

► **CPU+DCU 双轮齐发，铸造国产算力中坚力量。**海光是国内领先的高端计算芯片企业，聚焦 DCU 与 CPU 双产品并行研发，覆盖云计算、人工智能等主流应用场景。公司最大股东为中科曙光，截至 2025 年一季度持股比例 27.96%。受下游算力需求带动，公司 2025 年 Q1 实现营业收入 24.00 亿元，yoy +50.76%；归母净利润 5.06 亿元，yoy +75.33%。公司毛利率自 19 年以来持续增长，2025Q1 进一步提升至 61.69%，为历史较高水平。在信创政策持续推动与国产各大模型推出引发的算力需求增长的背景下，海光产品需求有望持续提升。

► **以 x86 架构为基础，彰显国产之光。**海光基于 x86 架构构建国产 CPU，近年来伴随 AI 产业加速发展，带动服务器市场扩容，2024 年全球服务器市场已增长至 1036.5 亿美元，上游核心部件迎来新一轮增长机遇。x86 作为当前生态最为完善的主流服务器架构，2023 年在全球服务器市场中占比高达 86%，凭借良好的软硬件兼容性在产业链中具备显著优势，IDC 预测 2028 年中国 x86 服务器市场收入有望增长至 620 亿美元，成长空间广阔。在美国技术限制与国内信创政策引导下，CPU 等核心部件有望在释放更大市场需求。海光持续推动架构自主升级，自研 C86 架构在保持对 x86 高度兼容的同时，完成国密算法集成，在数据安全与隐私保护层面持续发力。产品参数迭代不断缩小与国际领先竞品的性能差距，已跻身国内高性能 CPU 阵营。此外，海光联合国内 CPU 产业链上下游企业共建“光合组织”，加速构建完善的国产 CPU 生态体系，并与近 5000 家合作伙伴建立深度合作，为国产算力自主可控战略目标奠定坚实基础。

► **国产 GPU 迈入新阶段，技术创新驱动国产替代提速。**在 DeepSeek 等大模型广泛落地的背景下，AI 在多元场景的加速渗透推动智能数据量快速增长，据 IDC 预计，2028 年智能算力规模将增长至 2781.9 EFLOPS，并带动云厂商与运营商纷纷加码算力投入，算力资本支出屡创新高。与此同时，信创政策加速落地叠加形成国内训推算力缺口，为国产 GPU 提供关键增长窗口。海光深算系列产品基于 ROCm 平台开发，具备良好的 CUDA 兼容性，开发环境友好。同时，GPGPU 架构面向 AI 与高性能计算优化，突出并行计算能力。在应用端，海光积极与 DeepSeek 模型展开深度适配，通过软件层优化增加算力效率并弥补与国际竞品在参数层面的差距。配合信创加持和市场对国产方案的接受度提升，海光 DCU 有望加快渗透，助推核心高端部件国产化替代进入新阶段。

► **投资建议：**我们预计公司 2025-2027 年将实现营收 138.22/193.41/262.94 亿元，对应现价 PS 为 23/17/12 倍，我们看好公司在研发端的持续投入所带来的产品性能、扩展性与兼容性优势，在国产化替代提速、AI 产业景气度持续提升的背景下，维持“推荐”评级。

► **风险提示：**研发进展不及预期风险；知识产权风险；客户集中度较高风险。

盈利预测与财务指标

项目/年度	2024A	2025E	2026E	2027E
营业收入 (百万元)	9,162	13,822	19,341	26,294
增长率 (%)	52.4	50.9	39.9	36.0
归属母公司股东净利润 (百万元)	1,931	3,066	4,357	5,992
增长率 (%)	52.9	58.8	42.1	37.5
每股收益 (元)	0.83	1.32	1.87	2.58
PE	167	105	74	54
PS	34	23	17	12

资料来源：iFind，民生证券研究院预测；（注：股价为 2025 年 05 月 21 日收盘价）

推荐

维持评级

当前价格：

138.85 元



分析师 方竞

执业证书：S0100521120004

邮箱：fangjing@mszq.com



分析师 吕伟

执业证书：S0100521110003

邮箱：lvwei_yj@mszq.com

分析师 李萌

执业证书：S0100522080001

邮箱：limeng@mszq.com

相关研究

- 海光信息 (688041.SH) 2025 年一季报点评：业绩高速增长，合同负债环比大幅提升体现需求景气度-2025/04/24
- 海光信息 (688041.SH) 2024 年年报点评：业绩符合预期，国产算力龙头有望加速成长-2025/03/04
- 海光信息 (688041.SH) 2024 年业绩预告点评：业绩符合预期，收入利润同比大增-2025/01/15
- 海光信息 (688041.SH) 公司事件点评：国产算力龙头深耕根技术，AI 时代自身价值持续强化-2024/12/24
- 海光信息 (688041.SH) 2024 年三季报点评：高端处理器研发进展顺利，业绩表现亮眼-2024/10/16

目录

1 CPU+DCU 双轮齐发，铸造国产算力中坚力量	3
1.1 自主创新突破壁垒，国产 CPU&DCU 迈向前沿	3
1.2 股权分散稳定，科研国资背景浓厚	4
1.3 营收保持强劲增长，研发驱动价值释放	5
1.4 领导层技术背景深厚，带领团队稳健发展	8
2 以 x86 架构为基础，彰显国产 CPU 新风采	9
2.1 x86 架构领衔市场增长，国产化布局推动自主创新	9
2.2 打造国产高端 CPU，融合三大优势	13
3 国产 GPU 发展新篇章，技术创新寻求突破	18
3.1 AI 技术引爆 GPU 市场，国产厂商砥砺前行	18
3.2 构建国产 GPU 新标杆，成就国内计算新高度	24
4 盈利预测与投资建议	27
4.1 盈利预测假设与业务拆分	27
4.2 费用率预测	28
4.3 估值分析及投资建议	28
5 风险提示	30
插图目录	32
表格目录	32

1 CPU+DCU 双轮齐发，铸造国产算力中坚力量

1.1 自主创新突破壁垒，国产 CPU&DCU 迈向前沿

海光信息是国内领先的高端计算芯片设计企业，围绕 CPU 和 DCU 产品的双重布局。海光信息成立于 2014 年，是一家专注于高性能计算芯片和服务器研发的高科技企业，主要产品包括高端通用处理器（CPU）以及高端协处理器（DCU）。公司的 CPU 产品早年获得 AMD 在高端处理器的技术授权及相关技术支持，并基于 X86 架构研发，定位于高性能计算市场，涵盖服务器、工作站、高性能终端等多种应用场景；DCU 产品以 GPGPU 架构为基础，兼容“类 CUDA”通用环境，面向数据中心和云计算场景，解决传统架构在大规模并发、低延迟和高带宽需求下的瓶颈，随着 AI 算力时代的到来，DCU 产品发展前景广阔。

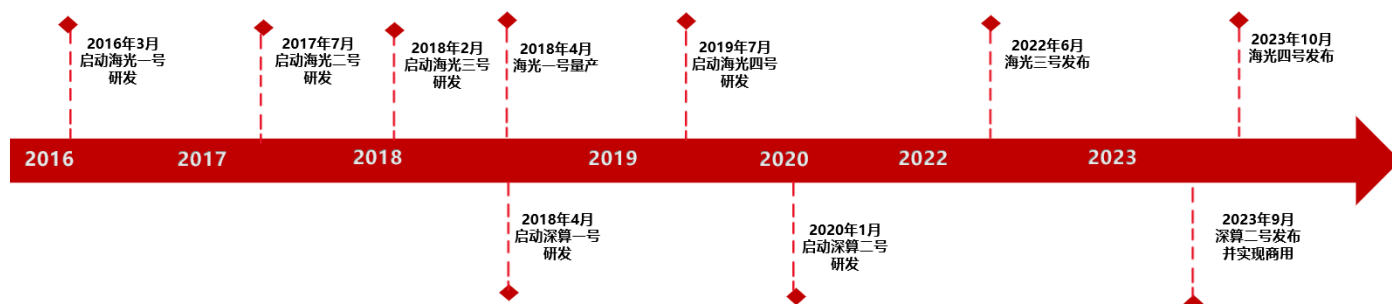
表1：海光核心产品系列介绍

产品类型	处理器类型	指令集	主要产品	产品特征	典型应用场景
海光 CPU	通用处理器	兼容 x86 指令集	海光 3000 系列	内置多个核心处理器，继承通用的高性能外设接口，拥有完整的软硬件生态环境和完备的系统安全机制，适用于数据计算和事务处理等通用性应用	云计算，物联网，信息服务等
			海光 5000 系列		
			海光 7000 系列		
海光 DCU	协处理器	兼容“类 CUDA”环境	海光 8000 系列	内置大量运算核心，具有较强的并行计算能力和较高的能效比，适用于向量计算和矩阵计算等计算密集型应用	大数据处理、人工智能、商业计算等

资料来源：海光信息招股书，民生证券研究院

公司发展历程可分为两个阶段，分别以“海光一号”CPU 和“深算一号”DCU 的启动设计为标志。在 CPU 研发进程中，2016 年公司启动了“海光一号”的研发，基于 AMD 授权技术进行产品设计，并于 2018 年成功实现量产与盈利。随着对市场对更高算力需求的预判，公司在 2017 年 7 月启动了“海光二号”的研发。该产品通过核心调整优化了性能和能效，于 2020 年实现量产。随后，2018 年公司开始研发“海光三号”，专注于满足 AI 大算力的新兴需求，并于 2022 年推出市场。2023 年 10 月，海光成功发布“海光四号”，该产品进一步提升了算力，并顺利进入市场。目前，公司正积极推进“海光五号”的研发，开发进程顺利。

在 DCU（深算系列）研发进程上，2018 年海光启动了“深算一号”的研发，产品主要面向大数据处理、商业计算等计算密集型应用，并于 2021 年完成结项并实现盈利。2020 年 1 月，海光启动了“深算二号”的研发，产品提升了浮点数据精度和常见数据计算类型的能力，性能提升超过 100%，并于 2023 年 9 月成功发布。2023 年，海光启动了“深算三号”的研发，产品在性能、兼容性和安全性方面进行了全面提升，相关开发进展顺利。同时，深算四号的研发也在积极推动中。

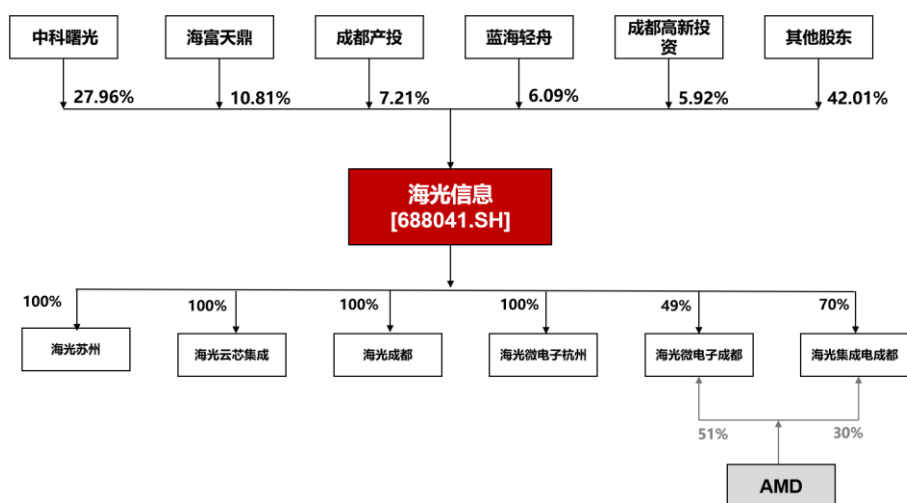
图1：海光信息 CPU 及 DCU 产品系列研发进程


资料来源：海光信息招股书，海光信息官网，民生证券研究院

海光信息始终秉持“销售一代、验证一代、研发一代”的产品研发策略，建立了完善的高端处理器的研发环境和流程，确保产品性能稳步提升，功能应用广泛丰富。同时，公司一直专注于 DCU+CPU 的双赛道开发，根据市场需求调整和持续产品迭代，在国产算力市场一直保持核心地位。

1.2 股权分散稳定，科研国资背景浓厚

中科院及国资基金为主要股东，股权较为分散。截至 2025 年一季度，海光信息技术股份有限公司股东结构分散，单一股东持股比例均未超过 30%，其中最大股东为中科院背景的曙光信息产业股份有限公司，持股比例 27.96%。此外，股东包括天津海富天鼎科技合伙企业（10.81%）、成都产业投资集团（7.21%）、成都蓝海轻舟企业管理合伙（6.09%）等，股东背景涵盖国有资本、产业基金及科研机构，科研与国资背景浓厚，为公司长期发展提供稳定支持。而海光与 AMD 拥有合资企业成都海光微电子有限公司与成都海光集成电路设计有限公司，为相关集成电路技术引进和和合作奠定基础。

