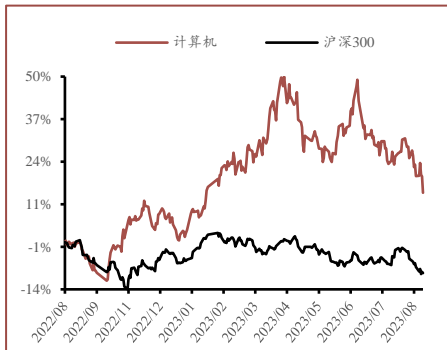


华为产业系列深度：凤凰磐涅，AI 昇腾

■ 证券研究报告

投资评级:看好(维持)

最近 12 月市场表现



分析师 杨焱

SAC 证书编号: S0160522050001
yangye01@ctsec.com

分析师 罗云扬

SAC 证书编号: S0160522050002
luoyunyang@ctsec.com

相关报告

- 《大模型推理算力知多少?》
2023-08-23
- 《“活跃资本市场”政策频出, 证券 IT 迎板块性机遇》
2023-08-17
- 《智能驾驶研究框架: (一) 总章》
2023-08-12

核心观点

- ❖ **自研高算力 AI 芯片, 构建全产业链生态。**昇腾计算产业是基于昇腾系列处理器和基础软件构建的全栈 AI 计算基础设施、行业应用及服务, 已发展为包括昇腾系列芯片、硬件、CANN、AI 计算框架、应用使能、开发工具链、管理运维工具等全产业链的完整体系。昇腾平台采用“一平台双驱动”模式, 坚持“硬件开放, 软件开源, 使能合作伙伴”战略, 截止 2023 年 7 月, 昇腾已认证了 30 多家硬件伙伴、1200 多家软件伙伴, 联合打造了 2500 多个场景方案。与 300 多家高校院所合作, 每年培养超过 10 万名专业 AI 人才, 昇腾开发者突破 180 万人。
- ❖ **内部: 昇腾计算产业架构, 软硬件夯实底座。**硬件层面, 作为 ASIC 芯片, 昇腾系列 N 一旦契合对应算法和场景, 具有高效低成本特点, 其中昇腾 910 在 FP16 和 INT8 算力高于英伟达 A100。软件层面, 昇腾 AI 产业链以 CANN 异构计算架构为底座, 支持 MindSpore 自研 AI 框架以及 TensorFlow 等第三方框架, 同时打造在应用使能平台, 助力行业开发者将 AI 计算融入各行各业, 目前已在富士康、中国移动等公司落地。2023 年 7 月, 昇腾推出首个万卡 AI 集群, 规模从最初的 4000 卡集群扩展至 16000 卡, 拥有更快训练速度和稳定训练周期。
- ❖ **外部: 联合合作伙伴, 打造产业生态。**昇腾与鲲鹏(通用计算)、盘古(多模态 AI 大模型)结合, 统合华为人工智能生态。对外, 昇腾生态伙伴网络目前已有宝德、华鲲振宇、长江计算、清华同方等合作伙伴。华为通过总经销商供货的销售支持、与总销售商双方的技术支持模式, 以及多种商务权益激励, 鼓励伙伴基于昇腾 AI 打造自有品牌的产品或解决方案, 实现共赢。目前昇腾开发者社区提供 100+种工具与样例和 50+种数据集、50+门直播课程和 70+门个人课程、20+名专家轮值服务旨在加速昇腾开发者成长、构建昇腾开发者生态。
- ❖ **投资建议:** 建议关注华为产业链: 整机合作伙伴: 神州数码、拓维信息、四川长虹、紫光股份; 垂直应用合作伙伴: 软通动力、赛意信息、东方国信、常山北明、格灵深瞳、云从科技、云天励飞、海量数据、润和软件、智洋创新。
- ❖ **风险提示:** AI 技术迭代不及预期的风险、商业化落地不及预期的风险、政策监管风险。

内容目录

1	自研高算力 AI 芯片，构建全产业链生态	4
2	内部：昇腾计算产业架构，软硬件夯实底座	6
2.1	硬件开放，高算力 AI 芯片提供澎湃动力	6
2.2	软件开源，简化 AI 场景开发及行业应用流程	10
3	外部：联合合作伙伴，打造产业生态	14
3.1	与鲲鹏处理器、盘古大模型结合，统合华为人工智能生态	14
3.2	使能外部合作伙伴，带动各行业“AI+”降本增效	17
3.3	深入高校与开发者社区，拓展昇腾 AI 影响力	19
4	投资建议	21
5	风险提示	21

图表目录

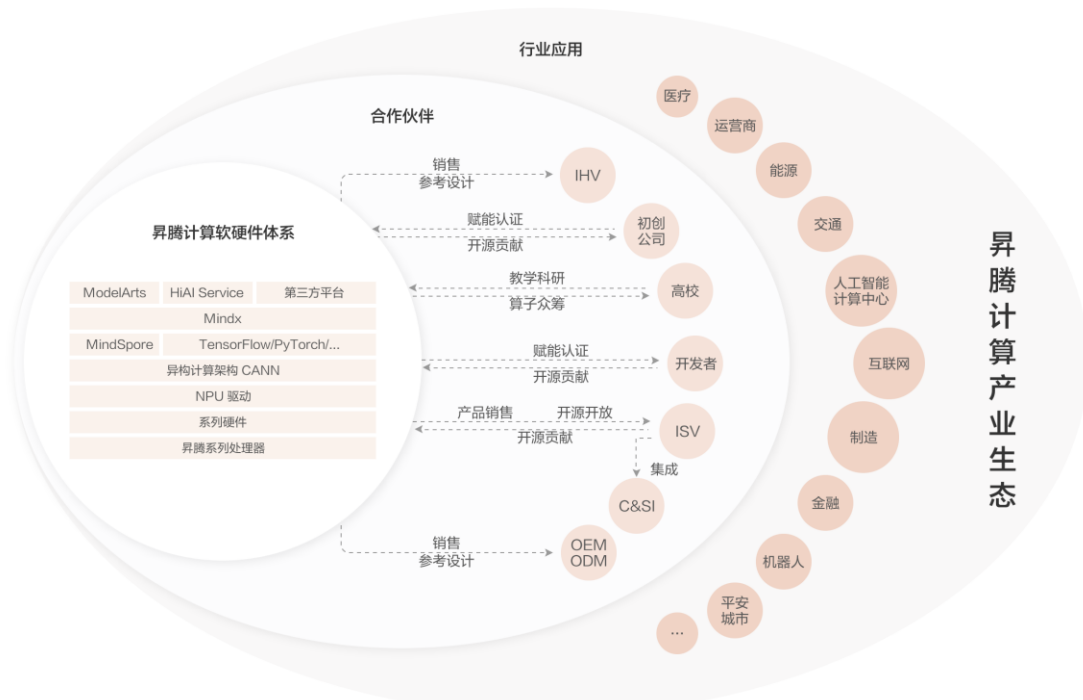
图 1.	昇腾计算产业生态	4
图 2.	昇腾计算架构	5
图 3.	昇腾一平台双驱动战略	5
图 4.	昇腾一体化解决方案发布	6
图 5.	华为智能基座育人基地	6
图 6.	昇腾 AI 芯片结构（昇腾 310）	7
图 7.	达芬奇 3D Cube 计算单元	8
图 8.	Atlas 系列集群基础单元特征	10
图 9.	AI 集群产品介绍	10
图 10.	CANN 异构计算架构	11
图 11.	昇思 MindSpore 四大优势	11
图 12.	MindStudio 开发流程	12
图 13.	MindX 应用使能	12
图 14.	鲲鹏计算产品	14
图 15.	华为开放昇腾 Atlas、鲲鹏主板等产品	14
图 16.	盘古大模型三层架构	15
图 17.	华为云盘古大模型 3.0 发布	16

