

行业研究

iPhone 时刻的开启，硅基时代的到来

——AI 行业深度报告之一——

要点

ChatGPT 引领人工智能进入以 AGI 为代表的新里程碑阶段，拉动算力需求，引领基础设施新周期。人工智能应用场景逐渐增多，模型参数规模和数据量也实现了大幅度增长，为 NLP、CV 等领域带来更强大的表达能力和性能。大模型是目前通往 AGI 的最佳实现方式。算力需求激增，推动算力基础设施迎增长新周期。据 IDC，AI 的应用与普及促使 2022 年我国智能算力规模近乎翻倍，达到 268EFLOPS，超过通用算力规模，2026 年我国智能算力规模将达 1271.4 EFLOPS，22-26 年复合增长率预计达 47.58%；硬件占 AI 支出比重 49.8%，其中最大投资是服务器，占 AI 硬件支出比重 84%以上，AI 将成为未来服务器市场新一轮快速增长的主要推动力。

全球科技巨头大力布局 AI 产业。（1）**微软：投资 OpenAI，Copilot 开启 AI 商业化里程碑。**2023 年 1 月中旬，微软对 OpenAI 投资 100 亿美元，占股达到 49%。微软推出全新的 Copilot 系统，发布 Dynamics 365 Copilot、Microsoft 365 Copilot、Power Platform Copilot 模块。（2）**英伟达：全球 GPU 龙头，产品需求量子价齐升。**全球 GPU 龙头开拓 AI 市场，助力服务器核心算力硬件性能提升。英伟达作为一家以 GPU 为主营业务的半导体设计公司，在 AI 行业发展初期就前瞻布局。FY4Q24 数据中心业务营收达到创纪录的 184 亿美元，环比增长 27%，同比增长 409%。算力需求迅猛增长，服务器 GPU 业务有望迎来量价齐升。（3）**华为：服务器产业深耕多年，打造“鲲鹏+昇腾”双引擎战略布局。**据 counterpoint 数据，2021 年全年华为服务器业务收入 43.43 亿美元，全球市场份额 4.5%。公司坚持计算产业战略，积极布局 AI 服务器领域。

AI 产业链：大模型加速落地，算力空间广阔，应用百花齐放。（1）大模型：国内和海外巨头加速布局，展望未来，随着底层技术逐步革新，基模型和领域大模型持续完善，大模型应用边界不断拓宽。（2）服务器：技术迭代催生新增长点，品牌与白牌之争趋于白热化。超算中心建设加速，服务器需求进一步提升；云边协同发展，增添服务器需求动能；AIGC 推动 GPU 服务器市场持续扩张；服务器品牌与白牌之争白热化，促进服务器市场优胜劣汰。AI 训练服务器方面，单个大语言模型训练驱动 AI 训练服务器市场约 2 亿美元需求。据 IDC 与浪潮信息数据，2023 年中国人工智能服务器市场规模将达到 91 亿美元，同比增长 82.5%，2027 年将达到 134 亿美元，五年复合增长率达 21.8%。（3）光模块：800G 需求释放，中国厂商厚积薄发。AI 大模型发展加速光模块迭代节奏，800G 光模块放量或将超预期：根据 Lightcounting 2021 年的预测，800G 光模块将从 2025 年底开始主导市场，主要原因是 AI 应用等带来的数据流量的增长，超预期的数据中心带宽需求以及光模块厂商技术的迭代。（4）边缘算力：AIoT 智能终端的“大脑”，端侧算法部署的核心。（5）PCB：AI 算力提升驱动行业快速成长。（6）智能终端：AI 赋能硬件入口，加速万物互联。

投资建议：推荐工业富联、大华股份、海光信息、寒武纪、芯原股份。（1）工业富联：消费电子龙头，AI 服务器打开成长空间；（2）大华股份：中国移动入股落地，智慧物联龙头扬帆起航；（3）海光信息：CPU/DCU 领军企业；（4）寒武纪：AI 芯片技术国内领先；（5）芯原股份：稀缺的国产 IP 供应商，Chiplet 布局领先。

风险提示：宏观经济风险、AIGC 发展不及预期风险、市场需求及上游供应不足风险。

电子行业
买入（维持）

作者

分析师：刘凯

执业证书编号：S0930517100002

021-52523849

kailiu@ebsecn.com

分析师：吴春阳

执业证书编号：S0930521080002

021-52523686

wuchunyang@ebsecn.com

分析师：林仕霄

执业证书编号：S0930522090003

021-52523818

linshixiao@ebsecn.com

联系人：孙啸

021-52523587

sunxiao@ebsecn.com

联系人：王之含

wangzhihan@ebsecn.com

行业与沪深 300 指数对比图



资料来源：Wind

投资聚焦

研究背景

近几年来人工智能高速发展，现已经进入了以 AGI 为代表的新里程碑阶段。随着 AI 基础设施逐步完善，深度学习技术不断进步，人工智能应用场景逐渐增多，模型参数规模和数据量也实现了大幅度增长。伴随 AI 技术升级和大模型成熟，AI 绘画与 ChatGPT 的成功破圈，生成式 AI 技术迎来发展拐点，行业关注度大幅提升。

我们区别于市场的观点

1、市场 AI 深度报告针对某一细分领域，我们本篇报告对服务器、GPU、PCB、云端算力、光模块、边缘算力、智能终端等领域进行深度分析。本文测算中国超算服务市场的规模，800G 光模块 2023-2025 年市场规模；此外本文还从智能手机、智能音箱、智能机顶盒、智能家居四个角度分析了 AI 对于智能终端行业的影响。

2、市场认为近两年算力建设完备，未来算力建设动力不足，我们认为受益于 AI 等相关新兴领域的应用以及“东数西算”政策下云计算、超算中心的蓬勃发展，数据计算、存储需求呈几何级数增长，算力需求持续释放带动算力基础设施产业迎来增长新周期。

股价上涨的催化因素

- (1) AI 场景落地超预期
- (2) AI 芯片研发超预期
- (3) 下游应用复苏超预期

投资观点

推荐工业富联、大华股份、海光信息、寒武纪、芯原股份。

工业富联：消费电子龙头，AI 服务器打开成长空间；

大华股份：中国移动入股落地，智慧物联龙头扬帆起航；

海光信息：CPU/DCU 领军企业；

寒武纪：AI 芯片技术国内领先；

芯原股份：稀缺的国产 IP 供应商，Chiplet 布局领先。

目 录

1、 AI: iPhone 时刻的开启，硅基时代的行业特征	7
2、 全球科技巨头大力布局 AI 产业	10
2.1、 微软:投资 OpenAI, Copilot 开启 AI 商业化里程碑	10
2.2、 英伟达:全球 GPU 龙头，产品需求量价齐升	12
2.3、 华为：服务器产业深耕多年，打造“鲲鹏+昇腾”双引擎战略布局	14
3、 AI 产业链：大模型加速落地，算力空间广阔，应用百花齐放	16
3.1、 大模型：国内和海外巨头加速布局	16
3.2、 服务器：AIGC 推动 GPU 服务器市场持续扩张	17
3.3、 光模块：800G 需求释放，中国厂商厚积薄发	19
3.4、 边缘算力：AIoT 智能终端的“大脑”，端侧算法部署的核心	21
3.5、 PCB：AI 算力提升驱动行业快速成长	22
3.6、 智能终端：AI 赋能硬件入口，加速万物互联	22
4、 投资建议	24
5、 重点公司推荐	25
5.1、 工业富联（601138.SH）：消费电子龙头，AI 服务器打开成长空间	25
5.2、 大华股份（002236.SZ）：中国移动入股落地，智慧物联龙头扬帆起航	30
5.3、 海光信息（688041.SH）：CPU/DCU 领军企业	36
5.4、 寒武纪（688256.SH）：AI 芯片技术国内领先	41
5.5、 芯原股份（688521.SH）：稀缺的国产 IP 供应商，Chiplet 布局领先	45
6、 风险提示	50

图目录

图 1：人工智能的发展	7
图 2：大模型发展及相关应用实现落地时间表	8
图 3：中国智能算力规模（EFLOPS）	9
图 4：Copilot 自动生成邮件	10
图 5：Dynamics 365 Customer Insights	11
图 6：Microsoft 365 Copilot 系统	11
图 7：Power Apps	12
图 8：Power Automate	12
图 9：2022 年 A100 及未来 H100 各领域应用数量统计	13
图 10：优秀的大模型能够赋能各行各业开放任务	16
图 11：大模型行业的国外厂商	16
图 12：大模型行业的国内厂商	17
图 13：英伟达 H100/A100 对比	17
图 14：全球服务器与中国出货量及增速	18
图 15：全球与中国服务器市场规模及增速（单位：亿美元，%）	18
图 16：2022 年中国服务器市场份额（按营收）	18
图 17：2022 年 AI 服务器采购量占比	18
图 18：数据中心光模块需求演进（21-26 年数据为预测值）	19
图 19：传统三层网络架构	20
图 20：叶脊架构	20
图 21：以太网光模块销售额增长率（23-28 年数据为预测值）	20
图 22：SoC 示意图	22
图 23：2021 年中国智能家居设备市场各品类出货量及同比增长率	23
图 24：中国 AI 产业链梳理	24
图 25：公司发展历程	25
图 26：公司股权结构（截至 2023 年 11 月）	25
图 27：2016 年-2023Q1-Q3 公司营业收入	26
图 28：2016 年-2023 Q1-Q3 公司归母净利润	26
图 29：2016 年-2023 Q1-Q3 毛利率及净利率	27
图 30：2016 年-2023 Q1-Q3 期间费用率	27
图 31：公司业务及产品	30
图 32：公司发展历程	31
图 33：公司股权结构（截至 2023 年三季报）	31
图 34：2016 年-2023 Q1-Q3 公司营业收入	32
图 35：2016 年-2023 Q1-Q3 公司归母净利润	32
图 36：2018Q1-2023Q3 公司单季收入	32
图 37：2018Q1-2023Q3 公司单季归母净利润	32
图 38：2016 年-2023 Q1-Q3 毛利率及净利率	33
图 39：2016 年-2023 Q1-Q3 期间费用率	33