图2：海光信息股权结构图（截至 2025 年 Q1）


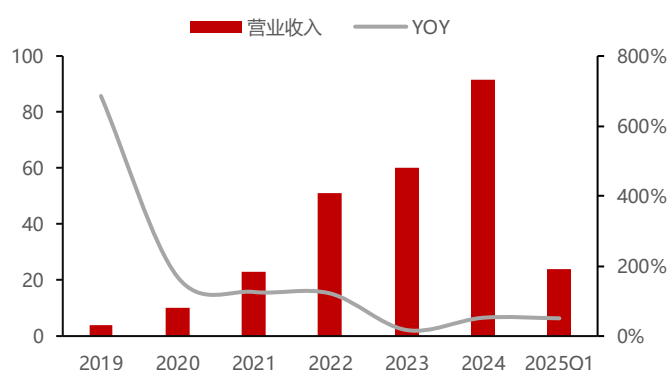
资料来源：iFind，民生证券研究院

1.3 营收保持强劲增长，研发驱动价值释放

公司营业收入总体呈现上升趋势。公司营业收入自 2019 年的 3.79 亿元大幅增长至 2024 年的 91.62 亿元，期间 CAGR 达 89.09%。增长的核心驱动在于产品迭代持续推进以及信创政策在各大关键行业国产化覆盖升级，并使公司于 2024 年创下历史营收新高。尽管 2023 年可能因信创需求疲软影响，CPU 业务增速承压，导致当年营收同比增速放缓至 17.30%，但受益于高端处理器产品生态持续完善及算力市场需求快速增长，2024 年公司实现收入同比增长 52.40%。而在公司最新公布的 2025 年第一季度财报中，海光信息实现营业收入 24.00 亿元，同比增长 50.76%；受益于地缘政治推动下的国产化进程及公司产品在国产大模型中的快速适配，整体需求加速持续加大并使得公司业绩继续维持高增长态势。同时公司 25Q1 合同负债 32.37 亿元，较去年末大增 23.33 亿元，订单需求持续旺盛，后续业绩释放具备较强增长动能。

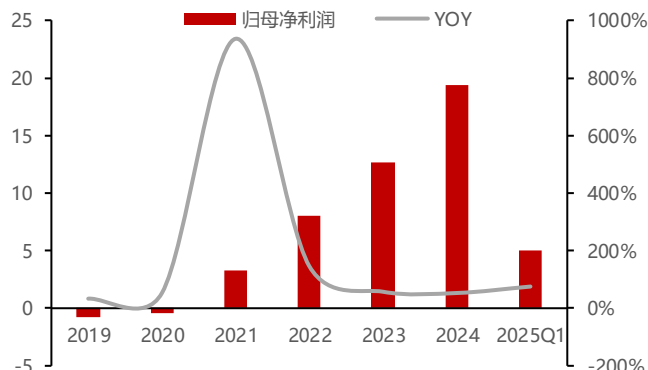
归母净利润方面公司改善明显，2019-2024 年，公司归母净利润从-0.83 亿元增长至 19.31 亿元。自 2021 年实现扭亏为盈起，公司净利润保持快速提升，2021-2024 年间 CAGR 达 80.91%，并在 2024 年公司归母净利润达到 19.31 亿元，同比增长 52.87%，创历史新高，增长因素主要包括高性能产品占比提升带动整体毛利率上升及国产化进程拉动业绩释放。而在 25 年 Q1，公司归母净利润达到 5.06 亿元，同比增长 75.33%，技术壁垒逐步成型，产品优势明显驱动公司利润大增。

图3：2019-2025Q1 公司营收（亿元）及增速



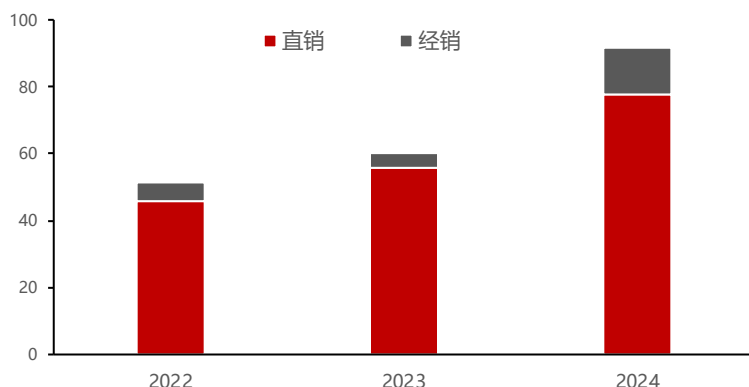
资料来源：iFinD，海光信息年报，民生证券研究院

图4：2019-2025Q1 公司归母净利润（亿元）及增速



资料来源：iFinD，海光信息年报，民生证券研究院

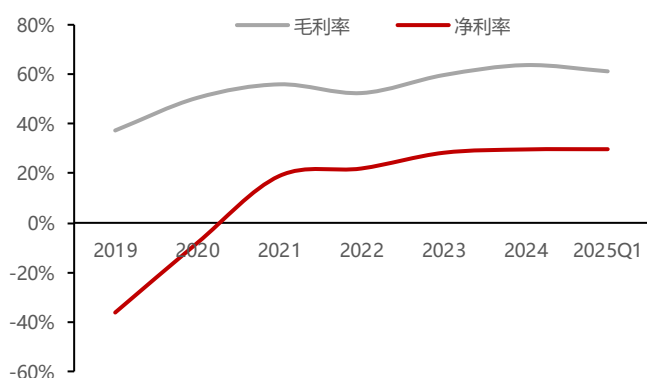
从收入结构来看，海光信息的高端处理器销售业务仍占主导，并以直销模式为主要销售渠道。公司 2024 年主要收入来源于高端处理器的销售，占总营业收入 99.70%。按销售模式拆分，2024 年直销业务实现收入 77.69 亿元，同比增长 39.36%，占比 84.79%；经销业务收入 13.93 亿元，同比增长 218.57%，占比 15.21%。目前，海光仍以直销模式为主，但经销业务的快速增长主要受下游经销商核心客户规模扩张推动，反映出公司在多元销售渠道的拓展能力不断增强。

图5：海光信息 2022-2024 年销售模式拆分（亿元）


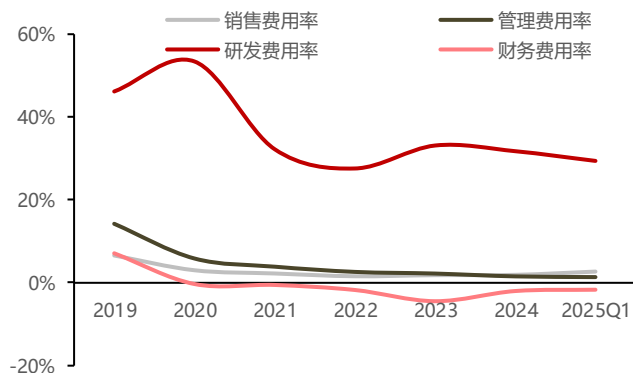
资料来源：海光信息年报，民生证券研究院

公司盈利能力持续增强，毛利率呈上升趋势，由 2019 年的 37.31%提升至 2024 年的 63.72%。尽管 2022 年受封装测试成本上涨，毛利率出现短暂下滑，但随着供应链管理效率提升、成本结构优化及规模化生产效应，2024 年毛利率同比+4.05 pct，创历史新高。在毛利率改善带动下，净利率同步上行，2024 年净利率达 29.65%，同比+1.35pct，同样为历史最高水平。高端芯片产品放量、产品结构优化及供应链协同能力增强是公司净利润持续增长的关键驱动因素。而 25 年 Q1 公司销售毛利率实现 61.69%，净利率 29.74%，成本与公司管理协同调整动能公司盈利能力持续上升。

随着产品销售规模拉大，顾客关系成熟及管理架构逐步优化，同时财务结构持续改善，利息收入增长叠加部分借款偿还，带动公司费用率改善明显。2019 至 2025 年 Q1，公司管理费用率由 14.25%下降至 1.38%，而销售费用率则由 6.60%下降至 2.71%，同时财务费用率则由 7.12%下降至-1.67%。

图6：2019-2025Q1 公司毛利率和净利率


资料来源：iFinD，海光信息年报，民生证券研究院

图7：2019-2025Q1 公司费用率


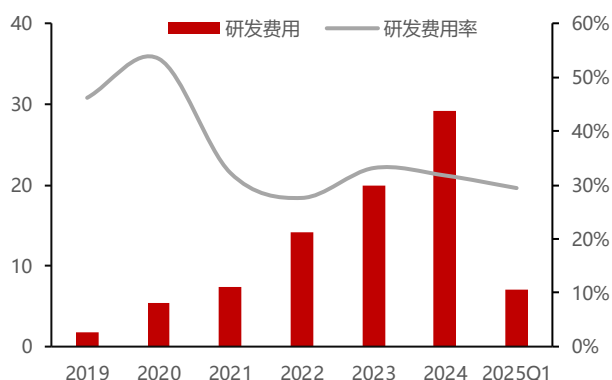
资料来源：iFinD，海光信息年报，民生证券研究院

海光信息持续注重技术创新研发。近年来公司持续加大研发投入力度，从 2019-2024 年，研发费用已经从 1.75 亿元增长至 29.1 亿元，期间研发费用率始终稳定在 25%以上，而 25Q1 研发投入达 7.64 亿元，同比增长 16.26%。公司基于长期发展战略扩大研发团队，2024 年研发人员数量较上年同期增长 31%。新产

品研发进程加速并增加研发项目投入，以强化技术储备并提升市场竞争力。同时，费用化项目较多，导致研发费用增长较快。海光围绕通用计算市场，持续保持高强度的研发投入，不断实现技术创新、产品性能提升，保持了国内领先的市场地位。

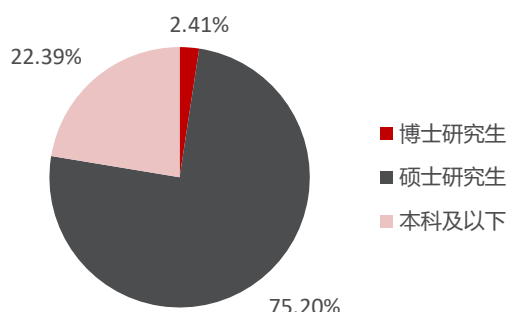
海光信息研发人员具备高学历及丰富研发经验。截至 2024 年，公司研发技术人员 2,157 人，占员工总人数的 90.18%，77.61%的研发技术人员拥有硕士及以上学历，同时在具有高学历的基础上，海光工程师对于芯片和处理开发均具有丰富的行业履历，并且拥有对于 x86 和 ARM 处理器的经验。随着高强度研发投入以及核心技术的积累，公司的技术“护城河”越来越巩固，产品组合也逐渐丰富。

图8：2019-2025Q1 公司研发费用（亿元）及研发费用率



资料来源：iFind，海光信息年报，民生证券研究院

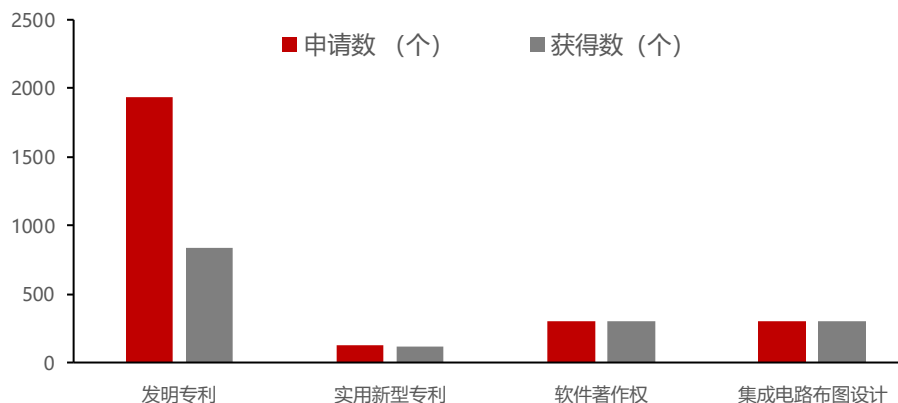
图9：2024 年公司研发人员学历构成



资料来源：海光信息 2024 年年报，民生证券研究院

公司通过多线程项目推进通用处理器和协处理器的持续迭代和产品拓展，并长期投入大量研发资金确保技术优势和市场竞争力。该策略不仅巩固了海光在国内高性能计算领域的领先地位，也在核心技术升级和产品布局上保持对国际龙头企业的紧密追赶。同时相应的研发投入支持也为海光带来了乐观的专利授权回报。截至 2024 年，海光信息累计取得发明专利 836 项、实用新型专利 115 项、外观设计专利 3 项、集成电路布图设计登记证书 299 项、软件著作权 304 项。