图 18. 盘古大模型 3.0 全栈创新.....	16
图 19. 华为云盘古大模型生态计划.....	16
图 20. 业内首个千亿级三模态大模型—紫东太初.....	16
图 21. 昇腾生态伙伴网络.....	18
图 22. 昇腾兼容性认证硬件目录.....	18
图 23. 昇腾万里合作伙伴计划.....	19
图 24. 昇腾开发者社区.....	20
表 1. 昇腾处理器与英伟达 A800 处理器.....	7
表 2. 昇腾加速卡产品图.....	9
表 3. 昇腾 Atlas 系列集群基础单元.....	9
表 4. 昇腾 AI 解决方案及案例	13
表 5. 《生成式人工智能服务管理暂行办法》条例.....	17
表 6. 2023 年与昇腾相关的新成立人工智能中心.....	19
表 7. 昇腾科研创新使能计划享有权益.....	20

1 自研高算力 AI 芯片，构建全产业链生态

昇腾计算产业是基于昇腾系列处理器和基础软件构建的全栈 AI 计算基础设施、行业应用及服务。2018 年 10 月，华为全联接大会正式发布首款采用华为自研达芬奇架构的 AI 训练芯片昇腾 910 与 AI 推理芯片昇腾 310，并于 2019 年正式投入商用。2019 年 9 月，华为推出昇腾 AI 计算架构，包括昇腾处理器、昇腾 AI 加速模块和昇腾 AI 开发环境，初步形成 AI 计算解决方案。昇腾计算致力于完善 AI 计算产业生态，以自身软硬件体系为核心，联合制造、设计、集成、开发等软硬件合作伙伴与咨询、运营、人才培养等服务合作伙伴，投入各个行业应用实现生产力升级核心目标。

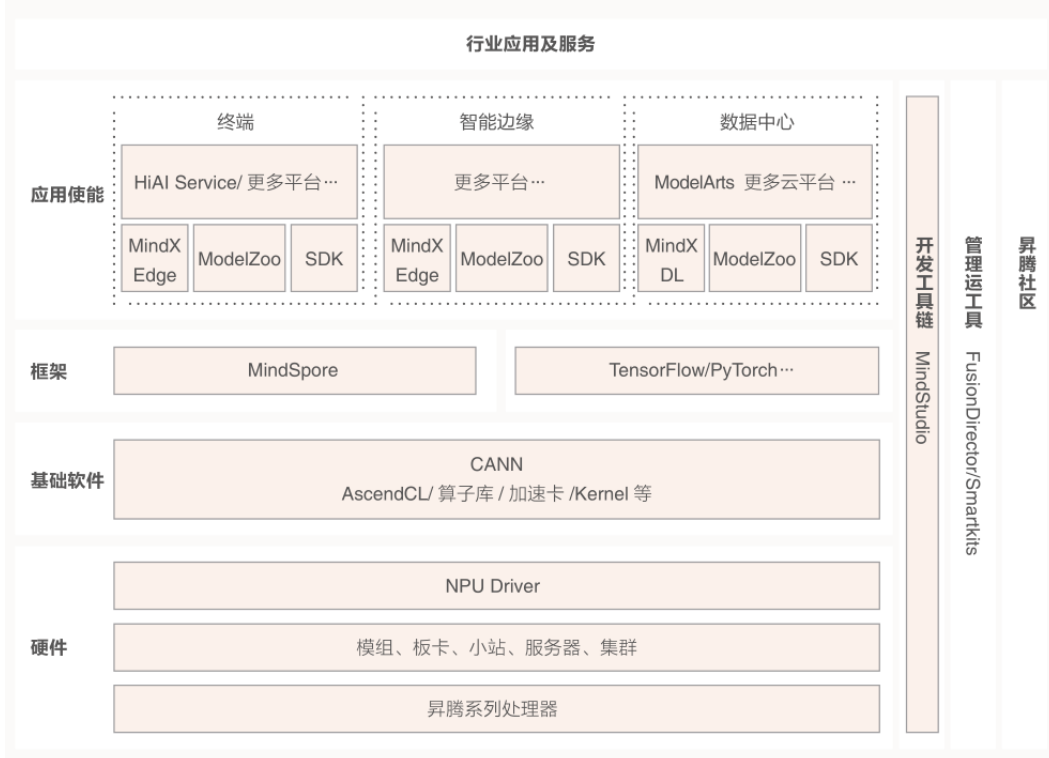
图1.昇腾计算产业生态



数据来源：《昇腾计算产业发展白皮书》华为\中国信通院等、财通证券研究所

昇腾计算架构由硬件、基础软件、应用使能与行业应用及服务四层构成。昇腾 AI 应用开发过程以芯片及处理器等硬件为基础，CANN（异构计算框架）承接硬件与计算框架，搭配自有开发与管理工具，联接应用使能以投入具体行业应用。目前昇腾计算产业已发展为包括昇腾系列芯片、硬件、CANN、AI 计算框架、应用使能、开发工具链、管理运维工具等全产业链的完整体系。

图2.昇腾计算架构



数据来源：《昇腾计算产业发展白皮书》华为\中国信通院等、财通证券研究所

昇腾平台以软硬件体系为基础，使能产业生态共建。昇腾平台采用“一平台双驱动”模式，其中“一平台”指基础软件平台，包括 AI 处理器、服务器硬件以及芯片使能软件、AI 框架，“双驱动”指平台要支撑的两大生态，即技术生态与商业生态。坚持“硬件开放，软件开源，使能合作伙伴”战略，聚焦自有计算架构、处理器和基础软件创新研发的同时发展软硬件与使能，共同开拓计算产业蓝海。