图 40：公司打造持续升级巨灵人工智能开发平台	33
图 41：海光信息主要产品简介	36
图 42：海光信息营业收入及增速	36
图 43：海光信息归母净利润及增速	36
图 44：海光信息毛利率及净利率情况	37
图 45：公司研发费用及其营收占比	37
图 46：海光信息部分客户	38
图 47：寒武纪思元 370-S4 智能加速卡	41
图 48：寒武纪玄思 1000 智能加速器整机	41
图 49：寒武纪 2019 年-2023 Q1-Q3 营业收入及同比情况	42
图 50：寒武纪 2019 年-2023 Q1-Q3 归母净利润情况	42
图 51：寒武纪 2019 年-2023 Q1-Q3 毛利率及净利率情况	42
图 52：寒武纪 2019 年-2023 Q1-Q3 研发费用及占比情况	42
图 53：芯原股份业务	46
图 54：芯原股份 2019 年-2023 Q1-Q3 营业收入及同比情况	46
图 55：芯原股份 2019 年-2023 Q1-Q3 归母净利润情况	46
图 56：芯原股份 2019 年-2023 Q1-Q3 毛利率及净利率情况	47
图 57：芯原股份 2019 年-2023 Q1-Q3 研发费用及占比情况	47
图 58：基于 Chiplet 的异构架构应用处理器的示意图	48

表目录

表 1：英伟达主要芯片产品.....	13
表 2：英伟达 AI 服务器主要产品对比	14
表 3：华为服务器产品	15
表 4：工业富联收入拆分.....	28
表 5：工业富联可比公司估值-PE 估值.....	29
表 6：工业富联盈利预测与估值简表	29
表 7：大华股份收入拆分（单位：百万元，%）	34
表 8：可比公司估值-PE 估值	35
表 9：大华股份盈利预测与估值简表	35
表 10：海光信息主营业务拆分及预测（单位：亿元，%）	39
表 11：可比公司估值-PS 估值.....	40
表 12：海光信息盈利预测与估值简表.....	40
表 13：寒武纪产品布局	43
表 14：寒武纪主营业务拆分及预测（单位：亿元，%）	44
表 15：寒武纪可比公司估值	45
表 16：寒武纪盈利预测与估值简表	45
表 17：芯原股份主营业务拆分及预测（单位：亿元，%）	49
表 18：芯原股份可比公司估值	50
表 19：芯原股份盈利预测与估值简表.....	50

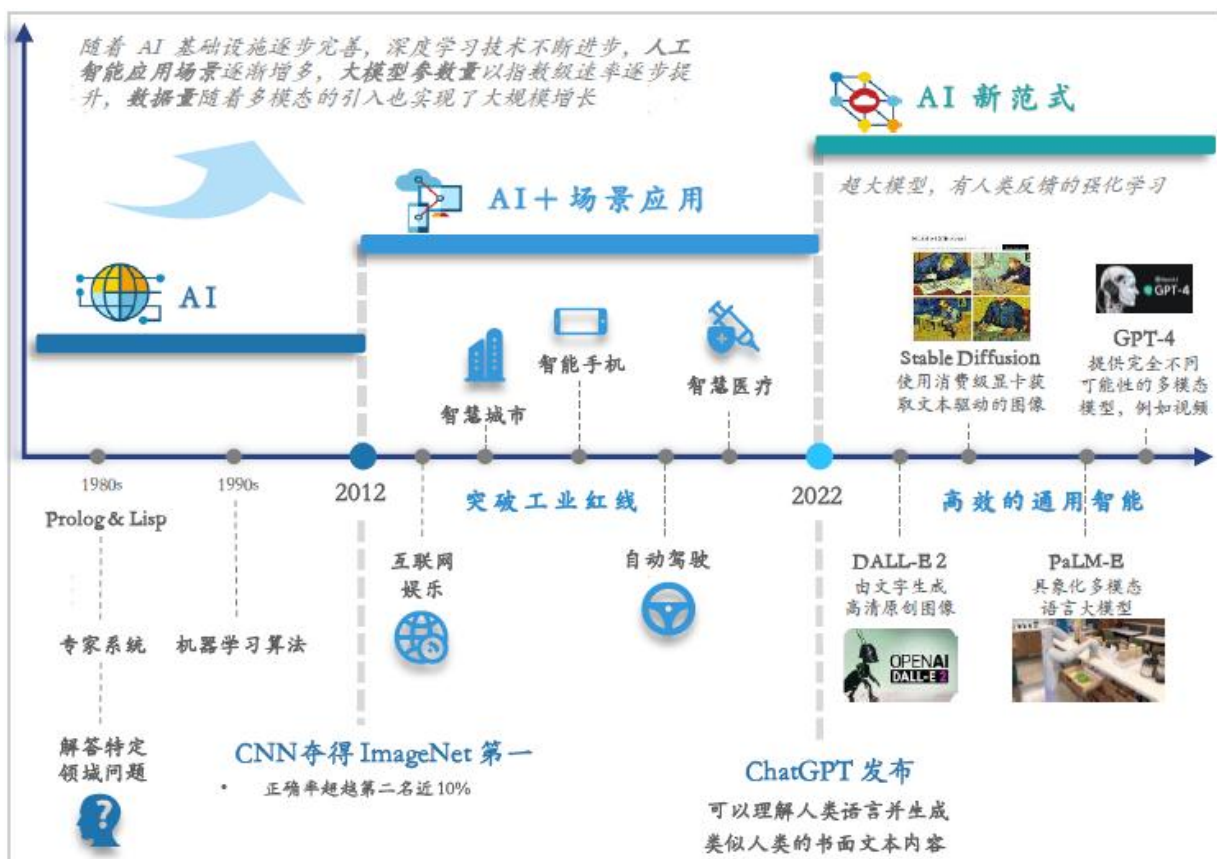
1、AI: iPhone 时刻的开启，硅基时代的行业特征

ChatGPT 引领人工智能进入以 AGI 为代表的新里程碑阶段。随着 AI 基础设施逐步完善，深度学习技术不断进步，人工智能应用场景逐渐增多，模型参数规模和数据量也实现了大幅度增长，为 NLP、CV 等领域带来更强大的表达能力和性能。人工智能发展历程中主要有两大里程碑：

里程碑一：2012 年 CNN 获得 ImageNet 第一，标志着机器视觉识别能力开始逐渐超越人眼识别准确率，开启了人工智能革命。随着深度学习技术不断突破，诞生了一批“AI+场景应用”的专属模型，但是整体研发成本比较高、研发时间比较长。

里程碑二：2022 年 ChatGPT 的出现，掀起了又一波人工智能发展热潮，以大模型+RLHF 为核心的技术落地意味着人工智能开启 AI 新范式。人工智能相关产业开始基于强大的基模型进行发展，通过人类反馈和强化学习不断解锁基模型的能力，以解决海量开放式任务，带来了新的研究范式。

图 1：人工智能的发展



资料来源：沙利文

AGI 技术能够精准识别人类情绪意图、理解人类语言、学习人类知识并进行类脑推理与创造，大模型是目前通往 AGI 的最佳实现方式。以 ChatGPT 为代表的人工智能技术已经具备 AGI 的核心技术和特征，能够自动化地学习任何可以符号

化的知识及信息，不断自我优化，充分理解和流畅表达人类语言，同时逻辑推理能力强，实现了具备一般人类智慧的机器智能。

图 2：大模型发展及相关应用实现落地时间表

	2020 年前	2020	2022	2025?	2030?	2050?
T 文字	垃圾邮件识别 翻译 基础问答	基础文案写作 起草初稿	更长的文字 完成第二稿	垂直微调 科学论文	高于人类平均水 平的终稿写作	产出比职业作家 写得更好的终稿
c++ 代码 生成	单行代码 自动完成	多行代码生成	更长的代码 更高的准确率	多程序语言 更多垂直领域	文本到产品 (草稿)	产出比全职开发人 员做得更好的文本 到产品 (终稿)
📷 图像			艺术作品 Logo 设计 摄影	产品设计模型 建筑概念模型	产品设计模型 建筑模型	产出比专业艺术家、 设计师、摄影师做 得更好的终稿
📺 视频/ 3D/ 游戏				基础/初稿视 频及 3D 文件	第二稿	AI Roblox 产出基于个性化梦 想的电子游戏和电影
可用的大模型： 初步探索 基本形成 准备阶段						

