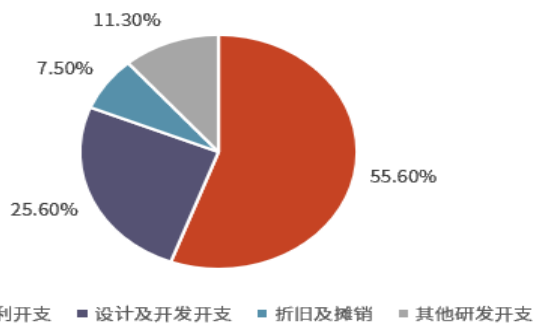
壁仞科技具备经验丰富且拥有深厚的行业知识的研发团队。

公司构建了高比例、高学历、经验丰富的研发团队体系。截至 2025 年 12 月 31 日，研发人员 689 人，占员工总数约 81.8%。

研发团队专业覆盖架构设计、SoC 设计、系统设计与软件工程等关键领域，多数成员曾任职于领先半导体与信息技术公司，具备丰富工程与产业经验。

公司围绕核心技术路线形成系统化的知识产权布局，覆盖 GPGPU 架构、SoC 设计、系统级硬件及软件与 AI 优化等关键领域。

图 20: 2025H1 公司研发开支结构



资料来源：公司招股说明书，光大证券研究所整理。公司 2025 年报未披露 2025 全年数据，因此以 2025H1 作为最新数据展示公司研发开支结构

图 21: 壁仞科技的研发成果和研发投入



资料来源：公司招股说明书，公司 2025 年报，光大证券研究所整理绘制。

3.3 客户和商业化：客户集中度较高，BR106 芯片为销售主力产品

3.3.1 尚处商业化早期阶段，客户集中度较高，前五大客户变动频繁

公司于 2023 年、2024 年及 2025 年的前五大客户采购的产品均为智能计算解决方案。

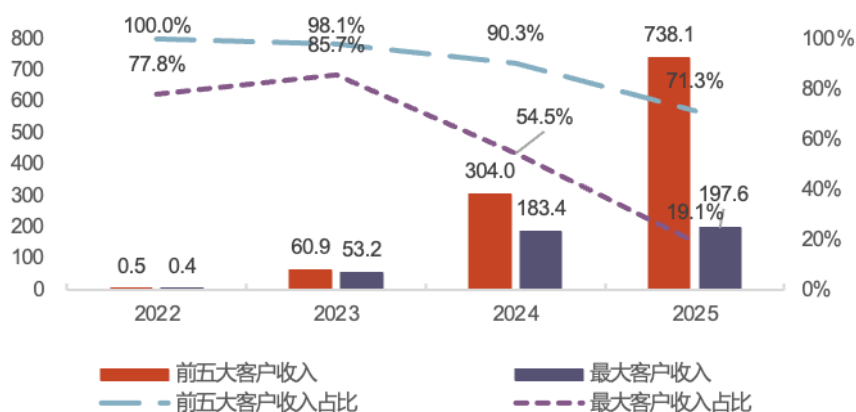
公司前五大客户变动较为频繁，且客户集中度很高、但多元化逐渐提升。2022 年~2025 年各时期内，期间贡献收入的客户主要为新客户，主要系公司现有客户根据其业务需求向公司采购，不必然导致重复购买。当前公司客户集中度很高，随着商业化进程的推进，收入将由更多客户贡献，2024 年和 2025 年客户集中度持续降低。

表 12: 公司各时期客户数量变化及交易情况

	2022 年	2023 年	2024 年	2024H1	2025H1
期间客户数量 (个)	-	12	14	4	12
期间新客户数目 (个)	-	12	13	3	6
交易数量 (个)	-	14	36	9	33
平均交易值 (百万元)	-	4.4	9.4	4.4	1.8

资料来源：公司招股说明书，光大证券研究所。注：2025 年报未披露 2025 全年客户数量和交易情况最新数据，因此最新数据截至 2025H1

图 22: 公司各时期客户采购规模及客户集中度变化 (左轴单位: 百万元)



资料来源: 公司招股说明书, 公司 2025 年报, 光大证券研究所整理绘制

3.3.2 销售: GPGPU BR106 为当前销售主力, 公司同时出售整体智能计算解决方案

公司旗舰 GPGPU 产品 BR106 及新型 BR166 的全形态量产和规模化商业落地推动公司 2025 年芯片销量的增长。BR20X 预计将于 2026 年实现商业化上市, 预期进一步推动芯片销售增长。公司 2025 年实现多个千卡级智算集群交付。

由于销售合约及订单履行的时间限制, 一般来言公司下半年录得的销量数据高于上半年。

表 13: GPGPU 芯片销售量 (单位: 颗)

	2022 年	2023 年	2024 年	2024H1	2025H1
BR106	-	590	9344	600	2216
BR110	-	-	298	-	22

资料来源: 公司招股说明书, 光大证券研究所。注: 2025 年报未披露 2025 全年 GPGPU 芯片销售量, 因此最新数据截至 2025H1

智能计算解决方案销售情况

公司自主开发了基于 GPGPU 的智能计算整体解决方案, 包括基于 GPGPU 的硬件系统和 BIRENSUPA 计算软件平台。公司为更有效满足客户对大规模智能计算能力的需求, 将硬件系统及软件平台与合作伙伴提供的服务器、存储及网络设备等其他硬件基础设施相集成为智能计算集群, 作为一个整体解决方案提供给客户。

公司以智能计算解决方案的形式出售给客户, 是除了出售 GPGPU 芯片外的商业形式。比如 2024 年 9 月, 公司与公司的 IT 合作伙伴紧密合作并成功为一家电信客户交付位于江苏省南京市的 1024-GPU 智能计算集群。

表 14: 公司智能计算解决方案的标杆项目「南京 1024-GPU 集群」项目细节

交付时间	客户	合同金额 (百万元)	集群配置	技术亮点	合作分工	战略意义
2024 年 9 月	电信客户 (位于南京)	180	在 128 台服务器上部署了 1024 个壁仞 106M GPU	针对大模型训练优化, 利用异步卸载和自动化优化, 实现 95% 的线性可扩展性, 减少内存开销。	公司: 提供核心 GPGPU 产品与技术方案。 公司 IT 合作伙伴: 提供存储、网络、集成服务, 并监督端到端集群部署。	证明了大规模部署关键任务智能计算基础设施的能力, 为未来数万至数十万 GPU 的超大规模 AIDC 项目奠定了竞争力基础。