图3.昇腾一平台双驱动战略



数据来源：华为官网、财通证券研究所

联合伙伴创新，共筑智能根基。2023年7月6日，在世界人工智能大会 WAIC 上，华为携手伙伴联合发布昇腾 AI 大模型训推一体化解决方案，加速大模型在各行业应用落地，并有 23 家昇腾 AI 伙伴推出 AI 服务器、智能边缘与终端新品，共同为行业智能化升级提供丰富的产品与解决方案。截止 2023 年 7 月，昇腾已认证了 30 多家硬件伙伴、1200 多家软件伙伴，与 300 多家高校院所合作，每年培养超过 10 万名专业 AI 人才，昇腾开发者突破 180 万人。

图4.昇腾一体化解决方案发布



数据来源：昇腾官网、财通证券研究所

图5.华为智能基座育人基地



数据来源：昇腾官网、财通证券研究所

2 内部：昇腾计算产业架构，软硬件夯实底座

2.1 硬件开放，高算力 AI 芯片提供澎湃动力

全球首个覆盖全场景 AI 芯片，构建昇腾全产业链底座。昇腾系列处理器是基于华为达芬奇架构的神经网络处理器 (NPU)，包括昇腾 310 和昇腾 910 两款针对不同应用领域的芯片：

- **昇腾 310 定位为推理处理器**，高效、灵活、可编程，能够以最大功率 8W 达到 16 位半精度 (FP16) 8TFLOPS、整数精度 (INT8) 16TOPS 的算力，能效比极高，主要用于智能手机、智能附件等电子终端。
- **昇腾 910 定位为训练处理器**，高度集成的片上系统在提供高算力的同时可减少与 Host CPU 的交互，最大功率 310W，运算能力可达 FP16 320 TFLOPS、INT8 640TOPS，比英伟达 A100 更高，主要应用于高端服务器和云计算。

表1.昇腾处理器与英伟达 A800 处理器

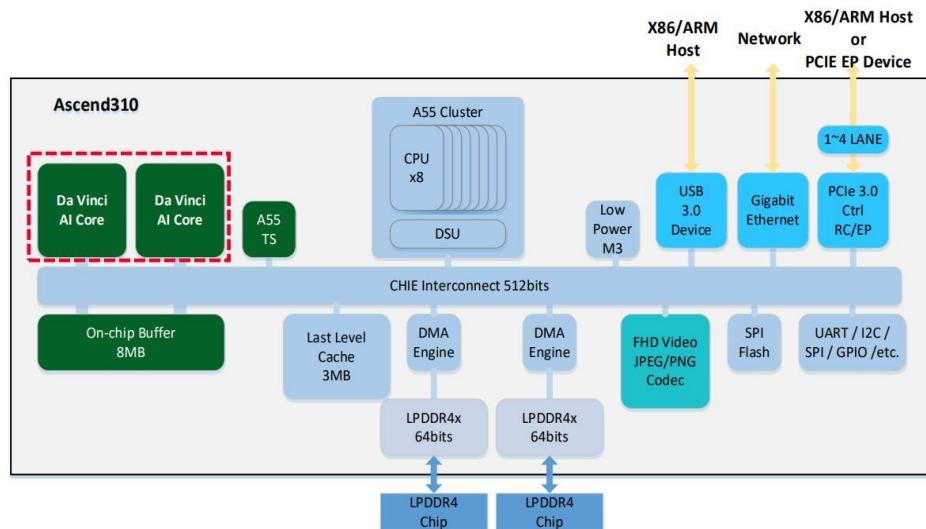


	昇腾 310	昇腾 910	英伟达 A100
架构	华为达芬奇	华为达芬奇	英伟达 Ampere
FP16 算力	8 TFOPS	320 TFLPOS	312 TFLOPS
INT8 算力	16 TOPS	640 TOPS	624 TOPS
最大功率	8W	310W	300 W

数据来源：海思官网、英伟达官网、财通证券研究所

量产 ASIC 高效低成本，独有片上系统结构保障灵活性。昇腾系列 NPU 属于 ASIC（即专用集成电路）的一种。相较于 GPU 和 FPGA 等其他类型 AI 芯片，ASIC 具有高度定制化特征，一次流片即定型，可编程性不足，但特定性能、规格、功耗等均显著优于前两者；初期成本高、开发周期长拉高生产进入门槛，但大产量规模可将单片成本压到极低。昇腾系列芯片采用“量产一代、研发一代、规划一代”的投产思路，保障每一代 ASIC 可用性的同时能够通过大规模生产。并且，昇腾 AI 芯片设计为 SoC（片上系统）类型，即单个芯片上集成了包括数据存储、计算单元等的整个系统，功能全面且支持芯片灵活组合。

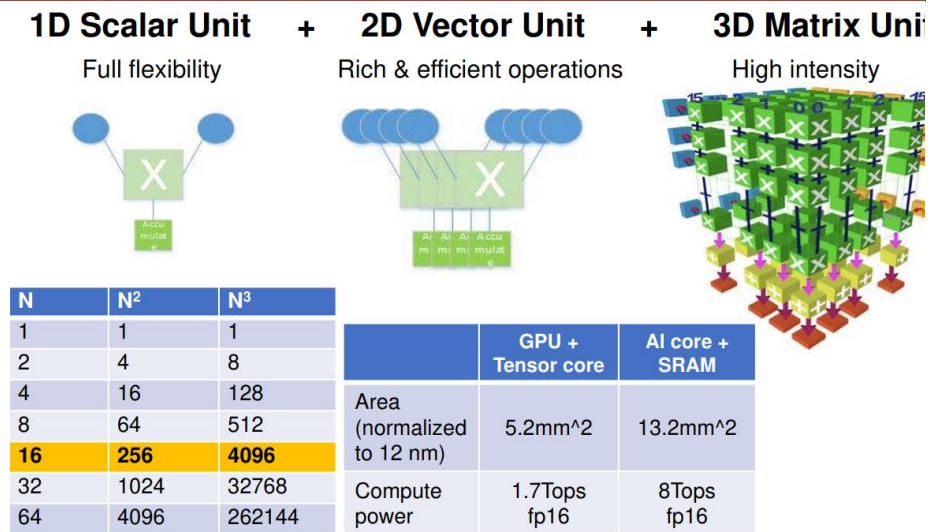
图6.昇腾 AI 芯片结构（昇腾 310）



数据来源：华为《DaVinci: A Scalable Architecture for Neural Network Computing》hot chips 2019、财通证券研究所