资料来源：沙利文

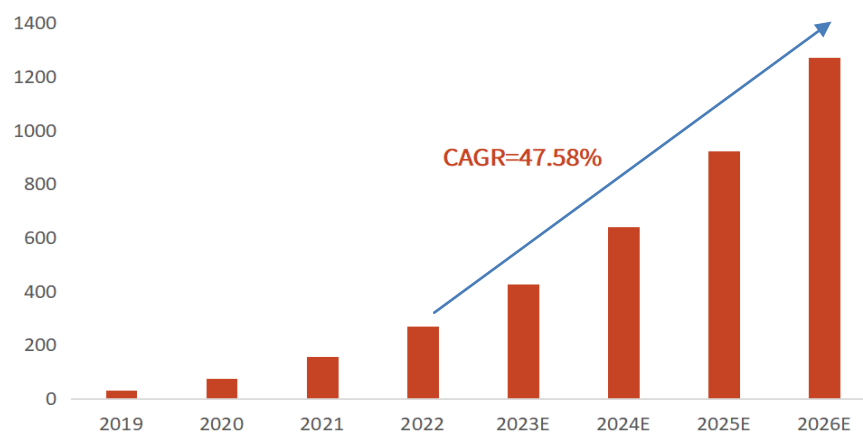
伴随 AI 技术升级和大模型成熟，AI 绘画与 ChatGPT 的成功破圈，生成式 AI 技术迎来发展拐点，行业关注度大幅提升。生成式 AI 是指基于大模型、生成对抗网络 GAN 等人工智能技术，通过已有数据寻找规律，并通过适当的泛化能力生成相关内容的技术，可生成如图像、文本、音频、视频等原创内容变体。例如，以 ChatGPT、Midjourney、文心一格、商汤商量、Codex 为代表的生成式 AI 应用拥有文本语言理解能力、涌现能力以及思维链推理能力，能够完成文学创作、新闻写作、数理逻辑推算、代码生成、图片生成等多项任务。目前，国内电商、游戏、文娱、设计等行业正在积极使用相关的生成式 AI 应用来提高自身工作效率，尤其以文生图应用为主。

生成式 AI 不仅能够增强并加速下游多领域的设计，而且有潜力“发明”人类可能错过的**新设计、新对象**。生成式 AI 有生成大规模、高质量、低成本内容优势，在算力和算法支持下生成大量内容，生成的内容质量将持续超越 UGC 与 PGC。未来有望为各行业提供内容支持并促进其内容繁荣，最大化释放内容生产力。

文字生成属于发展成熟、易于跨界转化的赛道，而跨模态生成赛道的发展潜力最大。生成式 AI 应用根据模态划分为文字生成、音频生成、图像生成、视频生成、跨模态生成。语音合成、文本生成图像属性编辑等技术应用目前较为成熟，跨模态生成、策略生成是高增长潜力的应用场景，在自动驾驶、机器人控制等领域有极高应用价值。

算力需求激增，推动算力基础设施迎增长新周期。受益于 AI 等相关新兴领域的应用以及“东数西算”政策下云计算、超算中心的蓬勃发展，数据计算、存储需求呈几何级数增长，算力需求持续释放带动算力基础设施产业迎来增长新周期。据 IDC，AI 的应用与普及促使 2022 年我国智能算力规模近乎翻倍，达到 268 EFLOPS，超过通用算力规模，2026 年我国智能算力规模将达 1271.4 EFLOPS，22-26 年复合增长率预计达 47.58%；硬件占 AI 支出比重 49.8%，其中最大投资是服务器，占 AI 硬件支出比重 84%以上，AI 将成为未来服务器市场新一轮快速增长的主要推动力。

图 3：中国智能算力规模（EFLOPS）



资料来源：IDC 统计及预测，光大证券研究所

2、全球科技巨头大力布局 AI 产业

2.1、微软:投资 OpenAI, Copilot 开启 AI 商业化里程碑

2023 年 1 月中旬,微软对 OpenAI 投资 100 亿美元,占股 49%。微软与 OpenAI 的合作始于 2019 年 7 月,当时微软宣布向 OpenAI 提供 10 亿美元的投资,并与其建立独家云计算合作关系。此外,微软还将为 OpenAI 提供技术支持和咨询服务,帮助其实现“通用人工智能”的愿景。

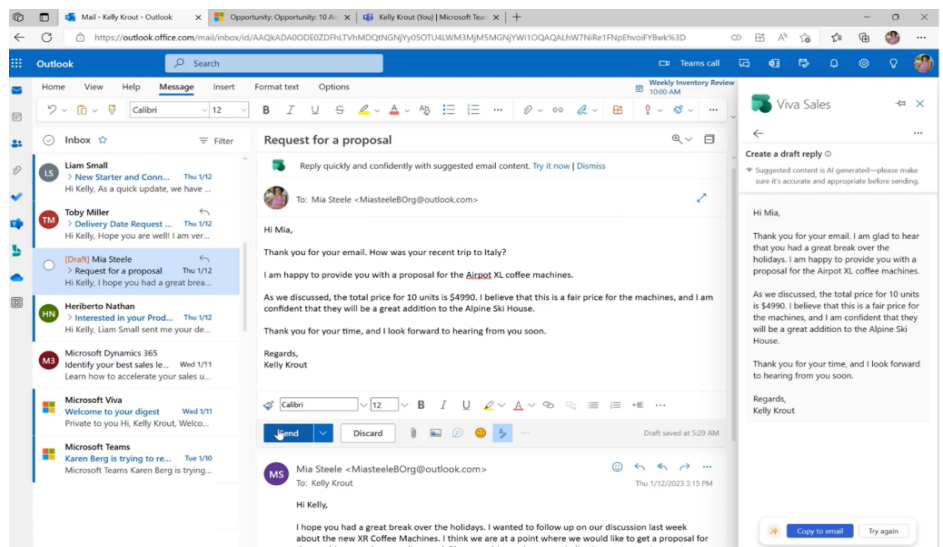
微软将 GPT-4 模型装进了 Office 套件,推出了全新的 Copilot 系统,发布 Dynamics 365 Copilot、Microsoft 365 Copilot、Power Platform Copilot 板块,开启 AI 商业化里程碑。

Dynamics 365 Copilot: GPT4 赋能业务全线

Dynamics 365 Copilot 赋能销售人员随时快速响应。在 Dynamics 365 Sales 和 Viva Sales 中, Copilot 可以编写给客户的电子邮件回复,在 Outlook 中自动生成一个 Teams 会议的总结邮件,从销售的 CRM 系统中自动提取产品、报价之类的细节,由 Teams 通话中总结出的核心要点一并汇总到邮件里。

Dynamics 365 Copilot 助力客服人员时刻对答如流。在 Dynamics 365 Customer Service 中,能够针对聊天对话和电子邮件中的问题,撰写出符合上下文语境的答案。

图 4: Copilot 自动生成邮件



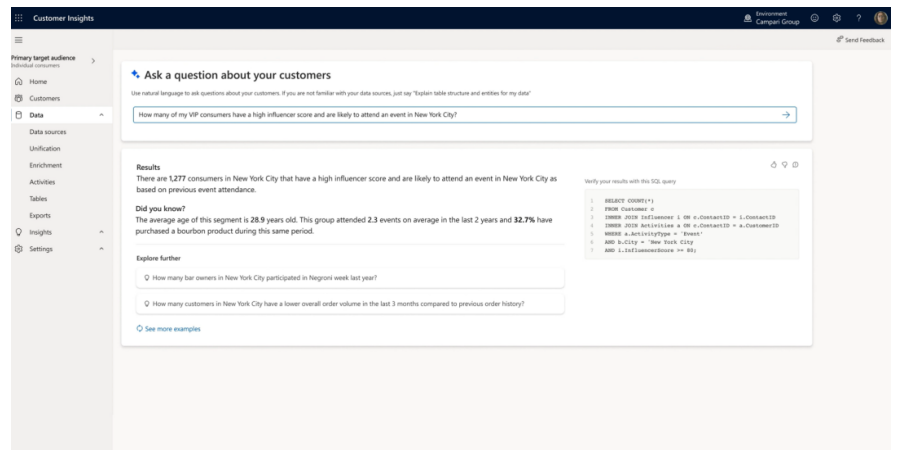
资料来源:微软官方公众号

Dynamics 365 Copilot 升级市场营销体验,轻松找到目标受众。在 Dynamics 365 Customer Insights 和 Dynamics 365 Marketing 中的 Copilot 使营销人员能够简化他们在数据探索、受众细分和内容创建方面的工作流程:

借助 Dynamics 365 Customer Insights 中的 Copilot,可以应对复杂的计算并且准确定位特定客户群体。

利用 Dynamics 365 Marketing 中的 Copilot,市场人员可以创建一个支持检索辅助功能的全新目标客户类别。

图 5: Dynamics 365 Customer Insights



资料来源：微软官方公众号

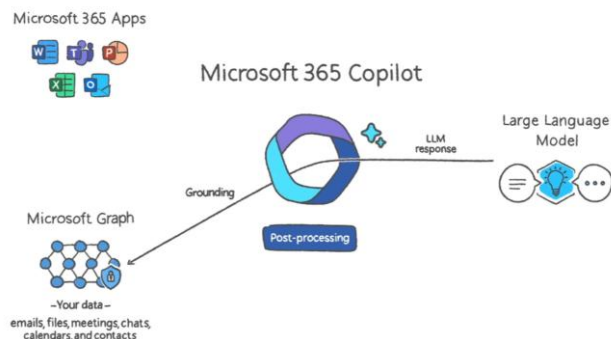
● Microsoft 365 Copilot: 开启全新工作方式

Copilot 是将大型语言模型 GPT-4 内建到 Microsoft 365。在程序里面可以用非常口语化的方式使用 AI，还可以通过各个程序的配合，达到预期效果。

Microsoft 365 Copilot 简化工作流程，释放创造力和生产力：

1. Copilot in Word 能够与用户一起撰写、编辑、总结和创作。
2. Copilot in PowerPoint 能够在创作过程中,通过自然语言命令将想法转化为设计好的演示文稿。
3. Copilot in Excel 能够帮助用户识别趋势,在短时间内创建专业型式的数据可视化。
4. Copilot in Outlook 能够帮助用户整合并管理收件箱。
5. Copilot in Teams 能够直接从对话上下文中提供实时摘要和待办事项。
6. Business Chat 汇集了来自文档、演示文稿、电子邮件、日历、笔记和联系人的数据,能够帮助用户总结聊天内容、撰写电子邮件、查找关键日期,根据其他项目文件制定计划。