资料来源: 公司招股说明书, 光大证券研究所

3.3.3 客户合作：当前以运营商、国产算力场客户为主，未来积极向互联网客户拓展

客户所属领域

公司行业聚焦，电信运营商是公司的重要客户。公司已与中国三大电信运营商建立合作伙伴关系，以应对多样化的采购需求。公司从服务现有客户中获得宝贵经验和行业知识，并将业务扩展到选定的垂直领域，如 AI 数据中心、电信、AI 解决方案、能源和公用事业、金融科技及互联网。

图 23：壁仞科技合作客户涵盖运营商、智算中心、互联网、行业客户等多领域



资料来源：公司官网

表 15：壁仞科技与是运营商客户重要战略合作的国产芯片厂商

运营商	战略合作性质	关键节点	合作层级判断
中国移动	重要生态合作伙伴	参与“芯合计划”（2023）；参与 OISA 开放互联战略（2025）；呼和浩特智算中心供货（2024）	生态共建 + 技术参与
中国电信	国产 GPU 集采入围	2023 年被认定为“云网基础设施安全国家工程研究中心”云计算合作伙伴；2024 年纳入新一轮国产 GPU 集采名录成为主要 GPU 供应商	集采入围供应商
中国联通	加入智算联盟	2024 年加入中国联通智算联盟	产业联盟成员

资料来源：中国工信新闻网，中国移动公司官网，东方财富网，界面新闻，光大证券研究所

壁仞科技已同运营商客户建立重要战略合作伙伴关系，积极推进国产数据中心/智算中心建设。壁仞科技分别参与“中国移动芯合计划”、“中国联通智算联盟”，并被中国电信纳入国产 GPU 集采名单，目前已经与三家运营商客户进行多地数据中心的商业合作。

表 16：壁仞科技与运营商客户的具体合作项目情况整理

运营商/主体	具体项目	参与方	时间	规模/数量披露
中国移动 (呼和浩特)	中国移动呼和浩特智算中心项目	壁仞参与	2024-07	BR100 芯片，花费 10 亿元，中心单体算力达 6.7EFLOPS (FP16)
中国移动	“芯合”异构混合并行训练系统	中国移动+壁仞科技+中兴通讯+其他	2024-11-30	ITD 算法+基于 GPU Direct RDMA (GDR) 的异构芯片
中国电信	“智算异构四芯混训解决方案”	中国电信研究院+江苏电信+中兴通讯+上海人工智能实验室	2024-12-07	
中国电信	中国电信“千卡集群”落地	中国电信+壁仞科技	2025-01	壁砺系列
中国联通 (浙江)、中兴通讯	浙江乌镇智算中心“国产算力集群”	浙江联通 + 中兴通讯 + 壁仞科技 + 优云科技	2025-07-25	
中国移动 (湖南)、中兴通讯	湖南移动“国产算力资源池”	湖南移动 + 中兴通讯 + 壁仞科技	2025-12-15	
中兴通讯	国内首个光互连光交换 GPU 超节点——光跃 LightSphere X	上海仪电+上海曦智科技+壁仞科技+中兴通讯	2025-07	

资料来源：壁仞科技公司官网，澎湃新闻，华尔街见闻，中国工信新闻网，C114 通信网，光明网，集微网，光大证券研究所整理

壁仞科技积极同国家训练场开展算力合作。公司已经同上海仪电达成合作，目标部署“壁砺 166 系列千卡智算集群”。

表 17: 壁仞科技与国家训练场的具体合作项目情况整理

合作主体	合作性质	时间	合作内容	战略意义判断
壁仞科技 + 上海仪电(智算平台)	千卡集群支持	2026 年初	部署壁砺 166 系列千卡智算集群	形成模型适配与算力验证闭环
壁仞科技 + 智算服务	模型适配	2026-02	壁砺 166 系列完成与阶跃星辰最新发布 Step 3.5 Flash 模型适配	提升国产 GPU 模型适配能力

资料来源: 公司官网, 光大证券研究所整理

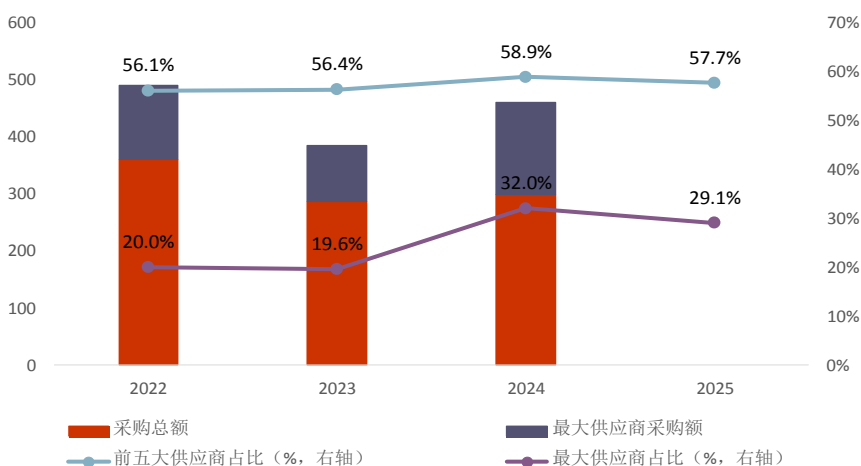
3.4 供应链: 晶圆制造&封测由第三方提供, 实体清单后采购转向国产替代

壁仞采用“无晶圆厂”模式运营, 研发资源聚焦于芯片设计等核心环节。

公司不自建产线, GPGPU 的晶圆制造、封装、测试等环节均由第三方合约制造商完成。在研发阶段, IP、EDA 及部分后端与物理设计环节依托第三方供应与外包支持。截至 2025 年上半年, 前五大供应商主要提供原材料、设计及开发服务、合同生产及服务(2025 年报未披露前五大供应商提供的产品和服务类别)。