华为自研达芬奇架构形成独特竞争优势，有效提高神经网络运算效率。达芬奇架构是华为自研的面向 AI 神经网络卷积计算设计的架构，针对矩阵运算进行加速，算力部分由 3D Cube 矩阵乘法单元、Vector 向量计算单元、Scalar 标量计算单元三部分提供，共同组合提供并行化多样化算力。独创 16*16*16 的高密度 3D Cube 使得单个 AI Core 每时钟周期可进行 4096 个 16 位半精度浮点 MAC 计算，而 4*4*4 结构的英伟达 Tensor Core 单个 AI 核心仅支持 64 个。并且，基于其灵活可扩展的特性，达芬奇架构可以通用硬件架构实现从低端到高端的全覆盖，能够满足端边云多侧全场景算力要求。

图7.达芬奇 3D Cube 计算单元



数据来源：华为《DaVinci: A Scalable Architecture for Neural Network Computing》hot chips 2019、财通证券研究所

坚持“硬件开放”策略，提供多样化算力选择。昇腾系列硬件基于同一达芬奇架构开发多用途硬件，具有以下特征：

- **产品针对性强：**昇腾系列硬件包括模组、板卡、小站、服务器、集群等丰的产品种类及形态，囊括“云、边、端”全场景 AI 基础设施方案。以加速卡为例，昇腾提供训练卡、推理卡、视频解析卡三类，分别针对模型训练、边端推理、高强度视频解析等不同需求。
- **扩展性强：**硬件基于统一的达芬奇架构，可灵活高效扩展，Atlas 900 PoDA2 集群基础单元最大可拓展至 3.2 EFLOPS，且保证全节点 200G 网络互联。
- **硬件搭配选择多样化：**除提供既有的 Atlas 系列硬件产品外，华为还将昇腾系列芯片等硬件开放给合作伙伴，让伙伴基于 Atlas 系列硬件进行集成和二次开发，联合打造有竞争力的产品。

表2.昇腾加速卡产品图



产品名称	Atlas 300T Pro 训练卡	Atlas 300I Duo 推理卡	Atlas 300V Pro 视频解析卡
FP16 算力	280 TFLOPS	140 TFLOPS	70 TFLOPS
INT8 算力	560 TFLOPS	280 TOPS	140 TOPS
最大功率	300W	150W	72 W
基准频率	2.2 GHz	1.9 GHz	1.9 GHz
工作温度	5°C~45°C	0°C~55°C	0°C~55°C
尺寸	266.7mm x 39.04mm x 111.15mm	266.7mm x 18.46mm x 111.15mm	169.5mm x 18.45mm x 68.9mm

数据来源：昇腾官网、财通证券研究所

表3.昇腾 Atlas 系列集群基础单元



产品名称	Atlas 900 PoD 集群基础单元	Atlas 900 PoD A2 集群基础单元
NPU	可使用昇腾 910	64 * 昇腾 910
CPU	32 * 鲲鹏 920	32 * 鲲鹏 920
带宽	1228 GB/s	1600 GB/s
FP16 算力	14.14 ~ 20.40 PFLOPS	25.6 / 24 PFLOPS
算力扩展	最高可扩展到 1.3 EFLOPS 全节点 100G 网络互联	最高可扩展到 3.2 EFLOPS 全节点 200G 网络互联
能效比	可达 0.44 PFLOPS / kW	可达 0.5 PFLOPS / kW

数据来源：昇腾官网、财通证券研究所

液冷技术加码高能效，算力集群大幅扩展。在数据中心侧，更高的能效比可大幅降低数据中心运营成本。Atlas 集群作为业界首个全液冷 AI 集群，采用板级液冷、柜级风液换热器等独特设计，单柜即可支持 46kW，16 个机柜即可提供 256 PFLOPS 算力，整个 AI 集群的 PUE 小于 1.1。在边缘侧，模组能效比可达 2 TOPS/W，适应低功耗和边缘部署需求。高能效使得算力集群密度增加成为可能，2023 年 7 月的世界人工智能大会上，华为宣布昇腾 AI 集群全面升级，推出首个万卡 AI 集群，规模从最初的 4000 卡集群扩展至 16000 卡，拥有更快训练速度和稳定训练周期。

图8. Atlas 系列集群基础单元特征

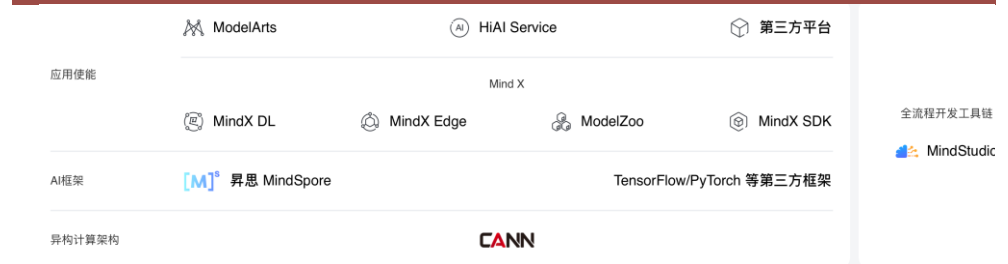
超强AI算力	47U 高度可最大提供20.4 PFLOPS FP16 AI算力 搭载超32颗超强算力的鲲鹏920处理器，高效加速应用
更优AI能效	能效比最高可达20.4 PFLOPS / 46 kW 全液冷散热，免冷机、免空调，PUE≤1.15
极佳AI拓展	支持机柜单元扩展，最大可拓展至4096颗昇腾910处理器集群， 总算力达1.3 EFLOPS FP16，全节点100G网络互联

数据来源：昇腾官网、财通证券研究所

2.2 软件开源，简化 AI 场景开发及行业应用流程

坚持全栈开放，开源开放原则，构建良好 AI 产业生态。在软件层面，昇腾 AI 产业链以 CANN 异构计算架构为底座，支持 MindSpore 自研 AI 框架以及 TensorFlow 等第三方框架，同时打造在应用使能平台，MindX 由“2+1+N”（2 个组件、1 个模型库、N 个 SDK）组成，助力行业开发者将 AI 计算融入各行各业，提供不同开发工具和套件，同时满足极简开发和极致性能两方面的需求。