图 6: Microsoft 365 Copilot 系统



资料来源：微软官方公众号

● Power Platform Copilot：低代码未来前景打开

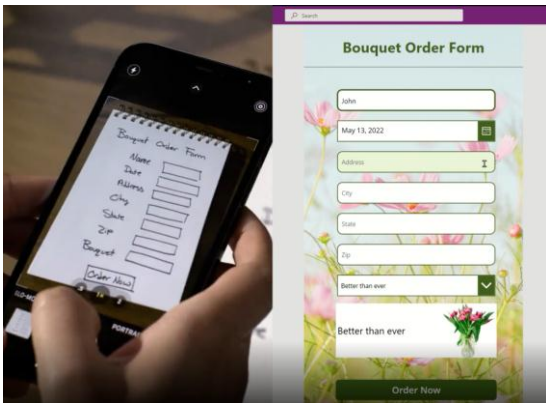
Copilot 是将 AI 结合到各种软件内的功能。Copilot 是 Microsoft Power Platform 的一个新功能，可以在 Power Apps, Power Virtual Agents 和 Power Automate 中基于 GPT 能力提供 AI-powered 的帮助，让制作者可以用自然语言描述他们想要的应用、流程或机器人，然后 Copilot 可以快速完成创建，并提供改进的建议。

(1) Power Apps Copilot 可以让制作者通过自然语言对话的方式来描述并且直接生成想要的应用程序，可以通过和 Copilot 的对话来对应用进行调整，Copilot 也可以为用户提供应用的改进建议来解决问题。

(2) 通过 Power Automate Copilot，可以通过自然语言来生成比以前更复杂的流程，Copilot 还可以以对话的方式不断优化和迭代更新流程，使用户无需专业知识也可以构建想要的流程。

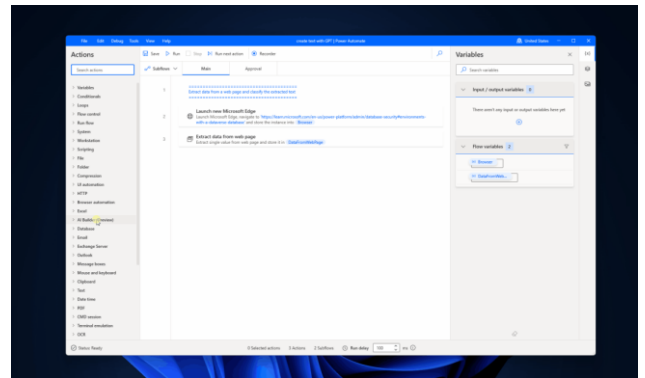
(3) 通过 Power Virtual Agents Copilot，可以直接使用自然语言来描述机器人流程，无需手动一个一个进行创建。

图 7：Power Apps



资料来源：微软官方公众号

图 8：Power Automate



资料来源：微软官方公众号

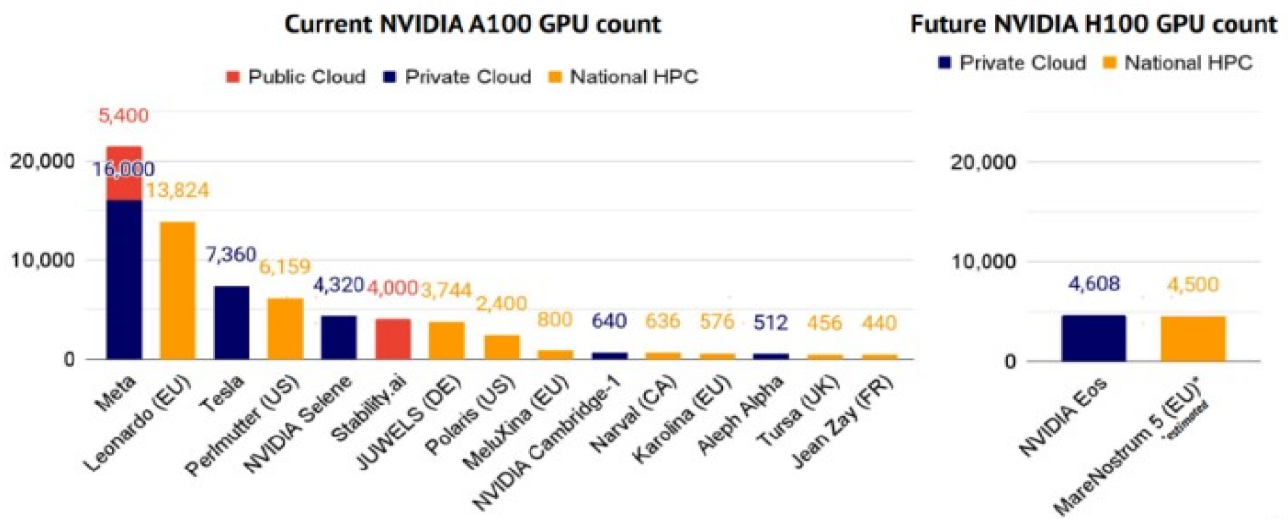
2.2、英伟达:全球 GPU 龙头，产品需求量价齐升

全球 GPU 龙头开拓 AI 市场，助力服务器核心算力硬件性能提升。英伟达作为一家以 GPU 为主营业务的半导体设计公司，在 AI 行业发展初期就前瞻布局。2022 年公司数据中心产品营业收入 150.05 亿美元，同比增长 41.38%，2020-2022 CAGR 高达 49.70%，增势迅猛。ChatGPT 发布引爆市场、Google AI 对话模型 Bard 开放公测、百度发布“文心一言”大模型等一系列互联网巨头纷纷发力 AI 领域，催生 AI 服务器需求增长。据 TrendForce 预测，2023 年 AI 服务器出货量同比增速将达 37.7%，2024 年将增长 38%。公司有望受益于 AIGC 带来的核心算力硬件性能需求大幅提升，AI 服务器系统业务未来将成为公司数据中心产品营业收入的新增长点。

GPU 性能领先全球，H100 拓展市场领先地位。公司深耕 GPU 芯片多年，据英伟达在 IEEE 会议公布数据显示，从 2012 年的 K20X 到 2020 年的 A100，GPU 的推理性能提高了 317 倍，远超摩尔定律的发展速度。A100 作为当前 AI 服务器 GPU 中的高端产品，性能相比上一代 Volta GPU 提升 20 倍，其 AI 推理性能是 CPU 的 249 倍，打破 16 项性能 AI 记录，在 AI、数据分析、高性能计算上应用广泛，目前仍是 AI 领域所用的主力 GPU，被 AWS、微软 Azure 和百度云等。

算力需求迅猛增长，服务器 GPU 业务有望迎来量价齐升。AIGC 产品与应用的逐步落地催生了巨大的算力需求，据 OpenAI 测算，自 2012 年以来，全球头部 AI 模型训练算力需求 3-4 个月提升一倍，每年头部训练模型所需算力增长幅度高达 10 倍。作为 AI 时代的算力核心，GPU 需求旺盛。同时，芯片性能的大幅提升带动 GPU 价格不断增长，英伟达 V100/A100/H100 三代 GPU 价格分别为 10000/15000/36500 美元。作为服务器 GPU 领导者，英伟达将直接受益于 AIGC 大热带来的核心算力硬件性能需求提升，服务器 GPU 业务有望迎来量价齐升。

图 9：2022 年 A100 及未来 H100 各领域应用数量统计



资料来源：Stateof.ai

表 1：英伟达主要芯片产品

英伟达数据中心芯片情况	芯片型号	上市时间	售价（美元）	售价（元）	中国供应情况
中低端	A10	2021 年 4 月 13 日	3200	22080	供货
	A16	2021 年 4 月 13 日	3500	24150	供货
	A30	2021 年 4 月 13 日	4700	32430	供货
	A40	2020 年 10 月 5 日	5300	36570	供货
	L40	2021 年 9 月 21 日	7600	52440	供货
高性能	V100	2017 年 5 月 11 日	10000	69000	供货
	A800	2022 年 11 月 8 日	12000	82800	缺货
	A100	2020 年 5 月 15 日	15000	103500	美国政府禁止供应中国
	H100	2022 年 3 月 21 日	36500	251850	

资料来源：Thinkmate，《财经》，光大证券研究所整理；汇率：1 美元=6.9 人民币

表 2：英伟达 AI 服务器主要产品对比

	NVIDIA DGX A100	NVIDIA DGX H100
GPU	8 个 NVIDIA A100 80GB Tensor Core GPU	8 个 NVIDIA H100 Tensor Core GPU
GPU 显存	共 640GB	共 640GB
性能	5 petaFLOPS AI 10 petaOPS INT8	32 petaFLOPS FP8
NVIDIA® NVSwitch™	6	4x
系统功耗	最大 6.5 kW	最高 10.2kW
CPU	双路 AMD Rome 7742、共 128 个核心、2.25 GHz（基准频率）、3.4 GHz（最大加速频率）	双路 x86
系统内存	2TB	2TB
网络	8 个单端口 NVIDIA ConnectX-7 200Gb/s 的 InfiniBand 端口 2 个双端口 NVIDIA ConnectX-7 VPI2 10/25/50/100/ 200Gb/s 以太网	8 个单端口 NVIDIA ConnectX-6 VPI4 400Gb/s 的 InfiniBand/ 以太网 2 个双端口 NVIDIA ConnectX-6 VPI2 10/25/50/100/ 200Gb/s 以太网 2 个双端口 NVIDIA BlueField-3 DPU VPI 1 个 400Gb/s 的 InfiniBand/ 以太网网卡 1 个 200Gb/s 的 InfiniBand/ 以太网网卡
网络管理		RJ45 接口 10Gb/s 板载网卡 50Gb/s 的以太网可选网卡 RJ45 接口主机基板管理控制器 (BMC) 2 个 NVIDIA BlueField-3 DPU BMC (均为 RJ45 接口)
存储	操作系统：2 个 1.92TB M.2 NVMe 驱动器 内部存储：30TB (8 个 3.84TB) U.2 NVMe 驱动器	操作系统：2 块 1.9TB NVMe M.2 硬盘 内部存储：8 块 3.84TB NVMe U.2 硬盘
系统软件	Ubuntu Linux 操作系统 同时支持：Red Hat 企业级 Linux CentOS 化。	DGX H100 系统预安装 DGX 操作系统，该操作系统基于 Ubuntu Linux，包含 DGX 软件堆栈（所有必要软件包和驱动均针对 DGX 优化）。 客户可以选择单独安装 Ubuntu Linux 或 Red Hat Enterprise Linux 以及必要的 DGX 软件堆栈。
运行温度范围	5°C 至 30°C	5°C 至 30°C