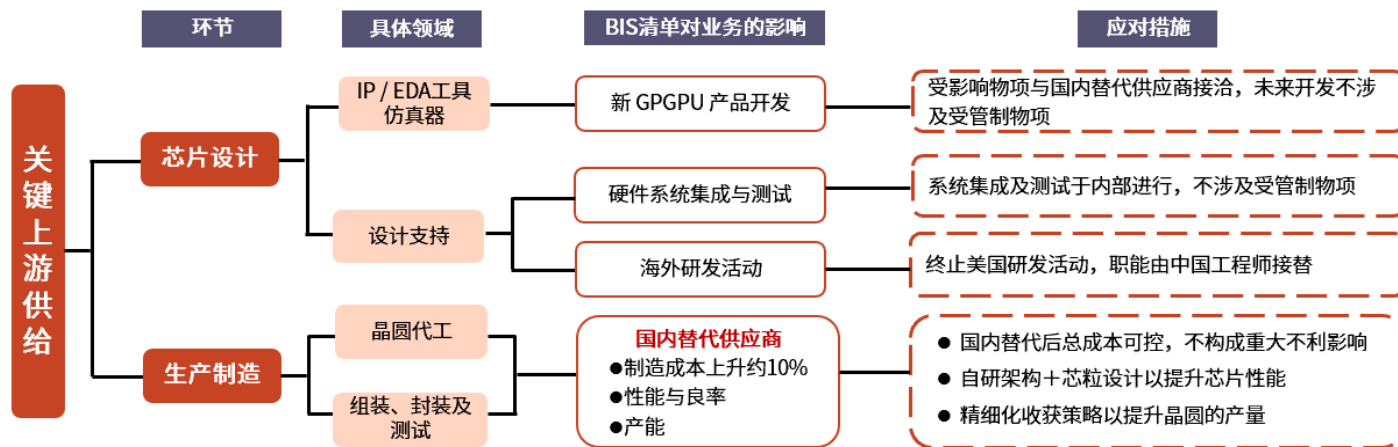
BIS 将本集团若干实体列入实体清单后, 公司受影响物项采购已转向国内替代。产品销售与订单。BR106、BR166、BR110 订单稳定, 未出现客户取消或缩减。研发与技术支持。销售、技术支持及 BIRENSUPA 开发不依赖受管制物项。

图 24: 公司采购规模及供应商集中度变化 (单位: 百万元)



资料来源: 公司招股说明书, 2025 年报, 光大证券研究所整理绘制

图 25: BIS 清单事件下关键上游供给影响及应对路径



资料来源: 公司招股说明书, 光大证券研究所整理绘制

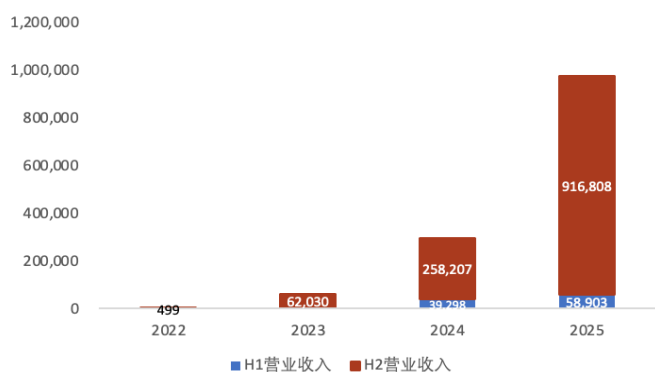
4、财务分析: 营收快速增长, 持续研发投入, 公司尚处于商业化初期

公司营业收入自商业化以来快速增长。收入由 2022 年的人民币 0.5 百万元增至 2023 年的 62.0 百万元, 并于 2024 年进一步增至 336.8 百万元。2025 年全年实现营业收入人民币 1,034.6 百万元, 同比增长 207.2%, 商业化放量明显加速。

收入放量节奏呈现下半年偏强特征。受销售合约签署及订单履行时间安排因素影响, 2025 年下半年收入高于上半年; 2025 年全年收入达 1,034.6 百万元, 较上半年的 58.9 百万元显著增长, 显示订单签署及履约确认主要集中于下半年, 收入释放弹性较强。

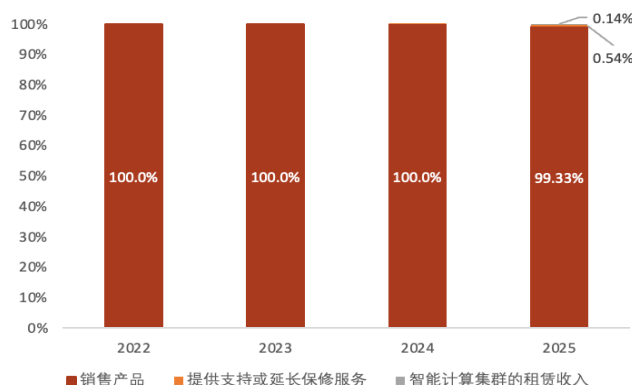
销售产品为主要收入来源, 其中智能计算解决方案占据主导。2022-2024 年销售产品收入占比均为 100.0%, 2025 年占比仍达 99.33%。支持或延长保修服务及智能计算集群租赁收入占比分别约 0.54%和 0.14%, 合计不足 1%, 对整体收入贡献仍较低。

图 26: 营收逐年增长, 2025 营收同比增速 207.2% (单位: 千元)



资料来源: 公司招股说明书, 2025 年报, 光大证券研究所; 营业收入 (红色) 为全年总收入, H1 营业收入 (蓝色) 为当年上半年收入

图 27: 销售产品业务持续贡献公司绝大部分营收



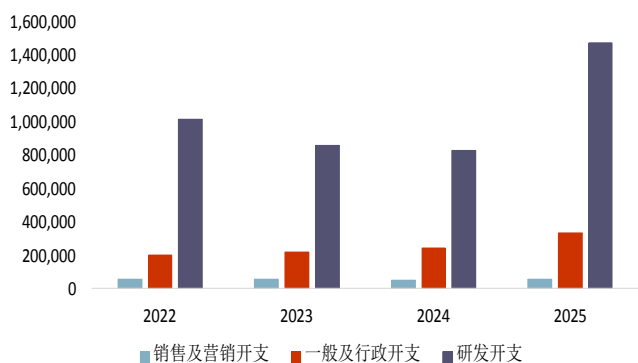
资料来源: 公司招股说明书, 2025 年报, 光大证券研究所; 注: 销售产品包括智能计算解决方案及代理费。

整体毛利率由智能计算解决方案毛利率决定：公司绝大部分收入来自销售产品（智能计算解决方案占主要），其他收入项目占比很低，对整体毛利率影响有限。2025 年公司实现毛利人民币 557.0 百万元，同比增长 210.8%，整体毛利率为 53.8%，较 2024 年的 53.2% 小幅提升 0.6 个百分点。

2025 年毛利率整体保持稳定，反映规模放量下盈利能力未明显恶化：结合收入结构看，公司 2025 年产品销售收入占比仍高达 99.3%，收入结构仍以智能计算解决方案为核心；在收入大幅增长背景下，整体毛利率较上年基本持平，说明公司商业化扩张并未显著压缩单位盈利空间。