图9. AI 集群产品介绍



数据来源：昇腾官网、财通证券研究所

CANN（异构计算架构）：针对 AI 场景开发，软硬件层面保障全方位兼容。该架构专为深度学习设计，向下使能昇腾处理器的并行加速能力，向上提供统一、多层次编程接口，助力用户快速构建基于昇腾平台的 AI 应用和业务。CANN 分为四层结构，下层实现硬件与操作系统的适配和支持，中层协调内存管理、算力分配和资源调动任务，上层提供多样接口实现高效率开发。该系统具有高兼容性，支持端边云全场景协同，支持 10+种设备形态、14+种操作系统及多种计算框架（包括 PyTorch、TensorFlow 等主流框架），支持向后兼容和演进。

图10.CANN 异构计算架构



数据来源：《昇腾计算产业发展白皮书》华为\中国信通院等、财通证券研究所

MindSpore（昇思）：支持多处理器架构的全场景开放 AI 计算框架。MindSpore 在动静态图转换、自动并行以及端边云协同等方面做出较大创新；最佳匹配昇腾 AI 处理器算力，支持端边云全场景灵活部署，降低 AI 开发门槛，致力于实现开发态友好、运行态高效、全场景按需协同三大目标。MindSpore 同样致力于构建良好开源生态，2020 年 3 月 28 日，华为于开发者大会 2020 上宣布该架构在码云正式开源。MindSpore 与社区合作，目前已覆盖逾 1.6 万用户，与 33 所高校开展 AI 课程、研究合作，提供 20 多种主流模型，覆盖 150 多种应用。

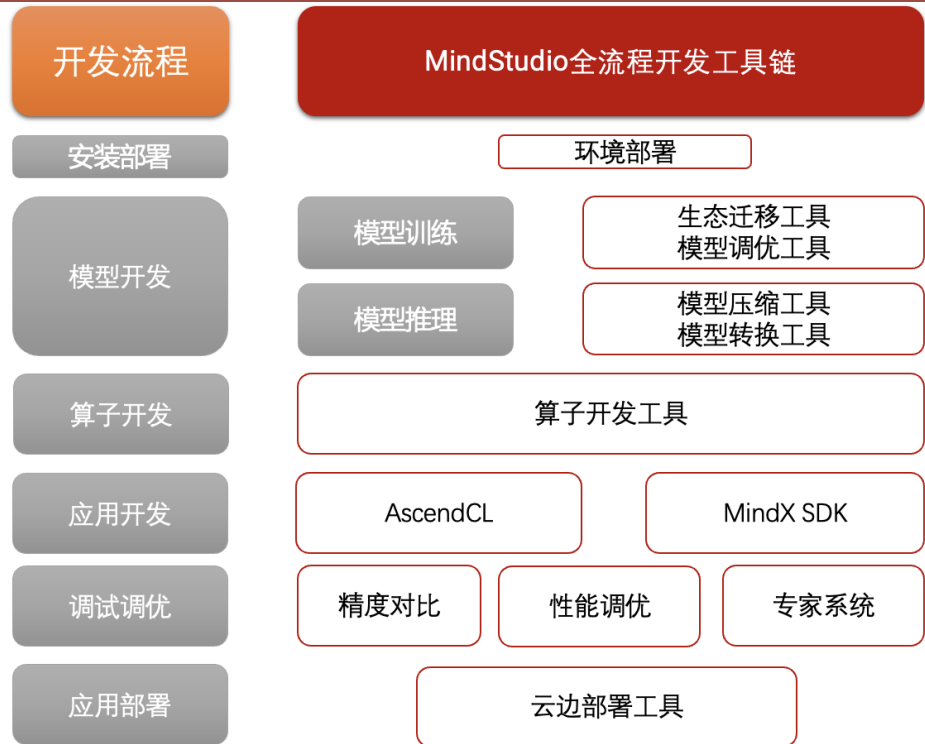
图11.昇思 MindSpore 四大优势



数据来源：昇思官网、财通证券研究所

MindStudio：全流程开发工具链，提供全新开发体验。Mindstudio 整合了算子开发、模型训练、模型推理、应用开发的所有工具流，提供四大开发体验：1) 插件化设计：提供开发接口，用户可自定义工具插件；2) 便捷安装：提供安装工具，针对难理解词汇提供实时注释；3) 开发辅助：提供语法一键补全功能；4) 精准分析服务：可视化调优；提供基于不同维度的蘑菇型精度比对结果。

图 12.MindStudio 开发流程



数据来源：昇腾官网、财通证券研究所

MindX：应用使能平台，“2+1+X”助力 AI 融入千行百业。 MindX DL 给业界伙伴提供深度学习系统参考架构，MindX Edge 实现将云端模型推送至边缘断部署，同时将边缘侧未识别数据上传至云端进行增量训练。ModelZoo 将 AI 开发需要的模型提前挑优并保障精度性能后提供给开发者。SDK 作为软件开发套件，凝聚了行业知识，结合 AI 最佳实践，可以提供给开发者行业所需的算法以及各种推理验证工作，帮助开发者提高开发效率。

图 13.MindX 应用使能



数据来源：《昇腾计算产业发展白皮书》华为\中国信通院等、财通证券研究所

多样化 AI 解决方案全方位使能行业应用。昇腾 AI 市场提供多行业的成熟的 AI 解决方案，涵盖数字政府、能源、金融、交通、电信、制造、医疗、教育等，同时提供应用案例，按多维分类展示，便于用户查询。昇腾 AI 为企业提供特定软硬件解决方案，硬件解决方案通常为向企业提供 AI 服务器算力（如下表案例 3 和 4），软件解决方案以图像和视频的识别为主（如下表案例 1 和 2），助力企业数值化管理运营、降本增效。

表4.昇腾 AI 解决方案及案例

分类依据	内容	案例
应用领域	计算机视觉（包括 OCR、图像分类、目标识别、图像识别等）和语音	（1）为 富士康 光伏控制器生产过程中的硅脂涂抹、标签检测等环节提供 智能质检 ，实现过程缺陷数量级降低
类型	训练或推理	（2）为 华鲲振宇 提供 推理+训练+计算一体服务 ，可实现图像、视频等多种类型的数据分析和推理计算，对充电站区域进行远程无死角监测
行业	能源、金融、交通、电信、制造、医疗、教育等	（3）基于 Atlas 500 Pro 智能边缘服务器，并结合 拓维信息 提供的 AI 稽核算法，共同实现实时 智能高速收费稽核
适配昇腾产品	加速模块、加速卡、智能小站、服务器、集群、昇思 MindSpore 等	（4）为 中国移动 提供 850 台 Atlas 800 服务器 ，用于开展对内、对外人工智能业务，如目标识别、OCR、工单稽核、经营分析、智慧营业厅等