资料来源：英伟达官网，光大证券研究所整理

2.3、 华为：服务器产业深耕多年，打造“鲲鹏+昇腾”双引擎战略布局

深耕服务器产业多年，打造“鲲鹏+昇腾”双引擎战略布局。华为从 2008 年开始对外提供服务器产品，目前形成了通用计算（鲲鹏）和 AI 计算（昇腾）两大业务布局。由于美国将华为列入实体清单，限制 x86 架构的芯片供给，华为在 2021 年 12 月出售了 x86 服务器业务。原本主营 x86 服务器业务的子公司超聚变分离后仍然稳定发展，2022 年国内市占率位列第三，达到 10.1%。目前华为致力于打造以“Arm 架构的鲲鹏+达芬奇架构的昇腾”为根基的双引擎战略布局，借助自研芯片优势，持续拓展服务器业务布局。

坚持计算产业战略，积极布局 AI 服务器领域。华为创新处理器架构“达芬奇”以匹配算力的增速，推出全场景处理器族，其中面向通用计算的鲲鹏系列与面向 AI 计算的昇腾系列主要应用于服务器产品，已服务政府、金融、运营商、电力、交通、医疗、教育、电力、油气、制造等领域，为全球合作伙伴开发服务器应用及解决方案。其中，采用鲲鹏 920/916 处理器的 TaiShan 服务器包括 2280E 边缘型、1280 高密型、2280 均衡型、2480 高性能型、5280 存储型和 X6000 高密型等产品，适配高性能计算、数据库、云计算等各种应用场景，现已应用于中科大千万亿次“瀚海 20 超级计算系统”。最高配置 2480 高端服务器基于鲲鹏 920 处理器，最高提供 256 核、3.0GHz 主频的计算能力和最多 25 个 SSD 硬盘。此外，昇腾系列 AI 服务器包括 Atlas 800 推理服务器与 Atlas 800 训练服务器，广泛应用于中心侧 AI 推理、深度学习模型开发和训练场景。Atlas 800 推理服务器搭载 64 核架构的鲲鹏 920 处理器，支持 8 张 Atlas 300I 推理卡，整机可提供 640 路高清视频实时分析（1080P 25FPS），Atlas 800 训练服务器达到 4U 高度，最高可提供 2.24 PFLOPS FP16 的算力，为 AI 推理、深度学习模型开发和训练提供超强算力支持。

表 3：华为服务器产品

机型	通用计算服务器		AI 计算服务器	
	2480 高性能型	2480 高端型	Atlas 800 训练服务器	Atlas 800 推理服务器
形态	2U 机架服务器	2U 机架服务器	4U AI 服务器	2U AI 服务器
处理器型号	4 个鲲鹏 920 处理器	4 个鲲鹏 920 处理器	4 个鲲鹏 920 处理器+8 个 昇腾 910 处理器	2 个鲲鹏 920，最大支持 8 个 Atlas 300I 推理卡
内存插槽	32 个 DDR4-2933 DIMM 插槽	32 个 DDR4-3200 DIMM 插槽	最多 32 个 DDR4 内存插槽，支持 RDIMM 内存速率最高 3200 MT/s 单根内存条容量支持 16GB/32GB/64GB	32 个 DDR4 内存插槽，最高 3200 MT/s
本地存储	25*2.5 SAS/SATA；8*2.5 SAS/SATA	25*2.5 SAS/SATA；8*2.5 SAS/SATA	2*2.5 SAS/SATA+3*2.5 NVMe 2*2.5 SATA+3*2.5 NVMe 2*2.5 SAS/SATA+6*2.5 NVMe 2*2.5 SATA+6*2.5 NVMe 2*2.5 SATA+8*2.5 SAS/SATA	25*2.5 SAS/SATA 12*3.5 SAS/SATA 8*2.5 SAS/SATA+12x2.5 NVMe
RAID 支持	支持 RAID 0, 1, 5, 6, 10, 50, 60，支持超级电容掉电保护	支持 RAID 0, 1, 5, 6, 10, 50, 60 支持超级电容掉电保护	支持 RAID 0, 1, 5, 6, 10, 50, 60	支持 RAID 0, 1, 5, 6, 10, 50, 60
板载网络	1 个灵活 IO 卡，支持 4*10GE 光口或者 4*25GE 光口；1 个板载 LOM，光口支持 4*GE 电口	1 个板载网络插卡，支持 4*10GE 光口或者 4*25GE 光口；1 个板载 LOM，支持 4*GE 电口	支持 8*100GE 光口+4*25GE 光口 /2*100GE	
PCIe 扩展	最多 9 个 PCIe 4.0 插槽，支持 3 个 PCIe 4.0 x16 和 6 个 PCIe x8 插槽	最多 4 个 PCIe 4.0 x8+2 个 PCIe 4.0 x16 标准插槽	最多支持 2 个 PCIe 4.0 扩展插槽	最多支持 9 个 PCIe4.0 PCIe 接口，其中 1 个为 RAID 扣卡专用的 PCIe 扩展槽位，另外 8 个为标准的 PCIe 扩展槽位
电源	2 个热插拔 900W 交流或 2000W 交流或 1200W 直流或 1500W 直流电源模块，支持 1+1 冗余	2 个热插拔 2000W 交流电源模块，支持 1+1 冗余	4 个热插拔 3 kW/2 kW 交流电源模块，支持 2+2 冗余	2 个热插拔 900 W 或 2000 W 交流电源模块，支持 1+1 冗余备份
供电	支持 100~240V AC, 240V DC, -48V DC	支持 220~240V AC, 240V DC	支持 200~240 V AC, 240 V DC	
风扇	支持 6 个热拔插风扇模组，支持 N+1 冗余	支持 6 个热拔插风扇模组，支持 N+1 冗余	支持 8 个热拔插风扇模组，支持 N+1 冗余	4 个热拔插风扇，支持 N+1 冗余备份
温度	5°C~40°C	5°C~35°C	工作温度（风冷）：5°C~35°C； 工作温度（液冷）：5°C~40°C	5°C~40°C
尺寸(宽 x 深 x 高)	447 mm*790 mm*86.1 mm	447 mm*790 mm*86.1 mm	175 mm * 447 mm * 790 mm	447 mm * 790 mm * 86.1 mm

资料来源：华为官网，光大证券研究所整理

3、AI 产业链：大模型加速落地，算力空间广阔，应用百花齐放

3.1、大模型：国内和海外巨头加速布局

在“基础设施支撑+顶层设计优化+下游需求旺盛”三轮驱动下，AI 大模型迎来了良好的发展契机。通过数据交互和任务反馈，优秀的大模型能够赋能各行各业开放任务，满足对未来 AI 应用的期待。展望未来，大模型“训练基础设施-底层技术-基础应用-垂直应用”发展路线逐渐清晰，随着底层技术逐步革新，基模型和领域大模型持续完善，大模型应用边界不断拓宽，将加速赋能交通、医疗、金融等各个行业和领域引发一场以强人工智能和通用人工智能为代表的新一轮智能革命浪潮，大幅提高生产和生活效率，带来深刻的经济、社会和产业变革。

图 10：优秀的大模型能够赋能各行各业开放任务



资料来源：沙利文，光大证券研究所

图 11：大模型行业的国外厂商



资料来源：沙利文，光大证券研究所

图 12：大模型行业的国内厂商



资料来源：沙利文，光大证券研究所

3.2、服务器：AIGC 推动 GPU 服务器市场持续扩张

AI 训练服务器方面，单个大语言模型训练驱动 AI 训练服务器市场约 2 亿美元需求。据英伟达测算，要完成一次 10000 亿参数大语言模型的训练，使用 4096 颗 A100 GPU 约需 4 周时间，使用 4096 颗 H100 GPU 约需 1 周时间。按每台服务器配备 8 颗 H100 GPU 计算，则完成一次万亿参数的大语言模型训练需约 500 台服务器，配备 8 个 H100 GPU 的服务器价格约为 40 万美元，则单个大语言模型训练驱动 AI 训练服务器约 2 亿美元需求。