图10：知识产权累计获得数（截至 2024 年）



资料来源：海光信息 2024 年年报，民生证券研究院

1.4 领导层技术背景深厚，带领团队稳健发展

海光信息的核心团队技术背景深厚。学历上大多具有硕士及博士学位，不乏清华大学、北京大学和海外求学学历背景；同时公司高管也拥有相应的高级工程和企业背景。公司董事及总经理沙超群为教授级高级工程师，历任中科曙光技术副总裁，高级副总裁。

表2：海光信息核心团队成员介绍

姓名	职务	背景及经历
沙超群	董事，总经理	硕士，北京理工大学工学，教授级高级工程师。 2011年1月至2020年4月，历任中科曙光技术副总裁、高级副总裁； 2019年12月起任公司总经理，现任公司董事，总经理。
历军	董事	博士，北京大学高级管理人员工商管理硕士，教授级高级工程师。 2006年至今，任中科曙光董事、总裁； 2016年1月至今，兼任中科三清科技有限公司董事、董事长； 2017年12月至2019年6月，兼任中科可控董事； 2019年10月至2021年7月，兼任国科控股有限董事。 现任公司董事，总裁。
潘宇	副总经理	硕士，华中科技大学微电子与固体电子学。 2012年6月至2017年9月，任AMD芯片设计高级经理； 2017年9月至2017年11月，任武汉晟联智融微电子科技有限公司副总经理、执行董事。 2017年11月加入公司，现任公司副总经理，公司核心技术人员，主要负责海光DCU的研发、设计与实现，为海光协处理器研发和产品量产做出了重要贡献。
刘新春	副总经理	博士，中国科学院电子学研究所信号与信息处理专业。 2002年4月至2008年12月，任中国科学院计算技术研究所副研究员； 2009年1月至2016年2月，任中科曙光研发中心负责人。 2016年加入海光信息，现任公司副总经理、核心技术人员，主要负责海光高端处理器硅后验证测试、软硬件系统开发和产品量产运营等工作，为海光高端处理器研发和产品量产做出了重要贡献。
黄河	CPU设计中心主任工程师	博士，中国科学院计算技术研究所计算机系统结构专业。 2012年7月至2014年5月，任AMD主管工程师； 2014年6月至2015年7月，任英国想象技术有限公司深圳代表处高级主管工程。 2015年加入海光信息，现任公司CPU设计中心主任工程师、核心技术人员，并主要负责海光CPU的浮点协处理器研发工作，领导海光CPU浮点协处理器微结构设计、验证与实现，并进行前瞻性基础研究，为提升海光CPU的浮点计算性能做出了重要贡献。
杨晓君	工程平台技术中心主任工程师	博士，哈尔滨工程大学通信与信息系统专业。 2000年1月至2010年11月，任中国科学院计算技术研究所副研究员； 2010年11月至2016年9月，任中科曙光研发中心副总工程师。 2016年加入海光信息，现任公司工程平台技术中心主任工程师、核心技术人员，主要负责海光CPU和海光DCU的系统硬件研发工作，包括芯片封装设计、硬件设计、处理器验证测试平台开发和运维，为海光高端处理器的研发和产品量产做出了重要贡献。

资料来源：海光信息 2024 年年报，海光信息招股说明书，民生证券研究院

2 以 x86 架构为基础，彰显国产 CPU 新风采

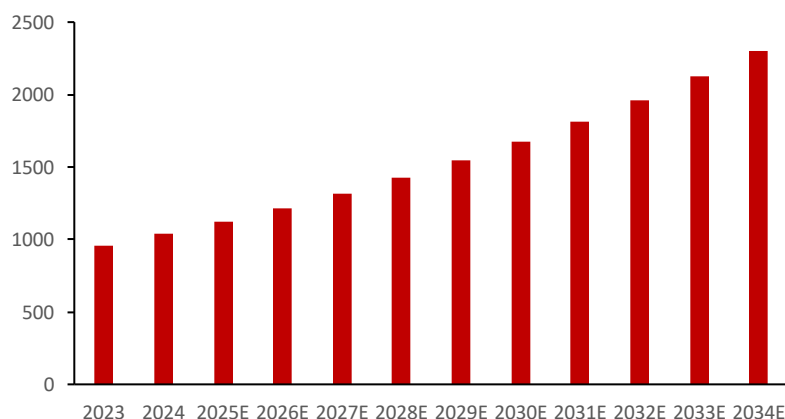
2.1 x86 架构领衔市场增长，国产化布局推动自主创新

2.1.1 全球市场平稳增长，x86 架构优势明显

CPU（中央处理器）是计算机的运算与控制核心，负责对存储器、输入输出单元等硬件资源进行调度，并执行通用计算任务。同时，计算机系统中各类软件的运行最终均需通过指令集转换为 CPU 操作，决定整体计算性能与系统效率。CPU 下游应用涵盖 PC、服务器、消费电子及物联网等多个领域。

而 CPU 下游核心服务器市场在 AI 需求带动下有望进一步扩张。据 Precedence Research 统计，2024 年全球服务器市场已增长至 1036.5 亿美元，并预计到 2034 年将达到约 2301 亿美元，2024-2034 年 CAGR 预计为 8.30%。受数据中心建设加速、云计算需求持续扩张及 AI 算力增长驱动，服务器市场有望保持稳定增长。

图11：2023-2034 年全球服务器市场（亿美元）



资料来源：Precedence Research，民生证券研究院

x86 架构符合服务器市场需求。目前 CPU 采用的指令集有复杂指令集和精简指令集两种，并根据所使用的指令集类型的不同，CPU 架构分为 x86、ARM、RISC-V 和 MIPS 四大架构。相较于其他架构而言，x86 更侧重于复杂指令采集，及单个指令即可借助指令数量存储多和超线程的优势来处理更为复杂的任务，并且相关架构设置的高性能能够支持更加高端处理的工作速度，同时在兼容性、硬件灵活拓展性及配套软件及开发工具成熟度、功能性上优势明显。并且 x86 架构能够高效利用主存储器，在处理复杂指令和商业计算任务时表现出色，同时许多先前的服务应用平台依赖于 x86 构架下的服务器支持，因而迁移至其他架构会产生不菲的技术匹配成本，因此 x86 被广泛应用于服务器领域并逐年实现增长。

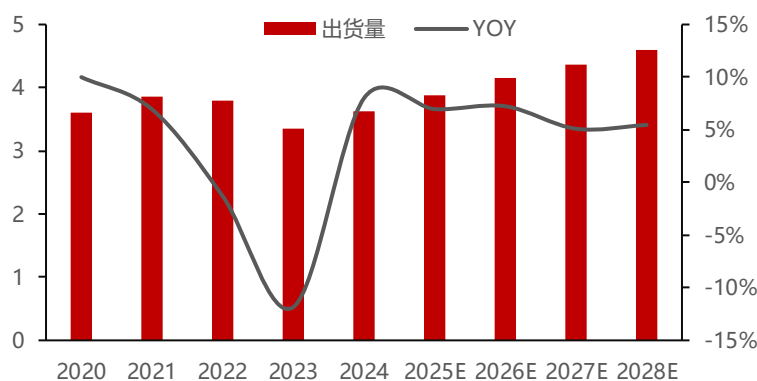
表3: CPU 四大架构对比

	x86	ARM	RISC-V	MIPS
指令集类型	复杂指令集	精简指令集	精简指令集	精简指令集
发布时间	1978 年	1983 年	2014 年	1981 年
特点	性能好, 速度快, 兼容性好	成本低, 功耗低	模块化, 极简, 可拓展	简洁, 优化方便, 高拓展性
代表厂商	Intel, AMD	苹果, 谷歌, 华为	三星, 英伟达	龙芯
应用领域	PC, 服务器	移动端	智能穿戴设备	网关, 机顶盒

资料来源: 华经情报网, 民生证券研究院

而 x86 服务器则占据服务器市场中的首要位置。据 IDC 统计, 2023 年 x86 服务器在全球服务器市场规模中的占比达到 86%, 依然是服务器中的主流选择。而在国产市场借助于 x86 服务器的对于平台的兼容性和灵活的扩展性, 打造的丰富的软件生态系统再加上可观的成本效益, 相应的服务器需求将会在国产化 AI 的浪潮化中持续走高, 并在 2024 年第三季度, 中国服务器市场出货量同比增长 16.6%。据 IDC 预计, 到 2028 年中国 x86 服务器市场出货量将达到 460 万台, 收入规模将增长至 620 亿美元。

图12: 2020-2028 年中国 x86 架构服务器出货量 (万台) 及增速

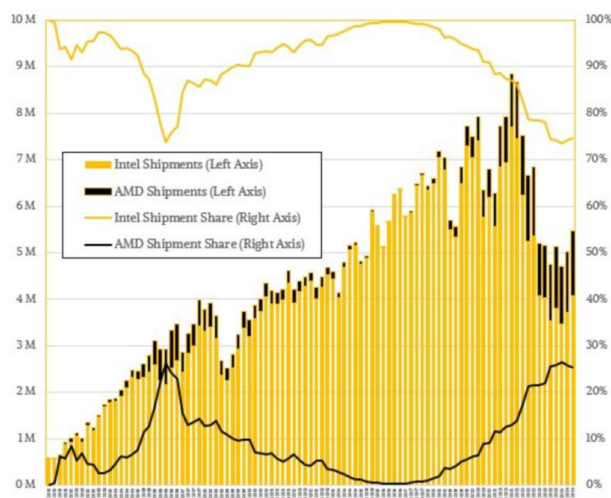


资料来源: IDC, 海光 2024 年财报, 民生证券研究院

2.1.2 海外巨头优势明显, 国产替代任重道远

海外公司垄断市场, Intel 依然保持霸主地位。根据智研咨询统计, 截至 2025 年初, 国产 CPU 市场英特尔和 AMD 分别占据 50%及 30%的份额, 外企占比依然较多并且行业垄断特征明显。Intel 在其中依然保持了绝对的领先地位, 根据 Mercury Research 数据显示, Intel 在 3Q24 出货 409 万颗 x86 服务器 CPU, AMD 出货 139 万颗 Epyc 处理器。尽管 AMD 的市场份额相较于 Intel 仍有一定差距, 但是凭借着过硬的产品力和性价比, 近年来 AMD 的市场份额逐渐上升, 未来 AMD 有可能与 Intel 平分秋色, 形成两强分庭抗礼的局面。

图13: Intel 和 AMDx86 服务器 CPU 出货量及市场份额变化情况

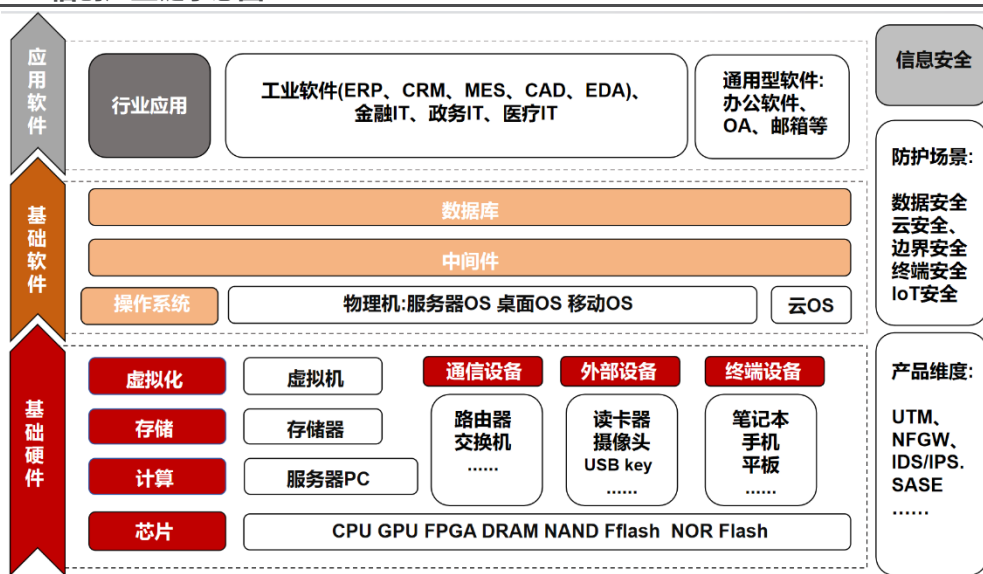


资料来源: Mercury Research, 民生证券研究院

2.1.3 国产化浪潮带动行业发展，自主创新势在必行

随着市场结构的演进与政策引导，CPU 正成为信息技术国产化的核心抓手，政企端在该领域的投入力度持续加大。尤其在 AI 与数字化渗透率不断提升的背景下，金融、教育等服务行业对信息安全与国产替代的需求愈发迫切。国家在“十四五”国家信息化规划明确提出推进信息化发展战略，强化以国内大循环为主体的数字政务建设，加快关键核心技术产品的研发与适配平台建设，推动软硬件协同适配，加速构建自主可控的信息基础设施。同时，国资委在 79 号文中明确提出，到 2027 年底央企国企 100%完成信创替代，涵盖基础软件、芯片、操作系统、中间件等关键领域，信创国产化进程加速。