数据来源：昇腾官网、财通证券研究所

3 外部：联合合作伙伴，打造产业生态

3.1 与鲲鹏处理器、盘古大模型结合，统合华为人工智能生态

鲲鹏展翅，共赢计算新时代。以鲲鹏基础芯片为底座，给予盘古大模型强大支撑。鲲鹏通用计算平台提供基于鲲鹏处理器的TaiShan服务器、鲲鹏主板及开发套件，可帮助开发者加速应用迁移和算力升级。软件厂商基于 openEuler 开源 OS 以及配套的数据库、中间件等平台软件发展应用和服务，鲲鹏与合作伙伴已进入各省市政务云、金融核心交易系统、运营商三朵云和电力调度等行业核心场景。

图14.鲲鹏计算产品



数据来源：鲲鹏官网、财通证券研究所

鲲鹏+昇腾生态稳步发展，硬件开放利好合作伙伴。华为完全开放鲲鹏主板接口规范和设备管理规范，合作伙伴可快速开发自有品牌的服务器和台式机产品。鲲鹏/昇腾生态领域使能诸多合作伙伴，四川长虹子公司华鲲振宇基于鲲鹏+昇腾推出“天宫”自主品牌系列产品，产品已规模应用于金融、运营商、交通、教育医疗等多个行业。截至2023年5月，鲲鹏和昇腾AI开发者已经超过350万，合作伙伴超过5600家，解决方案认证超过15500个。

图15.华为开放昇腾 Atlas、鲲鹏主板等产品



数据来源：华为官网、财通证券研究所

盘古大模型 3.0 问世，百“模”大战升级。截止 2023 年 7 月，全球已发布数百个大模型，我国年内已发布 80 多个大模型，面向消费者端的应用百花齐放。华为云盘古大模型 3.0 作为中国首个全栈自主的 AI 大模型，面向整个行业，包括“5+N+X”三层架构：Level 0 层包括自然语言、视觉、多模态、预测、科学计算五个基础大模型，Level 1 层是 N 个行业大模型，Level 2 层则专注于具体行业应用或特定业务场景的模型服务。

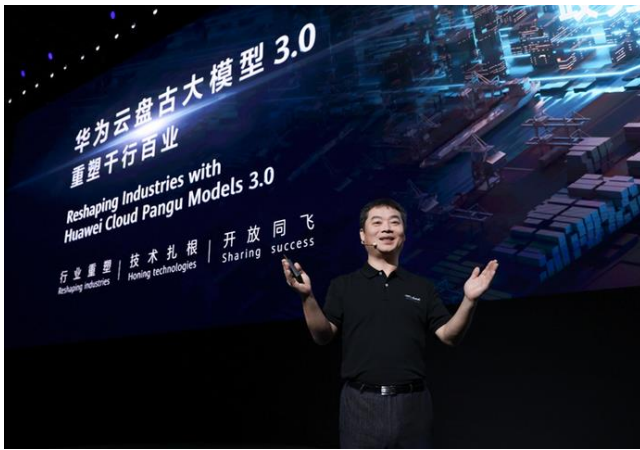
图 16. 盘古大模型三层架构



数据来源：华为云官网、财通证券研究所

盘古大模型全栈技术创新，服务千行百业。基于鲲鹏和昇腾为基础的 AI 算力云平台，以及异构计算架构 CANN、全场景 AI 框架昇思 MindSpore 等，为客户提供 100 亿参数、380 亿参数、710 亿参数和 1000 亿参数的系列化基础大模型，匹配客户不同场景、不同时延、不同响应速度的行业多样化需求。基于华为的 AI 根技术，大模型训练效能可以调优到业界主流 GPU 的 1.1 倍。目前，华为云盘古大模型已在金融、制造、医药研发、煤矿、铁路等诸多行业发挥着巨大价值。

图17.华为云盘古大模型 3.0 发布



数据来源：华为云官网、财通证券研究所

图18.盘古大模型 3.0 全栈创新



数据来源：华为云官网、财通证券研究所

开放同飞，携手联合共建大模型生态。基于昇腾 AI 澎湃算力，原生研发、适配的大模型超过 30 个，占据中国大模型近一半数量。为了进一步推动联合创新，华为携手 26 家 AI、金融、运营商、互联网行业领军企业、高校及科研院所，共同启动昇腾 AI 大模型联合创新，促进产业数智化转型升级。其中，武汉大学基于昇腾 AI 推出了大模型武汉.LuoJia；中科院自动化所发布基于昇腾 AI 的“紫东太初”2.0 全模态大模型等。

图19.华为云盘古大模型生态计划



数据来源：华为云官网、财通证券研究所

图20.业内首个千亿级三模态大模型—紫东太初



数据来源：中国科学院自动化研究所、财通证券研究所

人工智能产业迎来政策支持，鼓励大模型发展。2023年4月，国家网信办起草了《生成式人工智能服务管理办法（征求意见稿）》，公开征求意见。2023年7月7日，在世界人工智能大会上，国家人工智能标准化总体组宣布了我国首个大模型标准化专题组组长，由上海人工智能实验室与360集团、百度、华为、阿里等企业联合担任，正式启动大模型测试国家标准制订。2023年7月14日，国家互联网信息办公室近日联合国家发展改革委、教育部、科技部、工业和信息化部、公安部、广电总局公布《生成式人工智能服务管理暂行办法》，有助于推动人工智能产业结构升级优化。