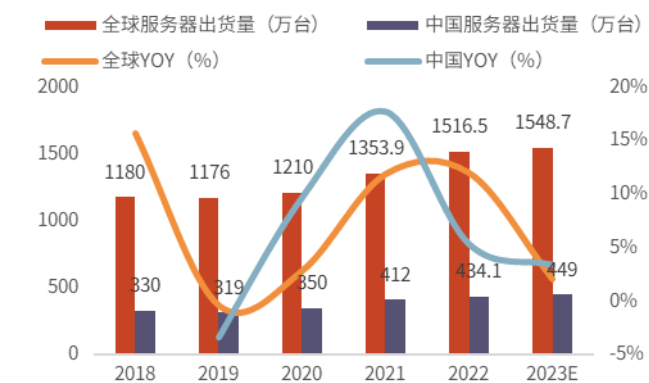
图 13：英伟达 H100/A100 对比



资料来源：英伟达官网，光大证券研究所

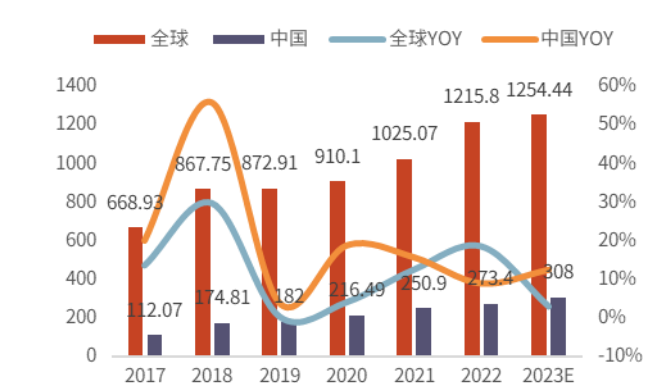
海内外服务器市场持续增长，国内市场一超多强。全球方面，根据 IDC 数据，2022 年全球服务器出货量突破 1,516 万台，同比增长 12%，产值达 1,215.8 亿美元，预计未来四年内仍保持增长，2026 年全球服务器出货量达 1885.1 万台，2022-2026 年全球服务器出货量 CAGR 为 5.59%，产值有望达到 1664.95 亿美元，2022-2026 年总产值 CAGR 为 8.18%。**中国方面**，据智研咨询数据，2022 年中国服务器出货量为 434.1 万台，同比增长 5.36%。据 IDC，2022 年中国服务器市场规模为 273.4 亿美元（1888.37 亿人民币），同比增长 9.1%。据中商情报网预测，2023 年中国服务器出货量有望达到 449 万台，同比增长 3.43%；中国服务器市场规模有望达到 308 亿美元，同比增长 12.66%。**从市场份额（按营业收入统计）**上看，据 IDC，浪潮、新华三、超聚变、宁畅、中兴位居 2022 年中国服务器市场前五，市场份额占比分别为 28.1%/ 17.2%/ 10.1%/ 6.2%/ 5.3%，同比变动-2.7pct/ -0.3pct/ +6.9pct/ 0pct / +2.2pct。

图 14：全球服务器与中国出货量及增速



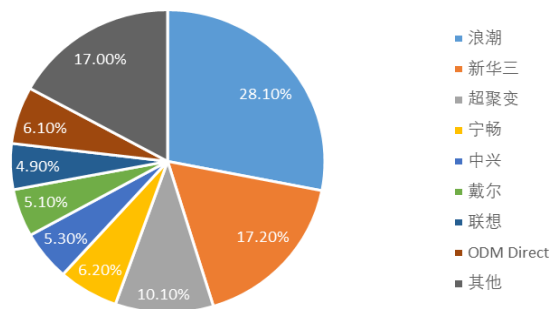
资料来源：IDC，中商情报网，光大证券研究所，IDC 与中商情报网预测

图 15：全球与中国服务器市场规模及增速（单位：亿美元，%）



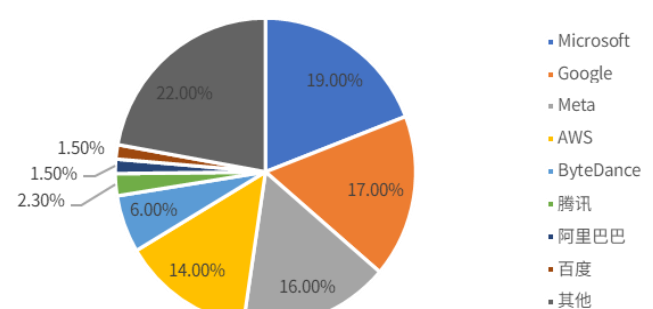
资料来源：IDC，华经产业研究院，中商情报网，光大证券研究所整理，中商情报网预测

图 16：2022 年中国服务器市场份额（按营收）



资料来源：IDC，光大证券研究所

图 17：2022 年 AI 服务器采购量占比



资料来源：TrendForce，光大证券研究所

AI 服务器有望迎高速增长，海内外云巨头加大 AI 服务器采购。ChatGPT、文心一言等 AIGC 大模型带来计算资源需求井喷，催生 AI 服务器需求增长，叠加配置升级带动单台 AI 服务器价值较通用型成倍增长，量价齐升。据 TrendForce 预测，2023 年全球 AI 服务器出货量同比增速可达 8%，2022~2026 年复合增长率将达 10.8%。据 IDC 数据，2021 年全球 AI 服务器市场规模 156 亿美元，预计到 2025 年全球 AI 服务器市场将达到 318 亿美元，21-25 CAGR 为 19.5%；2021 年中国 AI 服务器市场规模达 350 亿元，预计 2025 年中国 AI 服务器市场规模将达到 702 亿元，21-25 CAGR 为 19.0%。从 AI 服务器的采购量来看，据 TrendForce

统计，2022 年 AI 服务器采购量中，北美四大云端供应商 Microsoft、Google、Meta、AWS 合计占比约 66%，而中国近年来 AI 建设浪潮持续升温，字节跳动 2022 年采购占比达 6.2%，腾讯、阿里巴巴、百度紧接其后，分别约为 2.3%、1.5%与 1.5%。海外云巨头对 AI 服务器的需求更大，但随着国内 AI 大模型的开发及应用拉动更多 AI 服务器需求，中国 AI 服务器市场空间有望进一步提升。

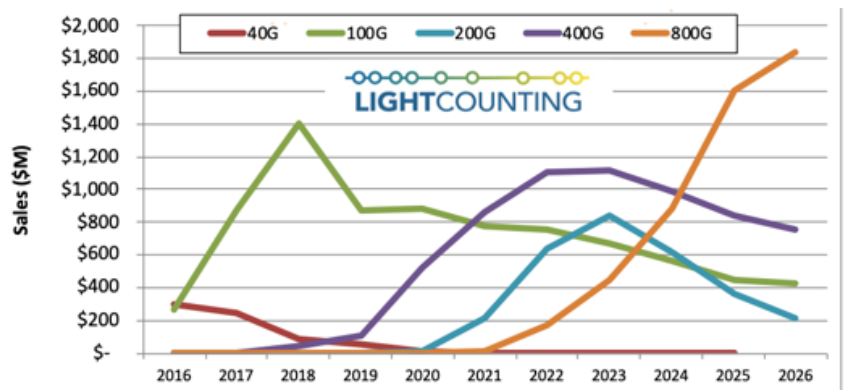
IDC 认为，中期来看 X86 架构仍然会是主流的服务器，未来将会形成以 X86 为核心多种架构并存的丰富算力市场。另外，随着互联网、云服务、电信、银行等各行业超大规模客户不断扩张和升级扩容对于服务器的强劲需求，以硬件为中心的集成硬件与软件支持多样化工作负载的解决方案将成为服务器厂商制胜未来的关键。当前服务器产业正往技术多样化、服务定制化、软件定义、异构计算与边缘计算方向发展。

芯片作为服务器算力的载体，与服务器整体性能紧密相关。传统服务器芯片组仅由 2 颗 CPU 构成，价格在 800-850 美元/颗，而 AI 服务器芯片组通常需要搭载 2 颗 1000-1100 美元/颗的 CPU，训练 AI 服务器与推理 AI 服务器还分别需要增加 8/4 颗 GPU 以提升服务器的综合算力。目前完整规格的英伟达 HGX H100 AI 服务器搭载配备 8 个 H100 GPU，总计 8 个 GPC 图形集群，72 个 TPC 纹理集群，144 个 SM 流式多处理器，共计 18432 个 FP32 CUDA 核心，每组 SM 配备 4 个专用于 AI 训练的第四代张量核心 TensorCore，共计 528 个。相较于同等 GPU 数量配置的 A100，H100 在 HPC 和 AI 计算 (FP64/TF32/FP16/INT8)上由 156TF/2.5PF/5PF/10POPS 提升至 480TF/8PF/32PF/32POPS，在多专家 (MoE) 模型训练速度上提升了 9 倍，大幅度提升万亿参数的 AI 大模型应用效率。在 AIGC 快速发展背景下，服务器芯片配置或成服务器厂商制胜关键。

3.3、光模块：800G 需求释放，中国厂商厚积薄发

AI 大模型发展加速光模块迭代节奏，800G 光模块放量或将超预期：根据 Lightcounting 2021 年的预测，800G 光模块将从 2025 年底开始主导市场，主要原因是 AI 应用等带来数据流量的增长，超预期的数据中心带宽需求以及光模块厂商技术的迭代。随着以 Chat GPT4 为代表的 AI 大模型问世以及相关应用的快速发展，AI 训练和推理服务器有望进一步推升高速率光模块的需求。目前头部云厂商已开始集中测试和采购 800G 光模块，我们预计 800G 光模块出货量将在 23 年下半年快速增长，放量节奏将超出此前 Lightcounting 预期。