图14: 信创产业链示意图

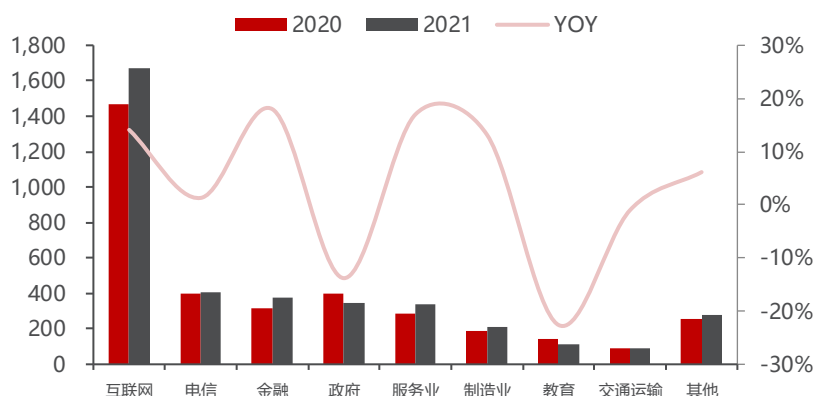


资料来源: 前瞻产业研究院, 民生证券研究院

在信创政策持续推进下,国产化进程加速,下游终端响应积极。IDC 数据显示,

2021 年我国 x86 服务器出货量达 382 万台，其中互联网企业采购 167 万台，占比 43.8%，电信企业采购 40.5 万台，占比 10.6%，显示其对服务器算力的强劲需求。当前信创国企已成为国产 CPU 的重点拓展方向。以中国电信为例，其 2024 年服务器集采项目国产化率已达 67.5%，中国移动与中国联通也在持续提升对国产芯片的采购占比。在政策与需求双重驱动下，国产 CPU 厂商将迎来广阔成长空间，海光有望凭借技术积累与产业协同优势，在新一轮国产替代进程中加速突围，确立行业领先地位。

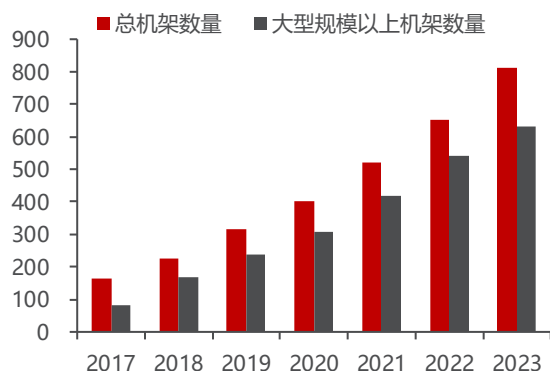
图15：2020-2021 年我国分行业 x86 服务器出货量（千台）及增速



资料来源：IDC，民生证券研究院

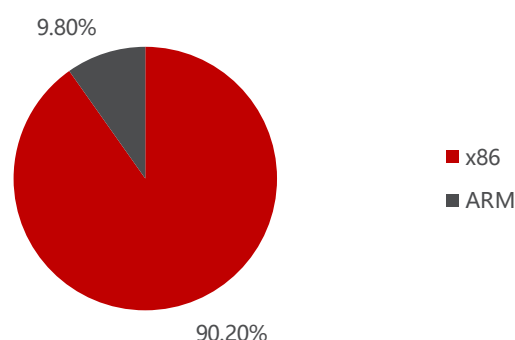
算力驱动产业升级，x86 仍是业界首选。对于日益增长的云计算需求，上游服务器供应链也享受到正面效应。据《中国算力发展报告（2024）》，2023 年我国在用算力中心机架数量达到 810 万架，较 2017 年的 166 万架增长接近 4 倍，其中大型规模以上机架占比达到 78.15%。而 x86 服务器在 23 年 CPU 服务器市场份额占比达到 90.20%。在科技革命和产业变革快速推进的背景下，计算能力正日益成为促进产业优化升级与社会发展的核心动力，x86 架构将继续发挥其关键作用，为我国算力的发展注入强大动力。

图16：2017-2023 年我国算力中心机架数量（万架）



资料来源：开放数据中心委员会，中国信息通信研究院，民生证券研究院

图17：2023 年我国 CPU 服务器市场份额分布



资料来源：开放数据中心委员会，中国信息通信研究院，Gartner，民生证券研究院

2.2 打造国产高端 CPU，融合三大优势

2.2.1 积极推动研发，融合三大优势

海光信息持续大力发展 CPU，产品具备三大优势。海光信息致力于在保证产品性能的基础上，研制出符合中国信息化发展要求的高端处理器。多年来，海光持续对 CPU 产品进行迭代改进，最终形成了集“性能、生态、安全”三位一体，满足中国用户需求的 CPU 产品：

1) 性能优异：海光 CPU 采用了尖端的微架构和缓存层次设计，通过优化分支预测技术，大幅提升了每个时钟周期内执行的指令数量。此外，基于先进的 SoC 架构和片上网络技术，海光信息在单颗 CPU 上集成了更多的处理器核心，并利用领先的工艺技术和物理设计手段，实现了处理器的高主频设计。

2) 兼容性好：x86 指令集在行业内拥有最强大的生态系统支持，公司成功掌握了 x86 处理器设计的关键技术，推出 C86 开放架构体系及产品，兼容国内外主流的操作系统、数据库、中间件等基础软件和众多行业应用软件。

3) 安全性高：海光 CPU 通过增加安全算法指令集、融合专门的安全算法加速单元等方式，大幅提升了数据处理和计算环境的安全级别，并且原生支持可信计算。

图18：海光信息高性能国产处理器



资料来源：海光信息官网，民生证券研究院

2.2.2 系列化产品布局，多维覆盖核心市场

海光 CPU 系列产品主要应用于服务器和 workstation，支持 x86 指令集，兼容国际主流操作系统和应用软件。海光 CPU 针对通用处理器应用场景，如复杂的逻辑运算和多任务处理，进行了深入优化，根据应用场景的不同，产品细分为 7000 系列、5000 系列和 3000 系列，分别对应高端、中端和低端需求。

表4：海光第三代 7000 系列、5000 系列、3000 系列 CPU 产品参数对比

	海光 7390	海光 7380	海光 5390	海光 5380	海光 3350
CPU 核心数量	32	32	16	16	8
线程数量	64	64	32	32	16
最高加速频 (GHz)	3.3	3.0	3.2	3.0	3.3
基准时钟频率 (GHz)	2.7	2.2	2.9	2.5	3.0
内存类型	DDR4	DDR4	DDR4	DDR4	DDR4
最高内存频率 (MHz)	3200	3200	3200	3200	3200
典型功耗 (W)	110	140	95	70	65

资料来源：海光信息官网，民生证券研究院

海光 7000 系列面向数据中心的旗舰级高性能处理器。其产品集成 16~32 核心，支持 128 路 PCIe 通道及 8 个 DDR4 内存通道，单颗 CPU 最大支持 2TB 内存容量。该系列针对数据中心、云计算中心进行了深度优化，为用户提供高能效的计算解决方案，广泛应用于云计算、大数据、数据库、分布式存储、人工智能等计算密集型领域。

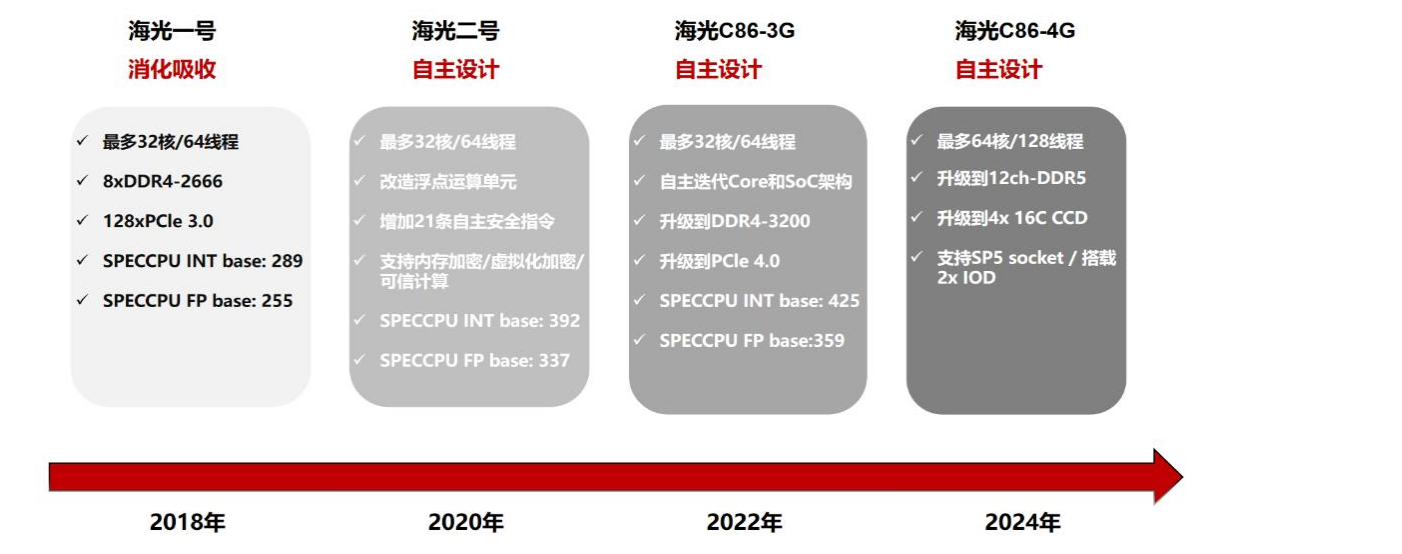
海光 5000 系列主要面向行业客户。其产品作为主流中端处理器，集成 8~16 核心，支持 64 路 PCIe 通道及 4 个 DDR4 内存通道，单颗 CPU 最大支持 1TB 内存容量。该系列适用于云计算、边缘计算、分布式存储等应用场景，满足互联网、金融、电信、交通、能源等行业的计算需求。

海光 3000 系列定位于多场景的高性价比处理器，广泛应用于入门级服务器、工作站、工业控制等市场。其集成 4~8 核心，支持 32 路 PCIe 通道及 2 个 DDR4 内存通道，最高加速频率达 3.3GHz，专为中小企业客户和专业用户提供高效、稳定的计算解决方案。

2.2.3 x86 技术自主化推进，自研 C86 加速性能突破

海光已全面掌握授权技术的整套源代码，并成功消化吸收 x86 相关全套技术，实现了独立的 CPU 设计与迭代开发。海光在微体系结构层面持续推进自主创新，基于完整的 x86 指令集源码实现 C86 架构的国产化研发，显著提升产品在计算性能、总线带宽、内存频率、安全性及可扩展性等方面的核心指标，持续强化产品竞争力。公司于 2022 年三季度正式推出 C86-3G 处理器，并在 2023 年实现规模化出货。2024 年发布的 C86-4G 处理器在架构性能进一步升级，支持最高 64 核心、128 线程，升级至 SP5 接口并兼容 DDR5 内存，配置最多两颗 IO 芯片，显著提升系统数据处理与调度能力，为信创行业提供具备高性能、强适配性的服务器算力解决方案。

图19：海光 CPU 产品升级路线



资料来源：赛迪项目组，Wccftch，民生证券研究院

而海光 CPU 自主迭代在更新微架构，包括计算能力的提升基础上，也增加安全应用方面能力，如对国密算法的支持，同时研究先进设计和封装技术，自主迭代核心与 SoC 结构，具备定制 IP 和开发封装基板的能力。同时，海光自研 C86 系列处理器也保持着每代至少 15%-30%的性能提升。