表5.《生成式人工智能服务管理暂行办法》条例

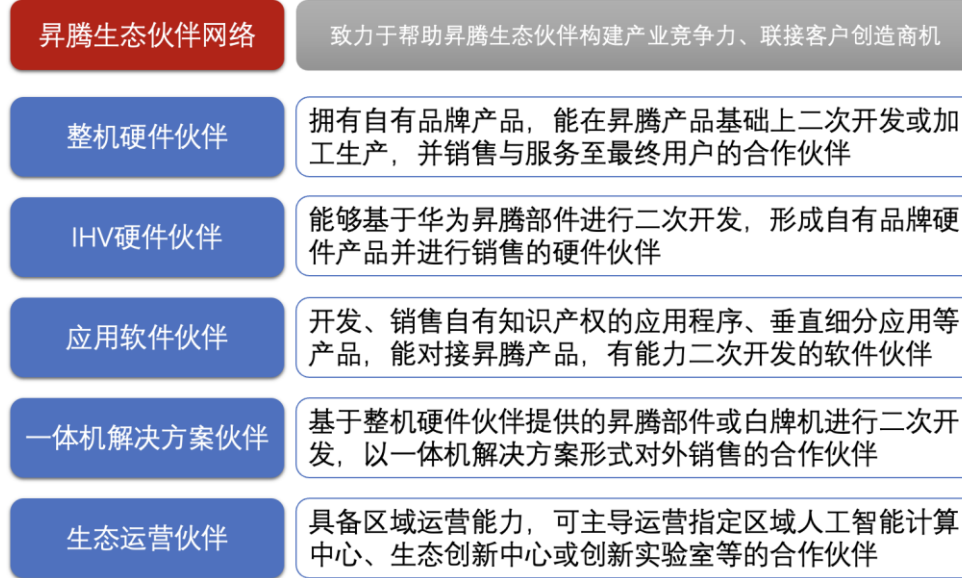
文件名称	《生成式人工智能服务管理暂行办法》
实施日期	2023年8月15日
相关内容	<p>秉持鼓励创新与规范发展的原则理念，要求采取有效措施鼓励生成式人工智能创新发展，对生成式人工智能服务实行包容审慎和分类分级监管。</p> <p>明确鼓励生成式人工智能技术在各行业、各领域的创新应用，生成积极健康、向上向善的优质内容，探索优化应用场景，构建应用生态体系。</p> <p>支持行业组织、企业、教育和科研机构、公共文化机构、有关专业机构等在生成式人工智能技术创新、数据资源建设、转化应用、风险防范等方面开展协作。</p> <p>鼓励生成式人工智能算法、框架、芯片及配套软件平台等基础技术的自主创新，平等互利开展国际交流与合作，参与生成式人工智能相关国际规则制定。</p>

数据来源：中国政府网、财通证券研究所

3.2 使能外部合作伙伴，带动各行业“AI+”降本增效

昇腾生态伙伴网络：联合AI走深向实。昇腾生态伙伴网络目前已有宝德、华鲲振宇、长江计算、清华同方等合作伙伴。生态伙伴分为五大类：1) 整机硬件伙伴，2) IHV 硬件伙伴，3) 应用软件伙伴，4) 一体机解决方案伙伴，5) 生态运营伙伴。华为通过总经销商供货的销售支持、与总经销商双方的技术支持模式，以及多种商务权益激励，鼓励伙伴基于昇腾AI打造自有品牌的产品或解决方案，实现共赢。并且，合作伙伴与昇腾系列合作产品的参数与有效期在华为官网完全公开，有效帮助伙伴与客户选择产品。

图21.昇腾生态伙伴网络



数据来源：昇腾官网、财通证券研究所

图22.昇腾兼容性认证硬件目录

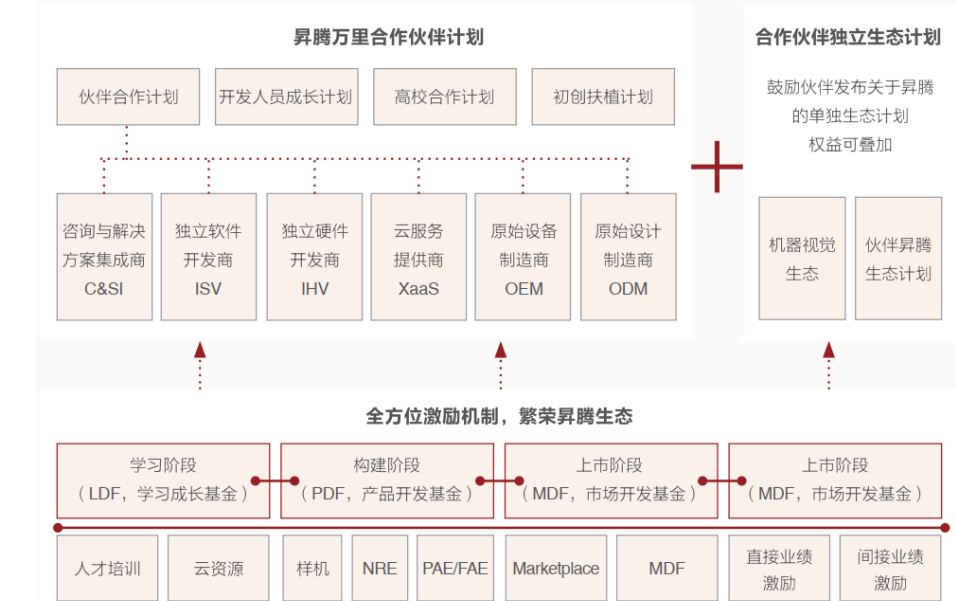
兼容性认证目录 Q 请输入合作伙伴名称 / 服务器型号 / 昇腾AI处理器 / CPU系列

合作伙伴	服务器型号	昇腾AI处理器	每节点最大AI处理器数	CPU系列	每节点最大CPU数	服务器形态	证书类型	有效期
宝德	PR215FI	Atlas 300I/V Pro 推理卡	4卡	飞腾处理器S2500	2	2U	Compatible	2022/08-2025.
宝德	PR215PI	Atlas 300V 视频解析卡	7卡	英特尔V5 (Skylake, Cascade Lake)	2	2U	Compatible	2022/11-2025.
北联国芯	BLX 800SI	Atlas 300V 视频解析卡	8卡	英特尔V6 (Ice Lake)	2	2U	Compatible	2022/11-2025.
北联国芯	BLK 800GI V2	Atlas 300T Pro 训练卡 (型号: 9000)	2卡	鲲鹏920	2	2U	Compatible	2022/08-2025.
北联国芯	AI中心推理型服务器 BLK 800GI (model 7000I)	Atlas 300I Pro 推理卡	8卡	鲲鹏920	2	2U	Compatible	2022/03-2025.
长江计算	G220K V2	Atlas 300T Pro 训练卡 (型号: 9000)	2卡	鲲鹏920	2	2U	Compatible	2022/09-2025.
长江计算	G220X V6	Atlas 300V 视频解析卡	8卡	英特尔V6 (Ice Lake)	2	2U	Compatible	2022/11-2025.
广电五舟	S800F2	Atlas 300I/V Pro 推理卡	4卡	飞腾处理器S2500	2	2U	Compatible	2022/08-2025.
华鲲振宇	天宫之印 AT800 (型号: 8000)	Atlas 300T Pro 训练卡 (型号: 9000)	2卡	鲲鹏920	2	2U	Compatible	2022/08-2025.