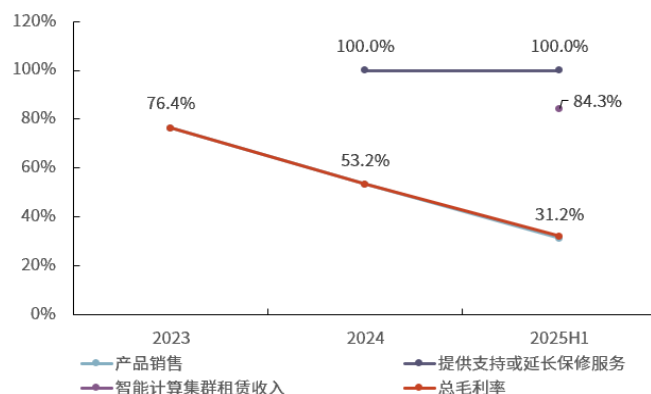
费用结构体现商业化初期特征，研发投入占据主导：2025 年公司销售及营销开支、行政开支和研发开支分别为 57.5 / 331.5 / 1,476.1 百万元，同比分别增长 11.6% / 35.8% / 78.5%；其中，行政开支增速相对较快，主要由于全球发售相关上市开支及其他专业服务费用增加所致；整体来看，公司当前费用仍主要投向研发与组织能力建设，而非单纯的销售投放，仍处于高研发投入驱动商业化扩张的阶段。

图 28：期间费用结构以研发投入为主（单位：千元）



资料来源：公司招股说明书，2025 年报，光大证券研究所

图 29：总毛利率走势与产品销售毛利率保持一致



资料来源：公司招股说明书，光大证券研究所。注：公司总毛利率和产品销售毛利率在 2023~2025H1 时期数据均一致，对应折线重叠；2025 年报未披露各项业务具体毛利率，因此最新数据截至 2025H1

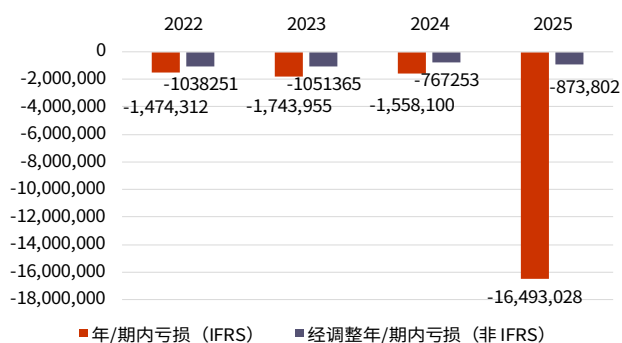
尽管实现强劲的收入增长，公司于往绩记录期仍录得亏损。公司于 2022 年、2023 年、2024 年及 2025 年分别录得年内亏损人民币 1,474.3/1,744.0/1,538.1/16,493.0 百万元。

亏损主要反映公司仍处于商业化的较早阶段。由于在实现规模化商业化之前需完成芯片设计、客户验证及市场渗透，公司于往绩记录期的收入尚未涵盖前期投入，相关规模经济效益仍需逐步释放。

公司披露经调整年内仍录得亏损（非 IFRS）。通过加回可赎回可转换优先股公允价值变动、以股份为基础的薪酬开支及上市开支，以剔除非现金及一次性项目的影响。调整后亏损规模较 IFRS 口径明显收窄，公司于 2022 年、2023 年、2024 年及 2025 年分别录得经调整亏损人民币 1,038.3 / 1,051.4 / 767.3 / 873.8 百万元。尤其是 2025 年，公司 IFRS 口径净亏损高达 164.9 亿元，但经调整后净亏损为 8.7 亿元，标准报表口径下的大额亏损主要受赎回负债相关非现金项目扰动，并不完全反映核心经营表现。

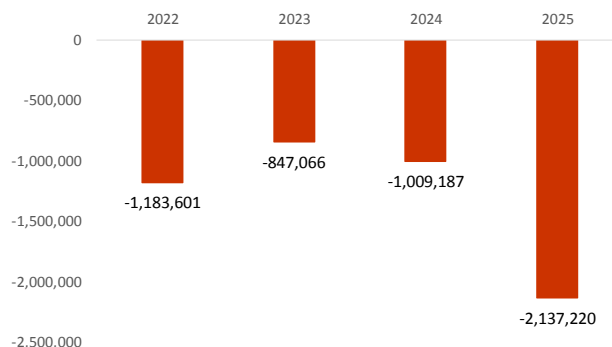
公司经营活动所用现金净额持续为负。2022 年~2025 年已披露的经营活动所用现金净额持续为负，主要由于经营亏损及营运资金占用；公司计划通过提升盈利能力及加强贸易应收款项管理，以改善经营现金流出净额状况。

图 30: 公司年内亏损与经调整净亏损对比 (单位: 千元)



资料来源: 公司招股说明书, 2025 年报, 光大证券研究所

图 31: 公司经营活动所用现金净额持续为负 (单位: 千元)



资料来源: 公司招股说明书, 2025 年报, 光大证券研究所

5、盈利预测与估值

5.1 关键假设及盈利预测

5.1.1 营收假设

壁仞科技营收结构方面, 公司按照产品销售、提供支持或延保服务、智能计算集群的租赁收入, 其中产品销售中的智能计算解决方案销售贡献公司绝大部分营收。2022/2023/2024/2025 年产品销售分别占据公司总营收的 100%/100%/100%/99.3%。

一、产品销售收入预测: 按照“量*价”的逻辑, 我们分别预测壁仞科技各类 AI 芯片的出货量和单价变化。

营收假设一: 产品出货节奏, 性能更优、产品竞争力更强、更高价值量的新品芯片出货是营收增长的关键。①BR106 方面, 我们假设 2025 年 BR106 仍然积极出货, 出货量较 2024 年同比提升; ②BR166 方面, 我们假设 2025H1 处于备货阶段, 2025H2 具备量产/出货能力, 且 2026 年出货量进一步提升; ③BR20X 方面, 我们预计将于 2026 年实现少量出货, 将于 2027 年实现大规模出货; ④BR30X 方面, 我们预计将于 2027 年推出并少量出货, 将于 2028 年实现大规模出货。

营收假设二: 季节性影响。公司普遍在上半年进行集中备货与产能准备, 而客户普遍在下半年完成交付和确收, 因此此前年份均出现下半年营收显著高于上半年度情况。我们假设季节性仍会存在, 但未来随着客户分散、供应链成熟等商业化进展加速, 上半年和下半年的营收季节性差异会减弱。

营收假设三: 订单与订单转换。截至 2025 年 12 月 15 日, 公司针对其特专科技产品订立五份框架销售协议及 24 份销售合同, 总价值约为人民币 12.41 亿人民币元, 其中未完成的具有约束力的订单总价值约 8.22 亿人民币元。通常订单转化成收入需要数个月到半年时间, 因此我们预计这将为 2026 年营收带来一定支撑。