海光信息在国产厂商中表现出色并加速追赶国际竞品。2024 年 9 月，Intel 率先发布旗舰级服务器 CPU 产品至强 6 系列处理器 6900P，核心数较之前产品至少提升两倍，随后 AMD 也发布霄龙 5 代处理器 9005 系列产品，最大支持 192 核、384 线程，参数表现全方位领先于竞品。在国内厂商中，海光和兆芯基于 x86 架构，海思和飞腾基于 ARM 架构，龙芯和申威则是基于自研的指令系统。相较于其他厂商而言，“海光三号”CPU 在参数上处于国产领先地位，性能表现优秀。

表5：各厂商服务器 CPU 参数对比

	Intel	AMD	海光	兆芯	海思	飞腾	龙芯	申威
产品型号	Xeon6980P	EPYC9965	海光 7390	KH-4000/32	鲲鹏 920-7260	S5000C-64	3D5000	申威 1621
指令集	x86	x86	x86	x86	ARM	ARM	LoongArch	SW_64
核心数	128	192	32	32	64	64	32	16
超线程	256	384	64	不支持	不支持	不支持	不支持	不支持
主频 (GHz)	2.0	2.25	2.7	2.0	2.6	2.1	2.0	2.0
内存类型	DDR5	DDR5	DDR4	DDR4	DDR4	DDR5	DDR4	DDR3
内存通道数	12	12	8	8	8	8	8	8
最高内存频率 (MHz)	6400	6000	3200	3200	2933	4400	3200	2133
PCIe 版本	5.0	5.0	4.0	3.0	4.0	5.0	/	3.0
PCIe 通道数	96	128	128	128	40	96	/	16

资料来源：各公司官网，海光信息招股说明书，民生证券研究院

注：龙芯 3D5000 并未披露 PCIe 相关参数信息

作为国产处理器的重要技术突破，海光 C86 正加速向行业主流标准迈进，进一步夯实其在国产计算生态中的核心地位。2024 年中央政府采购网发布《关于更

新中央国家机关台式计算机、便携式计算机批量集中采购配置标准的通知》，明确将海光 CPU 纳入配置标准，并要求其性能需达到或高于 C863350 规格。这一举措不仅体现了 C86 处理器在性能、广泛软件生态兼容性及安全性方面的技术优势，也反映出其在国家级采购体系中的认可度进一步提升。

2.2.4 兼容 x86 应用框架，服务提供实现全方面覆盖

海光 CPU 兼容国际主流 x86 架构，沿用 CISC 处理器技术路线，具备卓越的系统架构与广泛的软硬件生态。在主流指令集开发环境下，海光 CPU 支持平台的快速部署与扩展，同时凭借高度的灵活性，实现应用的平滑迁移，避免底层代码大规模重构，从而有效降低用户的总体拥有成本。基于 x86 架构，海光 CPU 在计算性能、生态适配性和安全性方面表现突出，赢得了国内市场的广泛认可，并获得联想、H3C、华硕等知名 OEM 厂商的支持。目前，海光产品已在电信、金融等关键行业实现规模化应用，并获得招商银行、中国邮政等采购订单，助力信创行业数字化转型与算力升级。

图20：海光 OEM 客户



资料来源：海光信息官网，民生证券研究院

为推动国产计算技术供应链与产业链的完善，海光产业生态合作组织（简称光合组织）应运而生。目前，光合组织已与近 5000 家合作伙伴建立深度合作，并在全国范围内布局 28 个生态适配中心、24 个区域分会及专业委员会。这一体系不仅提升了国产软硬件的兼容性与整体性能，也为国产计算产业的长期发展奠定了坚实基础，进一步增强了本土计算生态的自主可控能力。

光合组织在上下游领域与国内各厂商建立紧密联系。芯片领域与国产厂商深度合作，推动 CPU、GPU 及存储技术发展，并助力国产固态硬盘与大容量存储系统的应用。在服务器与网络设备方面，支持同方、联想开天等品牌创新，满足大规模计算与数据传输需求，已获金融、能源、电信等行业认可。在中间层软件上，光合组织积极推广统信、麒麟等国产操作系统，推动其与底层硬件深度适配，并支持达梦等国产数据库的应用落地。同时，推动中间件、虚拟化平台等关键基础软件的国产化进程。在应用软件层，光合组织推动 ERP、CRM 等企业级软件国产化，并在人工智能、大数据分析等前沿领域促进国产解决方案发展，构建高水平商业体系。

图21：光合组织合作伙伴



资料来源：光合组织官网，民生证券研究院

3 国产 GPU 发展新篇章，技术创新寻求突破

3.1 AI 技术引爆 GPU 市场，国产厂商砥砺前行

3.1.1 GPGPU 架构：高效并行计算的核心驱动

GPGPU 的编程模型旨在充分发挥 GPU 的并行计算能力，以加速人工智能、科学计算和数据分析等通用计算任务。相比传统图形渲染，GPGPU 的软件体系基于 GPU 硬件特性，为高效计算提供支持。现代 GPU 包含大量计算核心（core），这些核心以单指令多线程（SIMT）模式执行 kernel 程序。SIMT（单指令多数据）采用单一控制器操控多个计算单元，同时处理向量数据，实现空间并行计算。GPU 上的每个核心通常可运行上千个线程，并通过寄存器通信，利用高速屏障同步。同时，每个核心内含一级指令和数据缓存（L1cache），减少对低级内存的访问需求，优化延迟隐藏。

为保持高吞吐量计算，GPU 需在计算能力与内存带宽之间保持平衡，依靠多通道内存系统提升并行性。通常，每个内存通道与对应的 LLC（最后一级缓存）相关联，核心通过片上互连网络（如 crossbar）与内存分区连接。与超标量乱序 CPU 相比，GPU 将更多芯片面积用于计算单元，减少控制逻辑占比，从而提高单位面积性能，并在并行算力计算上赋予更多性能提升。

图22：现代 GPU 普遍架构

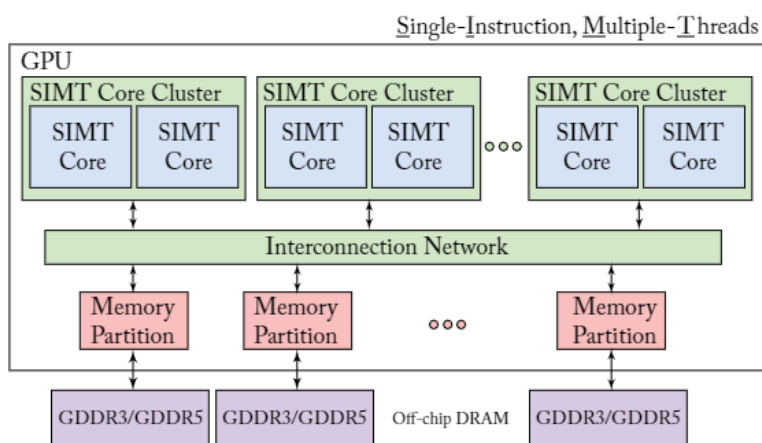


Figure 1.2: A generic modern GPU architecture.

资料来源：General-Purpose Graphics Processor Architecture，民生证券研究院

相比之下，另一种主流架构 ASIC 在研发投入和灵活性方面受到较大限制。ASIC 专为特定应用优化，在逻辑计算单元、控制单元比例及芯片架构上进行精确定制，使其在特定任务中具备极高的计算效率和低功耗，达到了性能和能效的极致。然而，这种高度定制化的设计需要额外的电路布局和架构开发，导致研发成本高昂，并对设计能力提出更高要求。

此外，ASIC 的灵活性较低，一旦设计完成，难以适应新的计算需求或算法变更。而 GPGPU 架构受益于成熟的软硬件生态和广泛的开发者支持，使其在开发周期、适配性和软件兼容性方面更具优势，并凭借更短的开发周期和强大的通用计算能力，在 AI、大数据分析和科学计算等领域持续占据主导地位。

3.1.2 DeepSeek 带动 AI 产业发展势头迅猛

近几年，AI 技术的快速发展推动了国内市场对生成式 AI 的广泛应用，引发了新一轮的市场浪潮。DeepSeekR1 模型的推出在降低算力芯片需求门槛的同时提升了算力效率进一步拉动各大服务厂商的适配需求。其发布后迅速登顶全球 140 个市场的移动应用下载榜单。这一现象不仅得益于免费开源策略，更关键在于其在数学代码推理等核心能力上已能媲美 OpenAI GPT-4o 正式版。此外 DeepSeekR1 训练成本仅 557.6 万元，不到 GPT-4o 训练成本的十分之一。凭借极高性价比成为众多企业和开发者的优先适配模型进一步推动 AI 生态的发展。

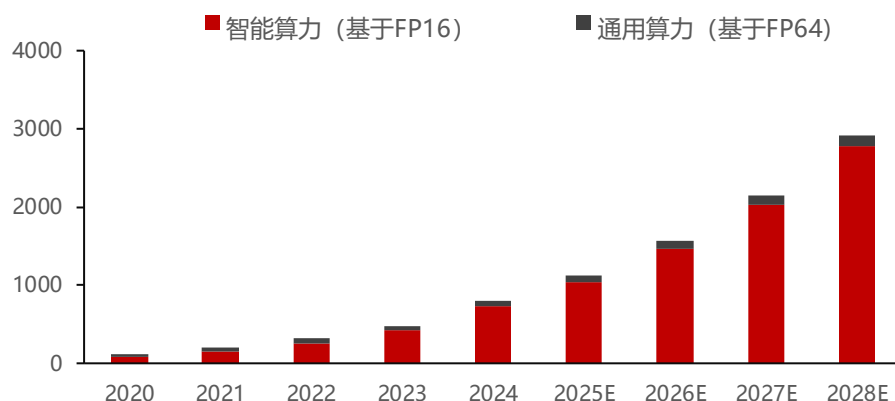
图23：DeepSeek 对于下游产业链的生态影响



资料来源：制造前沿，民生证券研究院

随着 DeepSeek 的推出并迅速普及，下游终端需求的增长及服务模式的演进推动了算力数据产业链迎来新的发展机遇。为满足 DeepSeek 的推理计算需求，云计算和算力租赁市场进入新一轮增长周期，算力需求占比从 30%跃升至 65%，同时加速了市场格局的变化，推动行业从高度依赖英伟达逐步向国产芯片方案过渡。同时国产芯片在集群部署 DeepSeek 方面，算力利用率已达英伟达 A100 的 92%，并且在同等算力水平下实现了 40%的成本降低，进一步增强了国产 AI 算力的竞争力。此外，DeepSeek 依托高效的模型压缩技术，使其能够在算力资源有限的边缘设备上高效运行，满足高隐私、低延迟的应用场景。与此同时，通过模型轻量化优化，DeepSeekR1 已能够在智能手机端运行，加速了终端设备 AI 计算的普及，进一步推动数据算力市场需求及算力数据规模的持续增长，据 IDC 预计，2028 年中国智能算力规模将增长至 2781.9 EFLOPS。

图24：2020-2028 年中国智能算力和通用算力规模（EFLOPS）

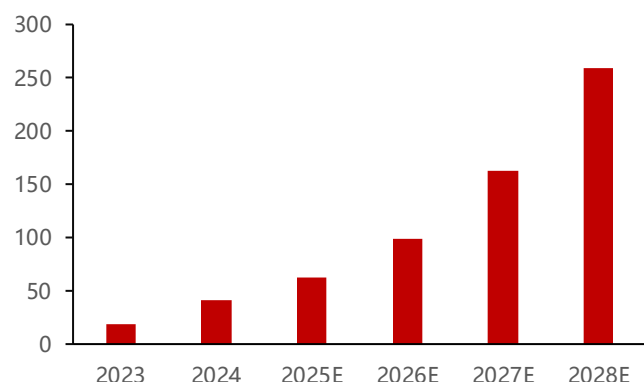


资料来源：IDC, 智东西, 民生证券研究院

国内 AI 产业前景广阔，加速技术突破与商业落地。近年来，AI 软件作为行业发展的核心驱动力，在政策支持、市场需求增长以及技术迭代的推动下，已从规模扩张阶段迈向质量升级。据中商产业研究院统计，2024 年国内 AI 软件市场规模已达 452.9 亿元，同比增长 20%，展现出强劲增长态势。