数据来源：昇腾官网、财通证券研究所

昇腾万里伙伴合作计划：共同繁荣昇腾 AI 产业生态。基于昇腾 AI 基础软硬件平台推出的一项合作伙伴计划，向合作伙伴提供培训、技术、营销和市场的全面支持。截止 2023 年 7 月，昇腾已认证了 30 多家硬件伙伴、1200 多家软件伙伴，联合打造了 2500 多个场景方案。昇腾计算产业生态秉承“硬件开放、软件开源，使能合作伙伴”的战略，与合作伙伴共建、共享全新的人工智能生态，共同促进 AI 赋能千行百业。

图23.昇腾万里合作伙伴计划



数据来源：《昇腾计算产业发展白皮书》华为\中国信通院等、财通证券研究所

3.3 深入高校与开发者社区，拓展昇腾 AI 影响力

“一中心四平台”产业发展，打通“政产学研用”。华为 AI 计算中心提供公共算力服务平台、应用创新孵化平台、产业聚合发展平台和科研创新人才培养平台，以此实现“政产学研用”五位一体打通，形成区域乃至全国的 AI 产业的汇聚。仅 2023 年 1-7 月，全国就已新成立 3 个昇腾相关人工智能中心。目前整个昇腾 AI 集群已支撑全国 25 个城市的人工智能计算中心建设，其中 7 个城市公共算力平台入选首批国家“新一代人工智能公共算力开放创新平台”（共 9 个）。

表6.2023 年与昇腾相关的新成立人工智能中心

名称	成立时间及详情
北京昇腾人工智能计算中心	2023 年 2 月 13 日在北京人工智能产业创新发展大会上正式成立
青岛人工智能计算中心	2023 年 2 月正式上线，基于华为昇腾 AI 打造
长沙昇腾人工智能创新中心	2023 年于湘江人工智能高峰论坛上举行长沙昇腾人工智能创新中心上线运营启动仪式

数据来源：华为计算、财通证券研究所

昇腾科研创新使能计划为国内高校与科研院所提供有力研发支持。通过向昇腾平台提交申请，国内科研团队可借助昇腾人工智能基础软硬件平台，开展科学技术和自主可信软件研发与技术攻关工作。通过审核后，该计划会向合作团队提供技术创新学习资源、软硬件及资金支持，并邀请团队成员参与相关影响力活动。

表7.昇腾科研创新使能计划享有权益

学习	• 技术赋能：参加昇腾关键技术开发者活动
	• 昇腾生态创新中心学习资源：课程学习资源补贴
	• 昇腾社区专属权益：技术资料下载权限、新品试用机会等
支持	• 科研创新基金支持：可申请年额度 最高 100 万元的 NRE 激励
	• 算力资源支持：配额 最高 5 万元算力/团队/年
	• 昇腾生态样机：最多可申请 Atlas 200I DK A2 开发者套件 10 套/团队/年
影响力	• 专题演讲：优先受邀作为演讲嘉宾参加由华为主办的产业活动
	• 荣誉奖励：优先受邀参加“昇腾科研创新卓越贡献奖”年度评选
	• 荣誉称号：优先受邀成为昇腾产业领域专家并获得证书和权益

数据来源：昇腾官网、财通证券研究所

高度重视开发者体验，昇腾开发者社区强力赋能服务开发者。开发者社区提供大量学习与开发资源，对业务开发者、AI 应用开发者与算法工程师、算子及网络开发者分层开放软硬件能力，针对性提供所需功能。昇腾开发者社区作为资源中心、赋能中心和支持中心，提供 100+种工具与样例和 50+种数据集、50+门直播课程和 70+门个人课程、20+名专家轮值服务，并有开发者成长计划、初创企业扶植计划、AI 人才培养计划等面向个人、企业、高校的扶持计划，旨在加速昇腾开发者成长、构建昇腾开发者生态。

图24.昇腾开发者社区



数据来源：昇腾官网、财通证券研究所

4 投资建议

建议关注华为产业链：

- 整机合作伙伴：神州数码、拓维信息、四川长虹、紫光股份；
- 垂直应用合作伙伴：软通动力、赛意信息、东方国信、常山北明、格灵深瞳、云从科技、云天励飞、海量数据、润和软件、智洋创新。

5 风险提示

AI 技术迭代不及预期的风险：若 AI 技术迭代不及预期，NLP 模型优化受限，则相关产业发展进度会受到影响。

商业化落地不及预期的风险：华为昇腾、鲲鹏、盘古以及 AI 工具链正在发展期，与全球领先的产品相比，商业价值仍有待进一步验证，若市场拓展不及预期，给予相关公司的增量业绩有限。

政策监管风险：目前 AIGC 相关技术还处于发展早期，后续若出台监管政策可能会对行业发展有一定影响。

信息披露

● 分析师承诺

作者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，具备专业胜任能力，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。本报告清晰地反映了作者的研究观点，力求独立、客观和公正，结论不受任何第三方的授意或影响，作者也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。

● 资质声明

财通证券股份有限公司具备中国证券监督管理委员会许可的证券投资咨询业务资格。

● 公司评级

买入：相对同期相关证券市场代表性指数涨幅大于 10%；

增持：相对同期相关证券市场代表性指数涨幅在 5%~10%之间；

中性：相对同期相关证券市场代表性指数涨幅在-5%~5%之间；

减持：相对同期相关证券市场代表性指数涨幅小于-5%；

无评级：由于我们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，致使我们无法给出明确的投资评级。

● 行业评级

看好：相对表现优于同期相关证券市场代表性指数；

中性：相对表现与同期相关证券市场代表性指数持平；

看淡：相对表现弱于同期相关证券市场代表性指数。

● 免责声明

本报告仅供财通证券股份有限公司的客户使用。本公司不会因接收人收到本报告而视其为本公司的当然客户。

本报告的信息来源于已公开的资料，本公司不保证该等信息的准确性、完整性。本报告所载的资料、工具、意见及推测只提供给客户作参考之用，并非作为或被视为出售或购买证券或其他投资标的邀请或向他人作出邀请。

本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的价格、价值及投资收入可能会波动。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。

本公司通过信息隔离墙对可能存在利益冲突的业务部门或关联机构之间的信息流动进行控制。因此，客户应注意，在法律许可的情况下，本公司及其所属关联机构可能会持有报告中提到的公司所发行的证券或期权并进行证券或期权交易，也可能为这些公司提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。在法律许可的情况下，本公司的员工可能担任本报告所提到的公司的董事。

本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议。在任何情况下，本公司不对任何人使用本报告中的任何内容所引致的任何损失负任何责任。

本报告仅作为客户作出投资决策和公司投资顾问为客户提供投资建议的参考。客户应当独立作出投资决策，而基于本报告作出任何投资决定或就本报告要求任何解释前应咨询所在证券机构投资顾问和服务人员的意见；

本报告的版权归本公司所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制、发表或引用，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。