营收假设四: 存货是未来营收的供给端保障。公司存货包括原材料、在制品和制成品三类, 2022/2023/2024/2025 年的存货分别为 0.39/1.73/1.53/9.49 亿人民币元, 其中 2025 年存货明显提升, 我们认为公司已针对 BR106、BR166 等产品进行充分备货, 为 2026 年营收高增长奠定供应链基础。

营收假设五: 国内公司 AI 基础设施建设需求伴随 AI agent、Openclaw 等不断带动, 公司有望加速向运营商、国家算力场、大型企业等客户出货, 并积极进行

互联网云服务商的适配和验证通过，开启向互联网等更多元客户的商业化进程。未来客户数和交易订单量均有望持续快速增长。

营收假设六：国产芯片行业正处于商业化放量的初期阶段，市场规模预计持续高速增长，灼识咨询预测 2029 年中国智能计算芯片的市场规模将达到 2012 亿美元，对应 2024-2029 年 CAGR 为 46.3%。公司作为高端国产 AI GPU 厂商，伴随供应链改善和客户验证推进，在中国 AI 算力市场份额有望不断提升。

根据公司招股说明书,2023/2024/2025H1 期间,公司主要出货 BR106 和 BR110 两款芯片产品, BR106 分别出货 590/9344/2216 颗, BR110 分别出货 0/298/22 颗。

出货量方面，根据以上六个营收假设，我们预计：

2026 年：分别销售 BR106、BR166、BR20X、BR110 芯片产品 1.50、2.25、0.10、0.02 万颗；

2027 年：分别销售 BR106、BR166、BR20X、BR30X、BR110、BR310 芯片产品 0.60、2.48、4.00、0.18、0.03、0.10 万颗；

2028 年：分别销售 BR106、BR166、BR20X、BR30X、BR110、BR310 芯片产品 0.12、0.99、4.40、5.00、0.04、0.25 万颗。

单价方面，我们在以上六个营收假设的基础上，进一步假设 1) 训推一体芯片 (BR106、BR166、BR20X、BR30X 等) 价格明显高于推理芯片 (BR110、BR310 等)；2) 使用 Chiplet 技术进行双 Die 合封的芯片价格会明显高于单 Die 芯片；3) 新型芯片随算力性能提升，价格会较老款芯片有所提升，但每代芯片产品价格逐年下滑。

我们预计 2026/2027/2028 年产品销售营收分别实现 23.27/70.12/122.41 亿人民币元，同比增长 126%/201%/75%。

二、其他收入预测：我们预计提供支持或延保服务、智能计算集群的租赁收入不会成为主要营收贡献项。我们假设 2026/2027/2028 年其他营收占比稳定在 0.5%。

表 18：壁仞科技营收预测

	2023	2024	2025	2026E	2027E	2028E
产品销售收入 (千元)	62,030	336,794	1,027,655	2,326,583	7,012,159	12,241,073
YoY	12330.9%	443.0%	205.1%	126.4%	201.4%	74.6%
营收贡献占比	100.0%	100.0%	99.3%	99.5%	99.5%	99.5%
其他收入 (千元)	0	9	6,959	11,691	35,237	61,513
YoY	/	/	77222.2%	68.0%	201.4%	74.6%
营收贡献占比	0.0%	0.0%	0.7%	0.5%	0.5%	0.5%
总营收 (千元)	62,030	336,803	1,034,614	2,338,274	7,047,396	12,302,586
YoY	12330.9%	443.0%	207.2%	126.0%	201.4%	74.6%

资料来源：公司招股说明书，2025 年报，光大证券研究所预测

5.1.2 毛利率预测

公司 2023/2024/2025 年毛利率分别为 76.4%/53.2%/53.8%，主要系营收确认的季节性影响、以及商业化早期毛利率水平较不稳定。

毛利率假设一：我们参考国产算力可比公司，发现平均毛利率水平在 50%-55% 区间。我们认为随着供应链打通、商业化开启，公司毛利率水平应逐渐稳定、且逐步向同行业优秀公司的水平靠拢。

表 19：壁仞科技与可比公司毛利率对比表

证券代码	公司名称	2023	2024	2025	2026E	2027E
688521.SH	芯原股份	44.6%	39.9%	34.1%	32.6%	32.6%
688802.SH	沐曦股份	61.4%	53.3%	56.4%	44.7%	41.3%
688795.SH	摩尔线程	24.4%	70.5%	65.5%	63.5%	61.6%
688041.SH	海光信息	59.2%	63.2%	57.4%	58.6%	58.3%
688256.SH	寒武纪	69.0%	56.6%	55.0%	51.9%	52.7%
300474.SZ	景嘉微	59.7%	42.2%	46.4%	63.2%	65.2%
	平均	53.1%	54.3%	52.5%	52.4%	52.0%
6082.HK	壁仞科技	76.4%	53.2%	53.8%	49.4%	50.9%

资料来源：wind，光大证券研究所预测，Refinitiv 预测。2026、2027 年预测值方面，壁仞科技为光大证券研究所预测，其余公司为 Refinitiv 一致预期。

毛利率假设二：公司销售成本涵盖已使用存货及消耗品、雇员福利开支、其他成本三项构成，其中 25H1 已使用存货及消耗品占总销售成本的 91.5%。考虑到公司切换至国产供应链，产品适配需要磨合，叠加产品放量初期或影响毛利率，我们预计 2026 年和 2027 年毛利率会相对下滑。但考虑到供应链（晶圆代工、封装测试）技术迭代和良率提升，以及产品出货规模化效应凸显，单颗芯片 Die 的成本增长幅度应低于价格提升幅度；且高端芯片产品毛利率更高，产品结构改善将持续利好后续毛利率提升。

我们预计 2026/2027/2028 年毛利率分别为 49.4%/50.9%/51.7%。

5.1.3 费用预测

由于公司处于商业化早期阶段，营收规模仍有限导致费用率高且不稳定，因此我们对销售及营销开支、一般及行政开支、研发开支分别假设同比增速：

- 1) 销售及营销开支假设：我们预计 2026/2027/2028 年同比增速分别为 40%/35%/30%，公司积极投入销售及营销开支，加强向运营商、算力训练场、ToB 客户等出货，并积极对互联网云服务商进行产品适配和运维服务。
- 2) 一般及行政开支：我们预计 2026/2027/2028 年同比增速分别为 -20%/+10%/+10%，我们预计保持稳健增长，2026 年预计同比下滑系一次性的上市开支推高了 2025 年基数。
- 3) 研发开支：我们预计 2026/2027/2028 年同比增速分别为 35%/30%/25%，公司积极迭代产品和技术，当前 BR20X、BR30X、BR310 等新产品待发布和量产，维持较高的研发开支增速水平。