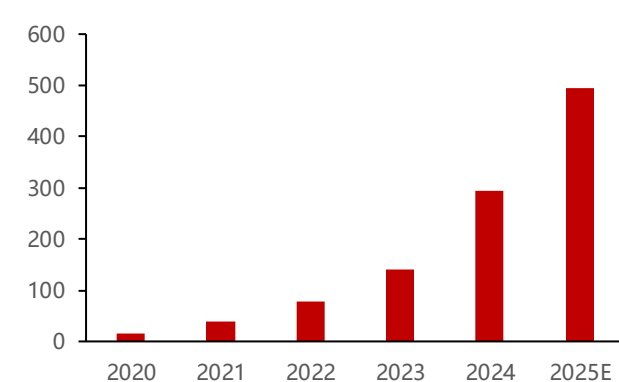
随着 DeepSeek 等大模型的开源与推广，生成式 AI、大模型、边缘 AI 成为市场增长的核心驱动力，并进一步推动新场景、新业态、新模式的不演化。据中商产业研究院预测，当前生成式人工智能产业正加速渗透各大行业，带动需求的快速提升，拓展 AI 在企业级应用领域的落地进程，预计 2025 年市场规模将达到 62.4 亿元；同时在技术层面，AI 大模型基于深度学习算法，能够处理大规模数据并具备更强的预测与决策能力，成为推动人工智能商业化落地的关键，并预计 2025 年相关市场规模将达到 495.39 亿元。未来，随着算法优化、算力基础设施完善及行业应用深化，AI 产业将迎来更广阔的发展空间。

图25：2023-2028 年中国生成式 AI 软件规模（亿元）



资料来源：中商产业研究院, IDC, 民生证券研究院

图26：2020-2025 年中国 AI 大模型市场规模（亿元）



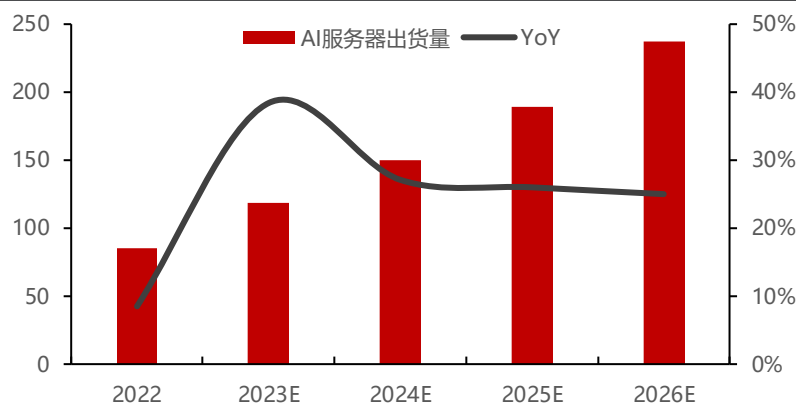
资料来源：中商产业研究院, IDC, 民生证券研究院

3.1.3 AI 服务器市场迅速扩张，政策驱动与云商运营商投资加码

AI 服务器在数据处理领域的战略地位持续提升。据 Trend Force 数据统计，2023 年全球 AI 服务器出货量预估已达 118.3 万台，预计到 2026 年将增长至

236.9 万台，2022-2026 年 CAGR 高达 22.6%。随着云计算、大数据、人工智能等前沿技术的广泛应用，数据规模呈指数级增长，传统服务器的计算能力逐渐逼近极限，无法满足日益增长的算力需求。相较于传统服务器，AI 服务器具备强大的并行计算能力，能够高效处理大规模数据集，显著提升计算性能，已成为下一代生产力计算基础设施的核心。

图27：2022-2026 年全球 AI 服务器出货量（万台）及增速



资料来源：Trend Force，民生证券研究院

而在外部政策制裁下，相关的技术限制也阻碍了国内 AI 产业发展。2025 年，美国出台更严苛的 AI 与半导体出口管控措施，明确将中国划分为“Tier3”级别，持续升级对高算力集成电路（如英伟达 H100）的出口限制，禁止训练超 1026 次计算操作的封闭式 AI 芯片出口，同时对晶圆代工领域实施更严格审批，确保 16nm 以下先进制程及 HBM 技术的封锁。美国在近几年里对于国内上游芯片供货的管制逐渐收紧，进一步影响国内服务器供应链市场稳定性。

国家出台相关政策缓冲制裁冲击。为了应对地缘政策紧张带来的政策性限制，如芯片和人工智能等的上下游生产的冲击，国家政府也出台相应政策鼓励算力发展国产化，以此缓冲外部影响带来的冲击并扶持国产算力市场的发展。

表6：国家主要信创政策汇总

发布年份	发布单位	发布政策与报告	算力产能相关政策条款解读
2023.8	工信部、财政部	《电子信息制造业 2023—2024 年稳增长行动方案》	推动先进计算产业发展和行业应用，开展先进计算在工业、城市管理等领域应用案例征集和应用对接，举办先进计算技术创新大赛等活动，加快先进技术和产品落地应用。鼓励加大数据基础设施和人工智能基础设施建设，满足人工智能、大模型应用需求
2023.12	发改委等五部门	《关于深入实施“东数西算”工程 加快构建全国一体化算力网的实施意见》	构建以东西部算力需求资源不平衡，开展相关的资源调配，即大力建设西部地区的大型数据中心算力设施，以满足东部地区对于紧缺数据处理的设施，同时建设西部也满足实惠的清洁和设备维护费用，以此为全国数字化经济高速发展打下筑基，同时达成东西部协调发展
2024.3	财政部	《关于 2023 年中央和地方预算执行情况与 2024 年中央和地方预算草案的报告》	中央财政产业基础再造和制造业高质量发展专项资金安排拨款，支持加快突破基础产品、核心技术等短板，增强产业链供应链韧性和竞争力。并优化投资基金功能，支持集成电路、人工智能、新一代信息技术等产业加快发展。

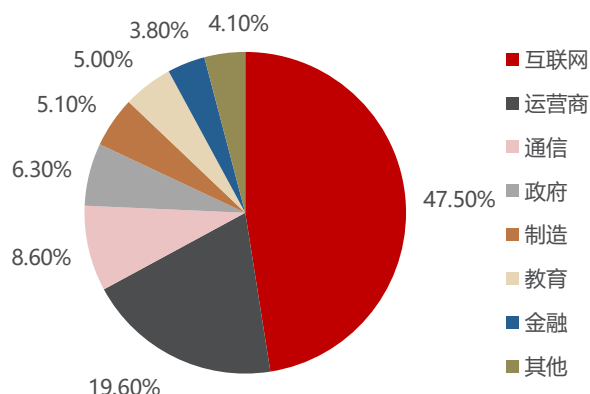
2024.7	工信部	《关于创新信息通信行业管理优化营商环境的意见》	开展算力互联互通技术研究和试点应用，推动公共算力资源标准化互联，加强算力统筹监测，打造智算生态圈，提升算力服务能力，助力传统产业智能化升级。
2024.9	中共中央办公厅、国务院办公厅	《关于加快公共数据资源开发利用的意见》	繁荣数据产业发展生态。聚焦算力网络和可信流通，支持数据基础设施企业发展。
2024.12	发改委等六部门	《关于促进数据产业高质量发展的指导意见》	提出打造全国一体化算力体系，发展通算、智算、超算等多元化算力资源，支持企业参与算力全产业链生态建设，构建一体化高质量算力供给体系，加强大带宽、低时延、高可靠的数据传输技术应用，加快算网融合等关键技术创新等。

资料来源：国家政府网，民生证券研究院

政策及市场双重驱动下，互联网算力投资需求拉大。2022 年我国 AI 服务器的下游应用市场中，互联网行业占比高达 47.5%，远超其他领域，而运营商、通信、政府等行业的市场份额也均超过 5%。得益于庞大的用户基数、政府推动行业数字化发展和不断扩展的 AI 应用场景，互联网企业持续加大算力投入，加速云计算及 AI 基础设施的建设。阿里巴巴在 2025 年 2 月业绩会上宣布，未来三年将投入超 3800 亿元，创下国内民营企业在 AI 和云计算领域的最高投资纪录，资金主要用于 AI 算力建设、云计算基础设施升级及原生 AI 应用拓展，并推动现有业务向 AI 技术深度融合。而腾讯近年来同样加快算力布局，2024 年全年资本开支突破 767 亿元，同比增长 221%，重点投向 AI 模型与 GPU 采购，在 AI 模型自研与适配开源模型的基础上，对应用层、模型层等进行全方位布局以此带动 AI 产业增长。

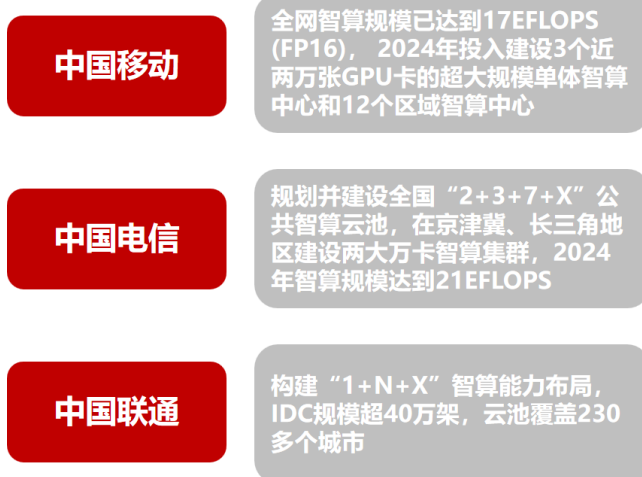
与此同时，三大运营商正在加速布局智算中心，推进业务从传统通信网络向 AI 算力及数智基础设施升级。以联通为例，公司目前已建成 300 多个训推一体的算力资源池，智算规模超过 17 EFLOPS。未来，运营商的投资重心由稳定基础的联网通信业务转向高增长的算网数智业务，借助云计算、AI 算力和边缘计算技术，提升整体服务能力。

图28：2022 年中国 AI 服务器下游应用市场分布



资料来源：观知海内咨询，民生证券研究院

图29：三大运营商智算中心规划

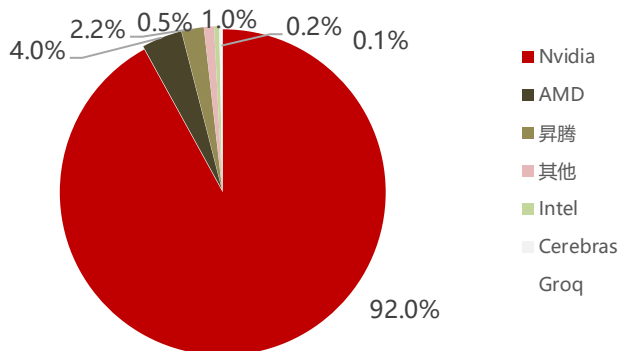


资料来源：三大运营商官网，数字中国建设峰会，民生证券研究院

3.1.4 英伟达主导 AI 芯片市场，国产技术奋力追赶

英伟达稳固 AI 芯片市场主导地位，市场份额持续扩大。2024 年，全球数据中心 GPU 市场规模已从 2022 年的 170 亿美元跃升至 1250 亿美元，主要受到云计算及 AI 算力需求激增的推动。而这一市场的增长几乎完全由英伟达主导，其凭借 CUDA 生态的垄断优势以及领先的硬件参数性能，持续满足全球对高性能 AI 计算的需求。英伟达推出的 H100 Tensor Core GPU 搭配 Blackwell 架构，在算力、能效和训练推理能力上大幅提升，进一步拉开与竞品的技术差距。得益于产品性能的领先性以及强大的软硬件整合能力，英伟达在 2024 年数据中心 GPU 市场的份额已达到 92%，继续稳居行业主导地位。

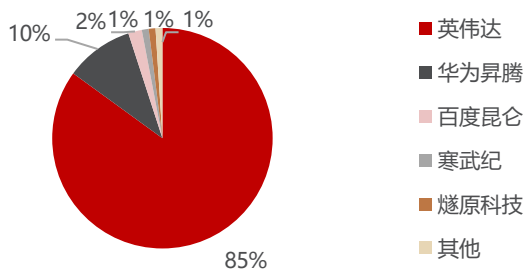
图30：2024 年全球数据中心 GPU 市场份额



资料来源：IOT Analytics，民生证券研究院

国产 AI 芯片替代率持续提升，技术差距仍待缩小。据 IDC 数据，2022 年中国 AI 芯片市场中，英伟达占据 85% 份额，而华为昇腾、百度昆仑、寒武纪、燧原科技等国产厂商合计市场份额仅达 14%。随着本土算力产业的快速发展，2024 年上半年我国 AI 芯片出货量已突破 20 万张，国产化率提升至 20%，国产替代进程持续加速。然而，在技术创新和产品迭代方面，国产厂商仍面临挑战。英伟达于 2024 年 3 月推出 B200 GPU，其计算性能相较 H100 大幅提升，进一步巩固领先地位，而 AMD 也在 2023 年 12 月发布 MI300X，加强在 AI 算力市场的竞争力。相比之下，国内企业在高端 AI 芯片的工艺制程、软件生态、算力优化等方面仍存在差距，未来仍需加快技术突破，以缩小与国际龙头企业的差距，实现更大规模的国产替代。