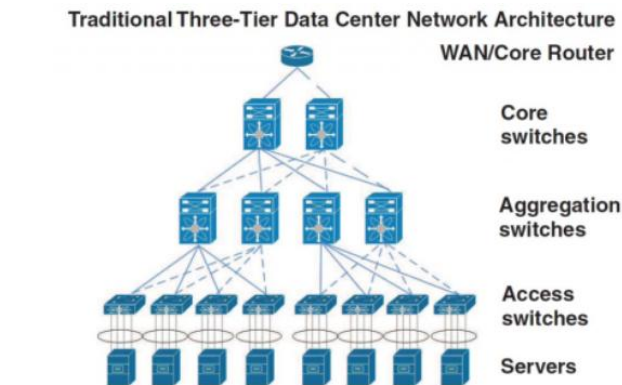
图 18：数据中心光模块需求演进（21-26 年数据为预测值）



资料来源：Lightcounting（2021 年预测）

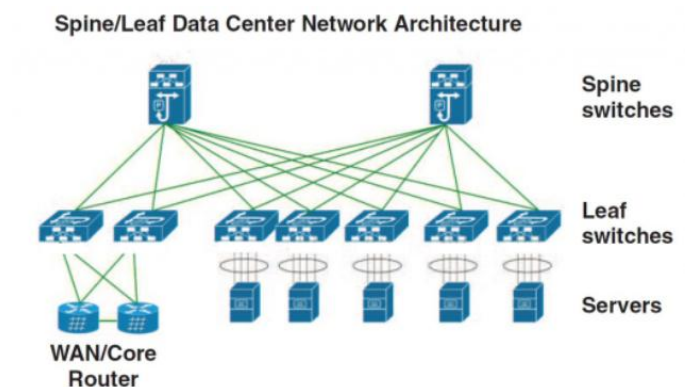
AI 服务器促使数据中心网络架构的演变，从而提升光模块需求：普通 GPU 服务器一般只要求单卡性能，而在 AI 训练中，GPU 卡间需要大量的参数通信，模型越复杂，通信量越大，所以 AI 服务器除了要求单卡性能外，还要求多卡间的通讯性能。这促使数据中心整体架构由传统的三层网络架构转化为叶脊式架构，数据流量从南北向到东西向转变。叶脊使用所有的互联链路，每台叶交换机都连接到了脊交换机上，脊交换机之间和叶交换机之间没有任何互连，相对于传统的三层网络架构而言，减少了设备寻找或等待连接的需求，从而减少了延迟及流量瓶颈。由于其特殊的拓扑结构，系统对于光模块的需求得到了提升。

图 19：传统三层网络架构



资料来源：fiber-optic-tutorial

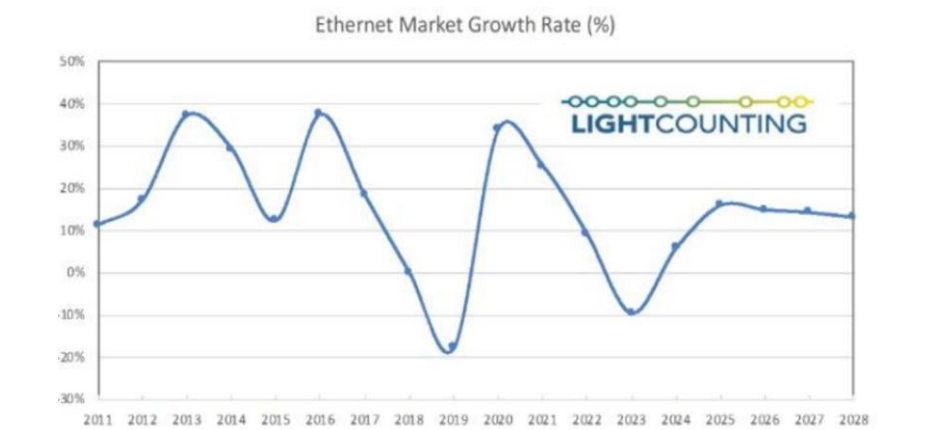
图 20：叶脊架构



资料来源：fiber-optic-tutorial

随着数据中心对于带宽需求的提升，叶脊架构中的光模块也在持续升级之中。速率方面，亚马逊、谷歌、微软、Facebook 等北美超大型数据中心内部互连已经在 2019~2020 年开始商用部署 400Gb/s 光模块；国内数据中心正由 100Gb/s 逐步向 400Gb/s 过渡。根据 IMT2020（5G）推进组预计，数据中心交换芯片吞吐量 2023 年将达到 51.2Tb/s，2025 年之后达到 102.4Tb/s，800Gb/s 和 1.6Tb/s 更高速率将成为实现高带宽数据交换的重要选择。

图 21：以太网光模块销售额增长率（23-28 年数据为预测值）



资料来源：Lightcounting 2023 年预测

Lightcounting 预测光模块市场 23 年将出现 10% 下滑，但 800G/1.6T 光模块预计仍维持高增：根据 LightCounting 23 年 3 月发布的最新报告，其下调了 2023 年以太网光模块销售的预测，从 22 年 10 月预测的 2% 的增长到现在 10% 的下降，24-28 年的复合增长率预期仍为 13%。下降的主要原因是云巨头 Meta 公司放缓了数据中心的建设部署，支持元宇宙业务的数据中心升级计划许多都被搁置

或缩小了规模。但 Meta 公司确认了其建立人工智能基础设施，并追赶亚马逊和谷歌的承诺，目前大多数 400G 光模块使用的场景，接下来将使用 800G 光模块。LightCounting 并未太多的削减 Meta 对 400G/800G 光模块的需求预测，LightCounting 对 1.6T 光模块的出货量预测没有改变，这些模块需要满足谷歌、亚马逊和微软运营的数据中心每年 40%带宽增长的需求。

800G LPO 方案或将成为 AI 服务器集群中最具性价比方案：LPO(Linear-drive Pluggable Optics)是线性驱动可插拔光模块，在数据链路中只使用线性模拟元件，无 CDR 或 DSP 芯片的设计方案。在 OFC 2023 上，多家厂商展示了其 linear-drive 方案，包括 Macom、Broadcom 和 Cisco 等公司。相较 DSP 方案，LPO 可大幅度减少系统功耗和时延（功耗相较 DSP 可下降接近 50%，与 CPO 的功耗接近），而系统误码率和传输距离有所牺牲，深度契合目前 AI 计算中心的短距离、大带宽、低功耗、低时延的数据连接需求。建议积极关注国内光模块公司在 800G LPO 方案的进展。

国内光模块厂商彰显强大竞争力，在 800G 时代有望持续扩大份额：根据 Lightcounting，2021 年全球前十大光模块厂商，中国厂商占据六席，分别为中际旭创（与 II-VI 并列第一）、华为海思（第三）、海信宽带（第五）、光迅科技（第六）、华工正源（第八）及新易盛（第九）。在 800G/1.6T 技术方面，国产厂商也都有深度布局，未来有望持续受益于 800G 光模块需求的放量。在 2023 年 OFC 大会上，各大国内厂商发布并展示最新 800G/1.6T 光模块产品，例如：中际旭创发布基于 5nm DSP 和先进硅光子技术的第二代 800G OSFP DR8+和 2xFR4 光通信模块，以及 1.6T OSFP-XD DR8+可插拔光通信模块；新易盛推出多款 800G 光模块，其中基于薄膜铌酸锂(TFLN)调制器的 800G OSFP DR8 模块功耗仅为 11.2W，是目前业界最节能的光模块；光迅科技首次展出业界最新的数据中心内部高速互联 1.6T OSFP-XD DR8 和 800G QSFP-DD800 SR8 光模块。

3.4、边缘算力：AIoT 智能终端的“大脑”，端侧算法部署的核心

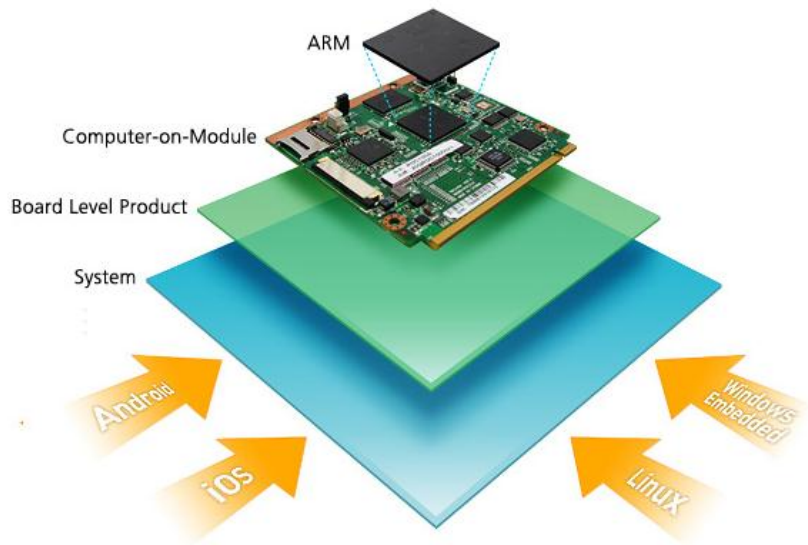
边缘算力应用和端侧 AI 算法部署需求巨大

智能应用处理器芯片均为系统级的超大规模数字集成电路，可以分为高性能应用处理器、通用应用处理器、人工智能视觉处理器、智能语音处理器、车载处理器、流媒体处理器等。上述处理器一般内置中央处理器（CPU），根据使用场景的需要增加图形处理器（GPU）、图像信号处理器（ISP）、神经网络处理器（NPU）及多媒体视频编解码器等处理内核。芯片内部设置高速总线负责各个处理器和外部接口的数据传输，配备闪存接口、动态存储器接口、显示接口、网络接口以及各种高速、低速外部设备接口。