财务收入和财务成本也对公司当前业绩带来显著影响，因此我们对财务收入和财务成本进行假设：

- 1) 财务收入：考虑到 2026 年公司上市后的募资金额将显著带动公司货币现金提升，而后货币现金将应用于研发支出等项目而略有下滑，我们预计 2026/2027/2028 年财务收入同比增速分别为+100%/-10%/-10%。
- 2) 财务成本：拆分公司财务成本，绝大多数为赎回负债的账面值变动（2022/2023/2024/2025 年赎回负债的账面值变动占财务成本的比例分别达到 98.8%/98.0%/94.6%/99.9%），而上市完成后此亏损将自动注销，因此我们预计 2026/2027/2028 年财务成本同比增速分别为-99.7%/+10%/+10%。

表 20: 壁仞科技盈利预测表

指标	2024	2025	2026E	2027E	2028E
营业收入 (百万元)	337	1,035	2,338	7,047	12,303
营业收入增长率	443.0%	207.2%	126.0%	201.4%	74.6%
净利润 (百万元)	(1538)	(16493)	(869)	926	2835
调整后净利润 (百万元)	(767)	(874)	(722)	1,099	3,043
P/S	217	81	46	15	9

资料来源: 公司公告, wind, 光大证券研究所预测, 股价时间为 2026-05-06; 1 HKD = 0.8717 CNY。

5.2 相对估值

我们选取国产算力芯片公司寒武纪、海光信息、景嘉微、芯原股份、沐曦股份、摩尔线程为可比公司。其中 寒武纪专注于人工智能领域的核心处理器芯片; 海光信息是国产 x86 架构处理器龙头, 核心产品包括海光通用处理器(CPU)和海光协处理器(DCU); 景嘉微是从事高可靠领域的国产化图形处理芯片(GPU)厂商; 芯原股份是一站式、平台化芯片定制服务和半导体 IP 授权服务厂商, 拥有 GPU IP、NPU IP、VPU IP 等六大类处理器 IP; 沐曦股份、摩尔线程、燧原科技和壁仞科技并称“国产 GPU 四小龙”, 均为国产 GPU 企业。

因壁仞科技尚未实现盈利, 我们预计 2027 年实现净利润和调整后的净利润为扭亏为盈, 我们采取 PS 估值法。可比公司当前股价对应 2026/2027 年的平均预测 PS 分别为 58x/34x, 公司对应 2026/2027 年预测 PS 为 46x/15x, 公司较其他可比公司而言, 2027 年 PS 估值仍具明显估值提升空间。

表 21: 壁仞科技与可比公司的盈利预测和估值比较

股票代码	公司	市值 (亿人民币)	PS(x)		营收 (百万人民币)		净利润 (百万人民币)		营收增长率	
			26E	27E	26E	27E	26E	27E	26E	27E
688521.SH	芯原股份	1,465	25.7	17.4	5,691	8,413	393	904	80.8%	47.8%
688802.SH	沐曦股份	3,245	81.4	44.3	3,987	7,332	-34	564	143.2%	83.9%
688795.SH	摩尔线程	3,474	133.8	74.1	2,596	4,687	-1,003	57	73.0%	80.5%
688041.SH	海光信息	8,005	34.7	25.0	23,040	32,075	4,769	7,068	61.7%	39.2%
688256.SH	寒武纪	7,696	40.9	22.1	18,836	34,876	5,752	9,636	191.1%	85.2%
300474.SZ	景嘉微	351	31.8	22.6	1,105	1,553	90	211	56.3%	40.5%
	平均		58.1	34.2	/	/	/	/	101.0%	62.9%
6082.HK	壁仞科技	1,082	46.3	15.4	2,338	7,047	-869	926	126.0%	201.4%

资料来源: 公司公告, wind, 壁仞科技盈利预测为光大证券研究所预测, 其余公司为 Refinitiv 一致预期。股价截至 2026-05-06, 1HKD = 0.8717 CNY

5.3 绝对估值

5.3.1 关于基本假设的几点说明:

- 1、长期增长率: 根据行业趋势分析, 我们认为壁仞科技作为国产算力芯片公司, 将受益于 AI 浪潮和国产替代趋势, 故假设长期增长率为 2.0%;
- 2、无风险利率 Rf: 选取美国 10 年期国债收益率作为无风险利率, 截至 2026 年 5 月 6 日, 美国 10 年期国债收益率为 4.371%;
- 3、 β 值选取: 采用半导体 (SW 港股) 板块的加权剔除财务杠杆原始 β 作为公司无杠杆 β 系数的假设值, 并结合其他参数给出公司的 β 系数为 1.07。

表 22: 绝对估值核心假设表

关键性假设	数值
第二阶段年数	8
长期增长率	2.00%
无风险利率 Rf	4.37%
β (β levered)	1.07
Rm-Rf	3.50%
Ke(levered)	8.11%
税率	10.00%
Kd	3.15%
Ve (百万元)	58277.91
Vd (百万元)	200.14
目标资本结构	0.34%
WACC	8.10%

资料来源: 光大证券研究所预测

表 23: 现金流折现及估值表

	现金流折现值 (百万元)	价值百分比
第一阶段	(1815.52)	-1.34%
第二阶段	48315.40	35.64%
第三阶段 (终值)	89060.30	65.70%
企业价值 AEV	135560.18	100.00%
加: 非经营性净资产价值	0.00	0.00%
减: 少数股东权益 (市值)	0.00	0.00%
减: 债务价值	200.14	-0.15%
总股本价值	135360.04	99.85%
股本 (百万股)	2439.00	
每股价值 (元)	55.50	

资料来源: 光大证券研究所预测

基于上述假设, 根据 FCFF 法估值得出公司合理价格为 55.50 人民币, 假设汇率为 1HKD = 0.8717CNY, 对应 63.67 港币。

5.4 估值结论与投资评级

我们预计壁仞科技 2026/2027/2028 年分别实现营收 23.38/70.47/123.03 亿人民币元, 同比增长 126%/201%/75%; 净利润分别为-8.69/9.26/28.35 亿人民币元, 调整后净利润分别为-7.22/10.99/30.43 亿人民币元。当前对应 2026/2027 年 PS 为 46x/15x。

目前壁仞科技的 2027 年 PS 估值水平仍显著低于同行业公司, 叠加 FCFF 法估值得出公司合理价格为 63.67 港币。未来潜在催化: 1) 产品迭代方面, 公司伴随 BR166 和 BR20X 等高性能芯片放量, 带来更高营收贡献; 2) 供应链方面, 2026 年将成为国产国造芯片的元年, 随着中国先进制程产能的持续爬升, 我们预计公司的供应链问题将在 2026 年底到 2027 年明显改善; 3) 客户拓展方面, 期待公司加速向运营商、国家算力场、大型企业等客户出货, 并积极进行互联网云服务商的适配和验证通过, 开启向互联网等更多元客户的商业化进程。考虑到