图31：2022 年中国 AI 芯片市场份额分布



资料来源：IDC，电子元件技术网，民生证券研究院

3.2 构建国产 GPU 新标杆，成就国内计算新高度

3.2.1 采用 GPUGPU 架构，兼容类 CUDA 生态，海光技术国内领先

海光 DCU 是一款 GPGPU 产品，兼容 ROCm GPU（类 CUDA）计算生态系统。基于传统 GPU 设计，GPGPU 去除了图形处理相关部件，保留了 SIMT 架构和通用计算单元，以高效满足 AI 训练、推理任务等通用计算需求，特别是在矩阵运算等数据密集型任务中，提供优异的计算性能。该架构有效解决了 CPU 通用算力的短缺，顺应人工智能技术发展的需求。

表7：GPUGPU 架构的计算优势

计算特征	计算优势
高效的并行性	GPU 通过多条流水线的并行计算实现高效计算能力。在当前主流 GPGPU 架构中，多条流水线既可以在单一控制部件的集中调度下运行，也可以独立执行任务。相比传统并行计算机，GPGPU 在降低硬件成本的同时，为适配 GPGPU 并行架构的应用提供了更优的并行计算解决方案。
高密集的运算	GPGPU 集成高速 GDDR 或 HBM 内存，具备 TB 级访存带宽，在数据密集型计算中表现优异，有效提升吞吐能力并减少内存瓶颈。
超长流水线	GPGPU 超长流水线的设计以吞吐量的最大化为目标，在对大规模的数据流并行处理方面具有明显的优势

资料来源：海光信息招股说明书，民生证券研究院

得益于 ROCm 与 CUDA 在生态环境和编程环境等方面高度相似，海光 DCU 拥有丰富的软硬件生态。公司围绕 DCU 产品积极推进开源生态建设，持续向开源社区贡献适配与优化方案，重点围绕人工智能框架、编程环境等领域开展研发，实现与国际主流的商业计算和人工智能软件兼容，并提升产品在主流软件生态中的可用性。通过自主研制一批具备国际影响力的基础软件与应用工具，海光已在金融、电信、交通等国民经济重点行业实现核心系统的基本自主可控，为高端处理器产品在开源生态体系中的落地应用奠定坚实基础。

图32：海光完整软件栈系统



资料来源：海光信息官网，民生证券研究院

3.2.2 性能达到同期国际高端产品水平，实现多领域应用

海光信息在持续深耕 CPU 领域的同时，也在 DCU 产品线上投入巨大研发力

度，打造出具备显著优势的 DCU 产品。英伟达 A100 是英伟达于 2020 年发布的高端 GPU，是当期国内商用性能最强的 GPU。与 A100 同期发布 GPU 有 AMD 的 MI100 和海光的“深算一号”，与两个国际品牌相比，“深算一号”在性能方面不落下风，相关指标均达到同等水平。目前海光信息 DCU 主力产品为“深算二号”，性能较“深算一号”提升 100%以上。海光 DCU 拥有自主开放的完整软件栈，与国内大模型全面适配，具备以下三大技术优势：

- 1) 强大的计算性能：**海光 DCU 依托于大规模并行计算微结构，在双精度、单精度、半精度、整型计算方面表现出色，拥有强劲的性能和较高的能效比。
- 2) 高效的并行处理能力：**海光 DCU 内置高带宽内存芯片，显著提升了在大规模数据计算任务中的数据处理效率，使得海光 DCU 能够适应多种不同的应用场景
- 3) 健全的软件生态环境：**海光 DCU 基于 GPGPU 架构，支持“类 CUDA”环境，在软件生态兼容性上表现良好，同时公司积极参与开源项目，确保 DCU 能够与 GPGPU 主流开发平台适配。

表8：英伟达、AMD、海光同期 GPU 参数对比

	英伟达	AMD	海光
产品型号	A100	MI100	深算一号
生产工艺	7nmFinFET	7nmFinFET	7nmFinFET
核心数量	2560CUDAprocessors 640Tensorprocessors	120CUs	4096 (64CUs)
内核频率	Upto1.53Ghz	Upto1.5Ghz (FP64) Upto1.7Ghz (FP32)	Upto1.5Ghz (FP64) Upto1.7Ghz (FP32)
显存容量	80GBHBM2e	32GBHBM2	32GBHBM2
显存位宽	5120bit	4096bit	4096bit
显存频率	3.2GHz	2.4GHz	2.0GHz
显存带宽	2039GB/s	1228GB/s	1024GB/s
TDP	400W	300W	350W
CPUtoGPU 互联	PCIeGen4x16	PCIeGen4x16	PCIeGen4x16
GPUtoCPU 互联	NVLink Upto600GB/s	InfinityFabricx3 Upto276GB/s	xGMIx2 Upto184GB/s

资料来源：海光信息招股说明书，民生证券研究院

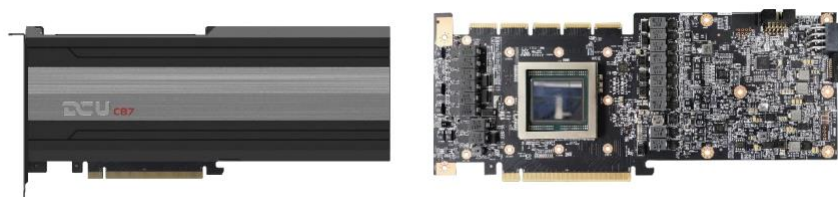
目前，海光信息的主力 DCU 产品为“深算二号”。产品具备全精度浮点数与多种标准整型数据类型的计算能力，已实现对主流通用计算任务的覆盖。在此基础上，公司进一步推出“深算二号”AI 专用版，针对深度学习模型训练与推理场景进行架构优化，增强对 FP16 和 INT8 等低精度计算格式的支持，提升 AI 运算效率并有效弥补“深算二号”至“深算三号”之间在智能计算能力上的过渡缺口。

3.2.3 DeepSeek 适配加速落地，进一步提升海光性能表现

自 2 月 2 日 DeepSeek V3 和 R1 模型与海光信息 DCU 达成适配，海光信息的硬件场景优势进一步显现。DeepSeek R1 基于 Transformer 架构，采用 Multi-Head Latent Attention (MLA) 和 MoE 两大核心推理技术，MLA 通过减少 KV 缓存降低内存占用并提升推理效率，MoE 则借助辅助损失优化专家负载平衡，提

高模型灵活性与计算效率，同时降低计算开销。海光信息 DCU 基于 Romc 生态开发，在 CUDA 类代码迁移方面表现突出，使其能够满足 DeepSeek 应用需求并发挥高效计算能力。在 DeepSeek 侧重于科学计算及 HPC+AI 融合的场景适配上，海光在此类应用中的竞争力进一步增强，借助 DeepSeek 模型优化行业解决方案，有望提升产品价值并强化市场渗透。

图33：海光 DCU 产品形态



资料来源：海光信息招股说明书，民生证券研究院

与此同时，DeepSeek 的 MTP（多令牌预测技术）技术优化数据处理流程，使海光 DCU 在单位时间内处理更大规模文本数据。FP8 混合精度训练技术与海光 DCU 架构深度结合，在硬件层面提升数据传输与计算效率，减少跨层数据搬运时间，缩短模型训练周期。在边缘计算与物联网融合的 AI 算力场景下，能耗与体积约束进一步收紧，海光 DCU 凭借轻量化设计、优化后的低功耗特性及强化的学习兼容能力，在这一领域具备更强市场竞争力。

4 盈利预测与投资建议

4.1 盈利预测假设与业务拆分

分业务进行营收预测：

1) 高端处理器收入：公司高端处理器业务为核心收入来源，历史占比稳定 99.5% 以上，核心产品涵盖 CPU（通用处理器）与 DCU（协处理器）双主线。CPU 产品基于 x86 架构自主研发，持续推动 C86 体系演进，产品在性能、兼容性、扩展性等维度持续升级，已实现与主流国际厂商的较小差距。依托公司产品迭代升级及信创政策驱动，公司在国产服务器 CPU 市场市占率已超 50%。DCU 产品聚焦 AI 推理、大数据等场景，搭载类 CUDA 生态，配套开发工具链逐步完善，具备良好的适配与拓展能力。受益于国内算力基础设施投入加速及 DeepSeek 等国产大模型应用落地，DCU 出货有望提速。在中美政策摩擦加剧背景下，具备自主可控属性的算力芯片需求将进一步提升。公司 CPU+DCU 双赛道驱动需求加速释放，叠加公司 2024 年高端处理器领域业绩表现亮眼，贡献营收 91.34 亿元，同比大增 51.94%，基于此，我们预计公司 2025-2027 年高端处理器业务将延续保持高增长态势，收入分别为 137.93/193.10/262.61 亿元，同比增长 51.00%、40.00%、36.00%。同时考虑到公司研发跟进，产品结构持续优化，随着核心产品在高端处理器市场渗透率不断提升，我们预计 2025-2027 年毛利率将保持上行趋势，分别达到 64.00%、64.50%、64.70%。

2) 技术服务支持：此类收入被认定为公司为客户提供技术支持、技术咨询、技术开发等服务内容。2024 年相关收入大幅增长主要系公司开展定制服务结算所致，对于后期收入增长趋势判断并不具备参考性。同时考虑到此类收入在公司营收占比较少，并非公司发展核心方向，基于此我们预估相关服务收入在 2025-2027 年期间将继续保持较低水平，并达到 0.29/0.31/0.32 亿元，同比增长也同步保持低位水平，达到 5.00%/5.00%/5.00%。考虑到该业务结构较为稳定、未来调整空间有限，我们预计该业务毛利率也将保持平稳，分别为 70.90%/70.90%/70.90%。

考虑到公司核心业务高端处理器在国产化替代趋势下成长空间开阔，我们预测海光信息 2025-2027 年总体营业收入 138.22/193.41/262.94 亿元，同比增长 50.86%/39.93%/35.95%，毛利率分别为 64.01%/64.51%/64.71%。

表9：公司分业务销售预测（百万元）

	2023A	2024A	2025E	2026E	2027E
高端处理器	6011.81	9134.30	13792.79	19309.91	26261.48
YOY	18.67%	51.94%	51.00%	40.00%	36.00%
毛利率	59.67%	63.70%	64.00%	64.50%	64.70%
技术服务	0.19	27.85	29.24	30.70	32.24
YOY	-99.68%	14660.00%	5.00%	5.00%	5.00%
毛利率	70.86%	70.96%	70.90%	70.90%	70.90%
营业收入	60.12	9162.15	13822.04	19340.61	26293.72
YOY	17.30%	52.40%	50.86%	39.93%	35.95%
毛利率	59.67%	63.72%	64.01%	64.51%	64.71%

资料来源：iFind，民生证券研究院预测

4.2 费用率预测

销售费用方面，随着公司业务规模的持续扩大、市场渗透率不断提升以及下游客户关系日益稳固，销售渠道逐步成熟，销售费用率有望进一步得到优化。我们预计公司 2025-2027 年销售费用率将分别为 1.90%、1.86%、1.84%。

管理费用方面，公司近年来持续推进管理体系建设，通过制度规范与组织架构优化逐步建立起较为成熟的内部管理机制。在公司治理水平不断提升背景下，预计管理费用率将于 2025-2027 年实现稳步下降，分别为 1.40%、1.38%、1.34%。

研发费用方面，公司持续加大在核心技术与高端人才上的投入，力求提升核心产品竞争力。考虑到未来市场竞争仍将保持高强度，公司预计仍将维持较高的研发强度以支撑产品迭代。然而，得益于营业收入的快速增长，以及技术框架逐步成熟，研发费用率有望小幅下降。我们预计 2025-2027 年公司研发费用率将分别为 29.80%、29.70%、29.60%。

表10：公司费用情况预测（百万元）

项目/年度	2024A	2025E	2026E	2027E
销售费用	175.51	262.36	356.64	473.13
销售费用率（%）	1.92%	1.90%	1.86%	1.84%
管理费用	141.85	193.72	264.61	344.56
管理费用率（%）	1.55%	1.40%	1.38%	1.34%
研发费用	2910.00	4118.97	5744.16	7782.94
研发费用率（%）	31.76%	29.80%	29.70%	29.60%

资料来源：iFind，民生证券研究院预测

4.3 估值分析及投资建议

海光信息主营高端处理器产品，涵盖 CPU 与 DCU 两大核心方向。寒武纪聚焦于人工智能芯片，产品主要应用于云服务器及边缘计算场景，与公司 DCU 产品在应用领域高度契合，具备较强可比性；龙芯中科专注于 CPU 处理器，其产品广

泛服务于电子政务等领域,与公司 CPU 的应用方向相近,营收结构具有参考价值。基于此,我们选择寒武纪,龙芯中科作为可比公司,对比其 2025-2027 年间的 PS 均值分别为 60/43/32 倍。根据我们对海光信息当前股价对应 PS 分别为 24/17/12 倍,显著低于可比公司平均水平。考虑到海光信息高端处理器在性能、兼容性、扩展性等维度表现优异,同时在国产芯片需求不断扩容及海外技术封锁愈发严苛的背景下,我们看好公司收入和盈利能力持续提升,维持“推荐”评级。

表11: 可比公司 PS 数据对比

股票代码	公司简称	市价 (亿元)	营业收入(亿元)			PS (X)		
			2025E	2026E	2027E	2025E	2026E	2027E
688256.SH	寒武纪	2782.33	52.88	77.18	107.44	53	36	26
688047.SH	龙芯中科	489.22	7.30	10.11	13.02	67	48	38
可比公司均值						60	42	32
688041.SH	海光信息	3227.35	138.22	193.41	262.94	23	17	12

资料来源: iFind, 民生证券研究院预测; 注: 可比公司数据采用 iFind 一致预期, 股价时间为 2025 年 05 月 21 日

5 风险提示

1) 研发进展不及预期风险。公司聚焦于高端通用处理器领域，产品更新迭代及新架构开发需持续投入大量人力与资金。作为技术密集型产业，核心技术突破与产品性能达成存在一定不确定性，若未来研发方向判断失误，或关键技术、性能指标未达预期，将对公司研发成果转化及盈利能力带来不利影响。

2) 知识产权风险。公司已积累较丰富的知识产权储备，包括专利、软件著作权、集成电路布图设计及专有技术，系构筑核心竞争力的关键。未来若出现知识产权被侵权、不当使用或与第三方产生争议，可能对公司业务造成干扰。同时，国际市场在知识产权法律适用与判定标准方面存在差异，公司面临一定的境外合规与潜在诉讼风险，或对业务拓展造成一定影响。

3) 客户集中度较高风险。公司客户主要集中于国内头部服务器厂商，客户集中度相对较高。受服务器行业集中度提升及头部客户议价能力较强影响，若主要客户经营出现波动或采购策略调整，且公司未能及时拓展新客户或实现客户结构优化，可能对公司业务稳定性与收入持续性带来不利影响。

公司财务报表数据预测汇总

利润表 (亿元)	2024A	2025E	2026E	2027E
营业总收入	9,162	13,822	19,341	26,294
营业成本	3,324	4,974	6,864	9,280
营业税金及附加	118	166	232	263
销售费用	176	262	360	484
管理费用	142	194	267	352
研发费用	2,910	4,119	5,744	7,783
EBIT	2,600	4,381	6,217	8,520
财务费用	-182	-52	-43	-84
资产减值损失	-96	-138	-181	-219
投资收益	7	14	19	26
营业利润	2,789	4,309	6,098	8,411
营业外收支	-4	0	0	0
利润总额	2,784	4,309	6,098	8,411
所得税	67	86	122	168
净利润	2,717	4,223	5,976	8,242
归属于母公司净利润	1,931	3,066	4,357	5,992
EBITDA	4,003	5,828	7,690	9,996

资产负债表 (亿元)	2024A	2025E	2026E	2027E
货币资金	8,797	6,485	8,472	11,541
应收账款及票据	2,275	3,567	4,938	6,640
预付款项	1,240	1,840	2,540	3,433
存货	5,425	6,770	9,352	12,669
其他流动资产	469	3,479	3,522	3,576
流动资产合计	18,207	22,141	28,823	37,860
长期股权投资	0	0	0	0
固定资产	537	577	607	627
无形资产	4,123	4,133	4,148	4,162
非流动资产合计	10,353	11,023	11,053	11,094
资产合计	28,559	33,164	39,876	48,954
短期借款	1,800	1,500	1,500	1,500
应付账款及票据	735	967	1,335	1,804
其他流动负债	1,853	2,847	3,841	5,096
流动负债合计	4,388	5,314	6,676	8,401
长期借款	899	700	700	700
其他长期负债	620	671	671	671
非流动负债合计	1,519	1,370	1,370	1,370
负债合计	5,908	6,684	8,046	9,771
股本	2,324	2,324	2,324	2,324
少数股东权益	2,401	3,558	5,178	7,428
股东权益合计	22,652	26,480	31,830	39,183
负债和股东权益合计	28,559	33,164	39,876	48,954

资料来源：公司公告、民生证券研究院预测

主要财务指标	2024A	2025E	2026E	2027E
成长能力 (%)				
营业收入增长率	52.40	50.86	39.93	35.95
EBIT 增长率	84.45	68.49	41.92	37.04
净利润增长率	52.87	58.76	42.11	37.54
盈利能力 (%)				
毛利率	63.72	64.01	64.51	64.71
净利率	21.08	22.18	22.53	22.79
总资产收益率 ROA	6.76	9.24	10.93	12.24
净资产收益率 ROE	9.54	13.37	16.35	18.87
偿债能力				
流动比率	4.15	4.17	4.32	4.51
速动比率	2.53	2.45	2.45	2.52
现金比率	2.00	1.22	1.27	1.37
资产负债率 (%)	20.68	20.15	20.18	19.96
经营效率				
应收账款周转天数	73.99	74.59	77.08	78.02
存货周转天数	351.95	441.33	422.79	427.16
总资产周转率	0.36	0.45	0.53	0.59
每股指标 (元)				
每股收益	0.83	1.32	1.87	2.58
每股净资产	8.71	9.86	11.47	13.66
每股经营现金流	0.42	1.48	1.80	2.39
每股股利	0.17	0.27	0.38	0.53
估值分析				
PE	167	105	74	54
PB	15.9	14.1	12.1	10.2
EV/EBITDA	79.40	54.54	41.33	31.80
股息收益率 (%)	0.12	0.19	0.28	0.38

现金流量表 (亿元)	2024A	2025E	2026E	2027E
净利润	2,717	4,223	5,976	8,242
折旧和摊销	1,403	1,447	1,473	1,476
营运资金变动	-3,171	-2,457	-3,519	-4,468
经营活动现金流	977	3,444	4,195	5,548
资本开支	-945	-1,430	-1,474	-1,488
投资	0	-2,904	0	0
投资活动现金流	-3,988	-4,979	-1,455	-1,462
股权募资	0	0	0	0
债务募资	1,758	-301	0	0
筹资活动现金流	932	-778	-753	-1,017
现金净流量	-2,079	-2,313	1,987	3,069

插图目录

图 1: 海光信息 CPU 及 DCU 产品系列研发进程	4
图 2: 海光信息股权结构图 (截至 2025 年 Q1)	4
图 3: 2019-2025Q1 公司营收 (亿元) 及增速	5
图 4: 2019-2025Q1 公司归母净利润 (亿元) 及增速	5
图 5: 海光信息 2022-2024 年销售模式拆分 (亿元)	6
图 6: 2019-2025Q1 公司毛利率和净利率	6
图 7: 2019-2025Q1 公司费用率	6
图 8: 2019-2025Q1 公司研发费用 (亿元) 及研发费用率	7
图 9: 2024 年公司研发人员学历构成	7
图 10: 知识产权累计获得数 (截至 2024 年)	7
图 11: 2023-2034 年全球服务器市场 (亿美元)	9
图 12: 2020-2028 年中国 x86 架构服务器出货量 (万台) 及增速	10
图 13: Intel 和 AMDx86 服务器 CPU 出货量及市场份额变化情况	11
图 14: 信创产业链示意图	11
图 15: 2020-2021 年我国分行业 x86 服务器出货量 (千台) 及增速	12
图 16: 2017-2023 年我国算力中心机架数量 (万架)	12
图 17: 2023 年我国 CPU 服务器市场份额分布	12
图 18: 海光信息高性能国产处理器	13
图 19: 海光 CPU 产品升级路线	15
图 20: 海光 OEM 客户	16
图 21: 光合组织合作伙伴	17
图 22: 现代 GPU 普遍架构	18
图 23: DeepSeek 对于下游产业链的生态影响	19
图 24: 2020-2028 年中国智能算力和通用算力规模 (EFLOPS)	20
图 25: 2023-2028 年中国生成式 AI 软件规模 (亿元)	20
图 26: 2020-2025 年中国 AI 大模型市场规模 (亿元)	20
图 27: 2022-2026 年全球 AI 服务器出货量 (万台) 及增速	21
图 28: 2022 年中国 AI 服务器下游应用市场分布	22
图 29: 三大运营商智算中心规划	22
图 30: 2024 年全球数据中心 GPU 市场份额	23
图 31: 2022 年中国 AI 芯片市场份额分布	23
图 32: 海光完整软件栈系统	24
图 33: 海光 DCU 产品形态	26

表格目录

盈利预测与财务指标	1
表 1: 海光核心产品系列介绍	3
表 2: 海光信息核心团队成员介绍	8
表 3: CPU 四大架构对比	10
表 4: 海光第三代 7000 系列、5000 系列、3000 系列 CPU 产品参数对比	14
表 5: 各厂商服务器 CPU 参数对比	15
表 6: 国家主要信创政策汇总	21
表 7: GPU/GU 架构的计算优势	24
表 8: 英伟达、AMD、海光同期 GPU 参数对比	25
表 9: 公司分业务销售预测 (百万元)	28
表 10: 公司费用情况预测 (百万元)	28
表 11: 可比公司 PS 数据对比	29
公司财务报表数据预测汇总	31

分析师承诺

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登记为注册分析师，基于认真审慎的工作态度、专业严谨的研究方法与分析逻辑得出研究结论，独立、客观地出具本报告，并对本报告的内容和观点负责。本报告清晰准确地反映了研究人员的研究观点，结论不受任何第三方的授意、影响，研究人员不曾因、不因、也将不会因本报告中的具体推荐意见或观点而直接或间接收到任何形式的补偿。

评级说明

投资建议评级标准		评级	说明
以报告发布日后的 12 个月内公司股价（或行业指数）相对同期基准指数的涨跌幅为基准。其中：A 股以沪深 300 指数为基准；新三板以三板成指或三板做市指数为基准；港股以恒生指数为基准；美股以纳斯达克综合指数或标普 500 指数为基准。	公司评级	推荐	相对基准指数涨幅 15%以上
		谨慎推荐	相对基准指数涨幅 5% ~ 15%之间
		中性	相对基准指数涨幅-5% ~ 5%之间
		回避	相对基准指数跌幅 5%以上
	行业评级	推荐	相对基准指数涨幅 5%以上
		中性	相对基准指数涨幅-5% ~ 5%之间
		回避	相对基准指数跌幅 5%以上

免责声明

民生证券股份有限公司（以下简称“本公司”）具有中国证监会许可的证券投资咨询业务资格。

本报告仅供本公司境内客户使用。本公司不会因接收人收到本报告而视其为客户。本报告仅为参考之用，并不构成对客户的投资建议，不应被视为买卖任何证券、金融工具的要约或要约邀请。本报告所包含的观点及建议并未考虑获取本报告的机构及个人的具体投资目的、财务状况、特殊状况、目标或需要，客户应当充分考虑自身特定状况，进行独立评估，并应同时考量自身的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专家的意见，不应单纯依靠本报告所载的内容而取代自身的独立判断。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而导致的任何可能的损失负任何责任。

本报告是基于已公开信息撰写，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、意见及预测仅反映本公司于发布本报告当日的判断，且预测方法及结果存在一定程度局限性。在不同时期，本公司可发出与本报告所刊载的意见、预测不一致的报告，但本公司没有义务和责任及时更新本报告所涉及的内容并通知客户。

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、咨询服务等相关服务，本公司的员工可能担任本报告所提及的公司的董事。客户应充分考虑可能存在的利益冲突，勿将本报告作为投资决策的唯一参考依据。

若本公司以外的金融机构发送本报告，则由该金融机构独自为此发送行为负责。该机构的客户应联系该机构以交易本报告提及的证券或要求获悉更详细的信息。本报告不构成本公司向发送本报告金融机构之客户提供的投资建议。本公司不会因任何机构或个人从其他机构获得本报告而将其视为本公司客户。

本报告的版权仅归本公司所有，未经书面许可，任何机构或个人不得以任何形式、任何目的进行翻版、转载、发表、篡改或引用。所有在本报告中使用的商标、服务标识及标记，除非另有说明，均为本公司的商标、服务标识及标记。本公司版权所有并保留一切权利。

民生证券研究院：

上海：上海市浦东新区浦明路 8 号财富金融广场 1 幢 5F；200120

北京：北京市东城区建国门内大街 28 号民生金融中心 A 座 18 层；100005

深圳：深圳市福田区中心四路 1 号嘉里建设广场 1 座 10 层 01 室；518048