国产算力出货和公司商业化开启，看好公司业绩增长和估值抬升。首次覆盖，给予公司“买入”评级。

6、风险分析

AI 大模型迭代及下游应用进展不及预期，大模型训练和推理的算力需求的下行风险：AIGC 浪潮提振算力需求，倘若后续 AI 大模型迭代及下游应用进展不及预期，算力需求下行将导致芯片、服务器、数据中心等产业链公司业绩增速下行。

AI 行业竞争加剧的风险。AMD、英特尔等其他芯片厂商积极布局 AI 芯片，同时云服务商联合博通、迈威尔科技等厂商开展定制化 AI 芯片研发，增加了整体 AI 算力行业的竞争水平，可能挤压厂商利润空间。

技术迭代不及预期的风险。公司业务的成功取决于能否及时推出具备良好功能和性能水平的产品或解决方案、为客户提供价值、同时支持并紧跟行业重大发展趋势，若研发进展或性能迭代效果不及预期，则可能影响客户商业化进展和公司业绩释放。

供应链风险：当前国产 AI 芯片在供应端仍面临供不应求局面，若公司第三方的供应链中断或延期，可能会延迟公司的产品研发计划和量产预期。

市场拓展和商业化进展不及预期的风险：公司产品及解决方案于日后的商业成功将取决于其市场接受度及客户需求。未能正确估计客户需求可能导致供需失衡。倘公司未能为产品及解决方案建立、扩展及优化有效的销售网络，可能无法按计划产生收入，而公司业务及经营业绩可能会受到不利影响。

次新股股价波动风险：公司初期公众持股比例较低且存在大股东禁售期，需警惕上市后 6 个月及 12 个月关键解禁节点带来的筹码供应压力及二级市场定价回归风险。

财务报表与盈利预测

利润表 (百万元)	2024	2025	2026E	2027E	2028E
营业收入	337	1,035	2,338	7,047	12,303
营业成本	158	478	1,182	3,461	5,938
折旧和摊销	124	149	157	190	227
税金及附加	0	0	0	0	0
销售费用	52	57	80	109	141
管理费用	244	332	265	292	321
财务费用	703	15,451	(17)	(4)	41
研发费用	827	1,476	1,993	2,591	3,238
资产减值损失	0	0	0	0	0
其他收益	36	12	70	141	246
投资收益	0	0	16	22	27
营业利润	(1,595)	(16,763)	(1,081)	740	2,876
利润总额	(1,538)	(16,493)	(869)	926	2,953
所得税	(0)	0	0	0	118
净利润	(1,538)	(16,493)	(869)	926	2,835
调整后净利润	(767)	(874)	(722)	1,099	3,043

现金流量表 (百万元)	2024	2025	2026E	2027E	2028E
经营活动现金流	(1,009)	(2,137)	(1,619)	(1,518)	1,023
净利润	(767)	(874)	(722)	1,099	3,043
折旧摊销	124	149	157	190	227
净营运资金增加	585	2,899	2,684	3,904	3,613
其他	(951)	(4,311)	(3,737)	(6,712)	(5,860)
投资活动产生现金流	1,218	(1,235)	(353)	(298)	(293)
净资本支出	101	351	(320)	(320)	(320)
长期投资变化	(2)	(248)	0	0	0
其他资产变化	1,120	(1,338)	(33)	22	27
融资活动现金流	217	3,330	5,646	1,212	840
股本变化	0	9	7	0	0
债务净变化	270	(59)	(199)	1,208	881
无息负债变化	828	19,355	(28,111)	1,639	1,371
净现金流	441	(63)	3,674	(604)	1,570

主要指标

盈利能力 (%)	2024	2025	2026E	2027E	2028E
毛利率	53.2%	53.8%	49.4%	50.9%	51.7%
EBITDA 率	-163.5%	-110.9%	-40.9%	13.1%	25.5%
EBIT 率	-200.2%	-125.3%	-47.6%	10.4%	23.7%
税前净利润率	-456.7%	-1594.1%	-37.2%	13.1%	24.0%
归母净利润率	-456.7%	-1594.1%	-37.2%	13.1%	23.0%
ROA	-50.6%	-278.7%	-7.8%	6.2%	14.2%
ROE (摊薄)	20.7%	69.2%	-9.0%	8.8%	21.2%
经营性 ROIC	-424.7%	-31.2%	-17.6%	7.1%	19.9%

偿债能力	2024	2025	2026E	2027E	2028E
资产负债率	344.1%	502.9%	13.1%	28.9%	32.8%
流动比率	1.52	4.74	7.90	3.30	2.91
速动比率	1.42	3.82	7.11	2.75	2.45
归母权益/有息债务	(26.31)	(106.67)	395.20	8.57	6.34
有形资产/有息债务	10.48	25.68	447.41	11.91	9.35

资料来源: Wind, 光大证券研究所预测。注: 股本方面, 公司 2020~2025 年进行多次上市前融资, 且 2026 年 IPO 融资, 2024/2025 年股本为 16.46/18.92 亿股, 2026~2028 年股本假设为当前股本 24.39 亿股。

资产负债表 (百万元)	2024	2025	2026E	2027E	2028E
总资产	3,041	5,918	11,091	14,865	19,953
货币资金	1,655	1,510	5,185	4,581	6,151
交易性金融资产	96	1,313	1,313	1,329	1,345
应收账款	87	525	1,146	2,763	4,220
应收票据	0	0	0	0	0
其他应收款 (合计)	36	75	263	794	1,386
存货	153	949	990	2,235	2,954
其他流动资产	47	46	46	46	46
流动资产合计	2,353	4,883	9,830	13,478	18,478
其他权益工具	0	0	0	0	0
长期股权投资	0	15	15	15	15
固定资产	366	537	556	570	569
在建工程	0	0	113	197	260
无形资产	84	176	178	179	180
商誉	0	0	0	0	0
其他非流动资产	54	125	190	190	190
非流动资产合计	688	1,035	1,262	1,387	1,475
总负债	10,465	29,762	1,451	4,299	6,551
短期借款	0	200	0	1,207	2,088
应付账款	33	135	292	856	1,469
其他流动负债	0	0	0	0	0
流动负债合计	1,553	1,030	1,244	4,090	6,341
长期借款	0	0	1	2	3
应付债券	0	0	0	0	0
其他非流动负债	148	194	194	194	194
非流动负债合计	8,912	28,731	207	208	209
无息负债	10,183	29,538	1,427	3,066	4,437
有息负债	282	224	24	1,233	2,114
股东权益	(7,424)	(23,844)	9,640	10,567	13,402
股本	33	42	49	49	49
公积金	0	0	25,044	25,068	25,068
未分配利润	(6,518)	(23,011)	(23,880)	(22,978)	(20,143)
归属母公司权益	(7,424)	(23,844)	9,640	10,567	13,402

费用率	2024	2025	2026E	2027E	2028E
销售费用率	15.30%	5.56%	3.44%	1.54%	1.15%
管理费用率	72.49%	32.04%	11.34%	4.14%	2.61%
财务费用率	208.74%	1493.42%	-0.74%	-0.06%	0.33%
研发费用率	245.53%	142.68%	85.23%	36.76%	26.32%
所得税率	0.01%	0.00%	0.00%	0.00%	4.00%

每股指标	2024	2025	2026E	2027E	2028E
每股红利	0.00	0.00	0.00	0.00	0.00
每股经营现金流	(0.61)	(1.13)	(0.66)	(0.62)	0.42
每股净资产	(4.51)	(12.60)	3.95	4.33	5.50
每股销售收入	0.20	0.55	0.96	2.89	5.04

估值指标	2024	2025	2026E	2027E	2028E
PS	217	81	46	15	9
PE	-	-	-	117	38
股息率	0.00%	0.00%	0.00%	0.00%	0.00%

行业及公司评级体系

	评级	说明
行业及公司评级	买入	未来 6-12 个月的投资收益率领先市场基准指数 15%以上
	增持	未来 6-12 个月的投资收益率领先市场基准指数 5%至 15%;
	中性	未来 6-12 个月的投资收益率与市场基准指数的变动幅度相差-5%至 5%;
	减持	未来 6-12 个月的投资收益率落后市场基准指数 5%至 15%;
	卖出	未来 6-12 个月的投资收益率落后市场基准指数 15%以上;
	无评级	因无法获取必要的资料, 或者公司面临无法预见结果的重大不确定性事件, 或者其他原因, 致使无法给出明确的投资评级。
基准指数说明:		A 股市场基准为沪深 300 指数; 香港市场基准为恒生指数; 美国市场基准为纳斯达克综合指数或标普 500 指数。

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。

分析师声明

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师, 以勤勉的职业态度、专业审慎的研究方法, 使用合法合规的信息, 独立、客观地出具本报告, 并对本报告的内容和观点负责。负责准备以及撰写本报告的所有研究人员在此保证, 本研究报告中任何关于发行商或证券所发表的观点均如实反映研究人员的个人观点。研究人员获取报酬的评判因素包括研究的质量和准确性、客户反馈、竞争性因素以及光大证券股份有限公司的整体收益。所有研究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的具体的推荐意见或观点有直接或间接的联系。

法律主体声明

本报告由光大证券股份有限公司制作, 光大证券股份有限公司具有中国证监会许可的证券投资咨询业务资格, 负责本报告在中华人民共和国境内(仅为本报告目的, 不包括港澳台)的分销。本报告署名分析师所持中国证券业协会授予的证券投资咨询执业资格编号已披露在报告首页。

中国光大证券国际有限公司是光大证券股份有限公司的关联机构。

特别声明

光大证券股份有限公司(以下简称“本公司”)成立于 1996 年, 是中国证监会批准的首批三家创新试点证券公司之一, 也是世界 500 强企业——中国光大集团股份公司的核心金融服务平台之一。根据中国证监会核发的经营证券期货业务许可, 本公司的经营范围包括证券投资咨询业务。

本公司经营范围: 证券经纪; 证券投资咨询; 与证券交易、证券投资活动有关的财务顾问; 证券承销与保荐; 证券自营; 为期货公司提供中间介绍业务; 证券投资基金代销; 融资融券业务; 中国证监会批准的其他业务。此外, 本公司还通过全资或控股子公司开展资产管理、直接投资、期货、基金管理以及香港证券业务。

本报告由光大证券股份有限公司研究所(以下简称“光大证券研究所”)编写, 以合法获得的我们相信为可靠、准确、完整的信息为基础, 但不保证我们所获得的原始信息以及报告所载信息之准确性和完整性。光大证券研究所可能将不时补充、修订或更新有关信息, 但不保证及时发布该等更新。

本报告中的资料、意见、预测均反映报告初次发布时光大证券研究所的判断, 可能需随时进行调整且不予通知。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。客户应自主作出投资决策并自行承担投资风险。本报告中的信息或所表述的意见并未考虑到个别投资者的具体投资目的、财务状况以及特定需求。投资者应当充分考虑自身特定状况, 并完整理解和使用本报告内容, 不应视本报告为做出投资决策的唯一因素。对依据或者使用本报告所造成的一切后果, 本公司及作者均不承担任何法律责任。

不同时期, 本公司可能会撰写并发布与本报告所载信息、建议及预测不一致的报告。本公司的销售人员、交易人员和其他专业人员可能会向客户提供与本报告中观点不同的口头或书面评论或交易策略。本公司的资产管理子公司、自营部门以及其他投资业务板块可能会独立做出与本报告的意见或建议不相一致的投资决策。本公司提醒投资者注意并理解投资证券及投资产品存在的风险, 在做出投资决策前, 建议投资者务必向专业人士咨询并谨慎抉择。

在法律允许的情况下, 本公司及其附属机构可能持有报告中提及的公司所发行证券的头寸并进行交易, 也可能为这些公司提供或正在争取提供投资银行、财务顾问或金融产品等相关服务。投资者应当充分考虑本公司及本公司附属机构就报告内容可能存在的利益冲突, 勿将本报告作为投资决策的唯一信赖依据。

本报告根据中华人民共和国法律在中华人民共和国境内分发, 仅向特定客户传送。本报告的版权仅归本公司所有, 未经书面许可, 任何机构和个人不得以任何形式、任何目的进行翻版、复制、转载、刊登、发表、篡改或引用。如因侵权行为给本公司造成任何直接或间接的损失, 本公司保留追究一切法律责任的权利。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。

光大证券股份有限公司版权所有。保留一切权利。

光大证券研究所

上海

静安区新闻路 1508 号
静安国际广场 3 楼

北京

西城区复兴门外大街 6 号
光大大厦 17 层

深圳

福田区深南大道 6011 号
NEO 绿景纪元大厦 A 座 17 楼

光大证券股份有限公司关联机构

香港

中国光大证券国际有限公司

香港湾仔告士打道 108 号光大中心 33 楼