SoC 的关键技术主要包括总线架构技术、IP 核可复用技术、软硬件协同设计技术、SoC 验证技术、可测性设计技术、低功耗设计技术、超深亚微米电路实现技术，以及嵌入式软件移植、开发研究，是一门跨学科的研究领域。SoC 意味着在单个芯片上实现以前需要多个芯片才能实现的系统功能，克服了多芯片板级集成出现的设计复杂、可靠性差、性能低等问题，并且在减小尺寸、降低成本、降低功耗、易于开发等方面也有突出优势。SoC 对研发设计、制造工艺以及软硬件协同开发技术的要求较高，主要体现在芯片验证和测试难度的提高，以及 IP 复用、混合电路设计的困难加大。任何 SoC 的设计都是性能、功耗、稳定性、工艺难度几方面的平衡。完整的 SoC 系统解决方案，除了提供硬件参考设计外，还需要提供系统级的软件参考设计，包括驱动软件、大型 OS（Linux、Android、

ChromeOS、国产 OS 等) 的移植、针对性的算法、中间件和上层应用软件的适配等, 并通过严格的可靠性、兼容性测试。

图 22: SoC 示意图



资料来源：华强电子网、光大证券研究所

3.5、PCB：AI 算力提升驱动行业快速成长

服务器市场增长带动 PCB 市场需求增长。PCB 是服务器的重要组成部分, 是承载服务器运行的关键材料。服务器出货量的大幅增长也使得服务器 PCB 市场规模扩容, 成为 PCB 市场中复合增长率最快的下游细分市场。我们认为, 随着 AI 算力需求的增长, PCB 市场有望迎来回暖。Prismark 统计数据显示, 目前全球 PCB 市场规模稳定增长。Prismark 测算 2022 年其产值达 821 亿美元, 同比 +1.5%。展望未来, 伴随宏观影响边际减弱, 整体需求稳步复苏, 叠加服务器及数据中心、汽车电子、AIoT(智能耳机、智能手表、AR/VR 等) 新兴应用放量及技术升级, PCB 产值有望稳健成长, Prismark 预计 2026 年全球产值达到 1016 亿美元, 2022-2026 年 CAGR 超 5%。受益于下游应用技术规格持续迭代升级, 对线路板要求亦不断提升, PCB 下游中高端化产品如 HDI、封装基板等产值占比显著提升。

3.6、智能终端：AI 赋能硬件入口，加速万物互联

生成式 AI 的发展或使消费电子产品与 AI 的结合成为新常态, 实现万物互联互通。生成式 AI 商业化进程的加快将吸引更多终端产品接入, 加快海量终端与人实现万物互联的进程, 实现 AI 全方位覆盖。人工智能交互技术和智能硬件产品的结合将广泛应用于智能家居、智能出行、智能办公等多种场景, 带来更为多元的增长。

(1) AI+智能手机: AI 赋能使手机应用更便捷。在苹果 2023 财年二季度财报会议上, 库克表示, AI 功能已经嵌入到各种苹果产品中, 苹果将继续有序地将 AI 融入设备中。iOS 17 系统可能会嵌入一个名叫“Personal Voice”的语音功能, 允许用户在 iPhone 和 iPad 上合成自己的声音, 用于面对面对话、电话、FaceTime 视频通话以及音频通话。谷歌在 I/O 开发者大会上宣布了安卓 14 系统引入两项生成式 AI 特性。这两项特性分别是 Magic Compose 和生成式壁纸

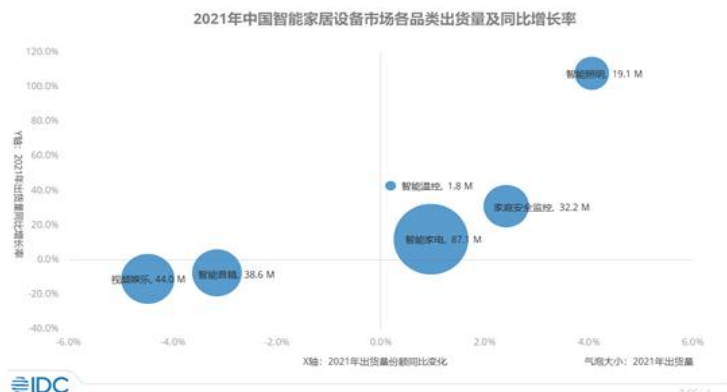
(Pixel 设备)。Magic Compose 是一种基于生成式 AI 技术的文本输入建议功能。使用者只需要输入一部分句子或单词，Magic Compose 就能够根据上下文自动推荐相关的文字，加快了输入的速度和准确性。生成式壁纸可以根据用户的喜好和主题选择，通过生成式 AI 技术创造出各种独具特色的壁纸。

(2) AI+智能音箱：智能音箱经历了三年极速发展期和两年疫情调整期后，智能音箱已完成首轮用户普及，2022 年市场出现下行趋势。据 RUNTO 数据，2022 年中国智能音箱市场销量为 2631 万台，同比下降 28%；市场销售额为 75.3 亿元，同比下降 25%。伴随人工智能交互技术与智能硬件产品结合的趋势，智能音箱逐步与智能家居的其他产品相结合，正在发挥智能家居的控制中心作用，有望为声学行业带来更为多元的增长。据 RUNTO 预测，2023 年中国智能音箱市场销量将达到 2715 万台，同增 3%，止跌回升可期。

(3) AI+智能机顶盒：机顶盒于全球市场的需求、迭代周期及各个国家所处的发展阶段具有较大的差异性，中国、美国的发展相对较快，东南亚、印度、非洲、中东、拉美、欧美等国家及地区仍处于数字化、高清化、超高清化、融合化的不同阶段，且智能化、多功能等在不断迭代演进。基于全球机顶盒行业十几年的发展及存量与迭代更新的视角，全球机顶盒行业仍处于数字化变革、融合并朝着更加智能化、多功能化方向递进的阶段，需求、产品及应用等具有多层次和多样性，未来全球数字机顶盒市场的需求量仍将保持稳定增长的态势。此外，随着新一代信息技术，IoT、AIGC 等技术的发展，机顶盒将逐步融合人工智能、物联网、云计算和大数据等技术，结合 NLP（自然语言处理）、CV（计算机视觉）等 AIGC 技术能力，搭载 AIGC 相关应用系统、平台，除电视直播与超高清视频之外，智能盒子具有的核心能力会成为智慧家庭的 AI 智慧助手、智能控制中心、物联网中心及多媒体信息中心，进一步推动机顶盒更新需求。

(4) AI+智能家居：据 Statista 数据，智能家居是目前 AIoT 场景中发展最快的领域，占物联网市场收入的 96%，将成为最具潜力的蓝海市场之一。此外，当前以智能音箱、智能机顶盒等产品为代表的智能家居大多接入判别式 AI，在交互模式上仅能对相应问题及指令做出简单回应，而以 ChatGPT 为代表的生成式 AI 模型提升了对用户意图的理解，对用户的反馈更加准确丰富，如对用户的个性化服务、快速收集与处理信息、提升场景交互效率等，有望大幅提升智能家居场景体验，生成式 AI 将赋予智能设备更强的产品竞争力。

图 23：2021 年中国智能家居设备市场各品类出货量及同比增长率



资料来源：IDC、光大证券研究所

4、投资建议

图 24：中国 AI 产业链梳理



资料来源：光大证券研究所整理

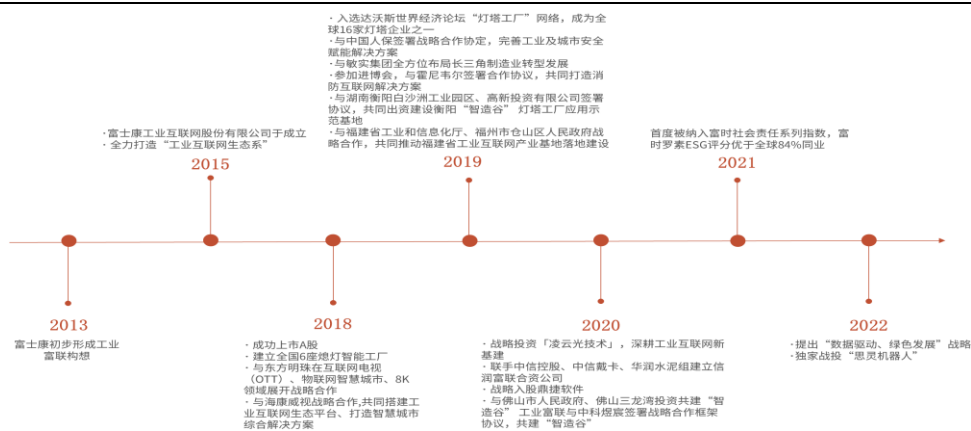
推荐：我们认为 AI 浪潮之下应重点关注 AI 硬件赛道。ChatGPT 开启 AI 发展新浪潮，算力紧缺和海量应用驱动 AI 硬件广阔空间。云端算力（GPU/光模块/CPU/AI 芯片/FPGA/CHIPLET/PCB）、边缘算力（SoC/AIoT/RISC-V）、AI 应用、视频视觉四条主线布局 AI 硬件赛道。推荐：工业富联（AI 服务器驱动增长）、大华股份（智慧物联龙头）、海光信息（CPU/GPU 领军企业）、寒武纪（A 股 AI 算力芯片龙头）、芯原股份（稀缺的国产 IP 供应商）。

5、重点公司推荐

5.1、工业富联（601138.SH）：消费电子龙头，AI 服务器打开成长空间

公司概述：工业富联为全球领先的高端智能制造及工业互联网解决方案服务商，主要业务包含云计算、通信及移动网络设备、工业互联网。2015 年，工业富联成立，宣布将全力打造“工业互联网生态系”；2018 年，工业富联成功上市 A 股。公司深耕中国市场，并布局全球，业务范围覆盖数据全流程，是全球领先的智能制造及工业互联网整体解决方案服务商。公司 22 年在受到国际政治、经济形势复杂多变的影响下，仍实现营业收入 5118.50 亿元，同增 16.4%；归母净利润 200.73 亿元，同增 0.3%，归母净利润创历史新高。

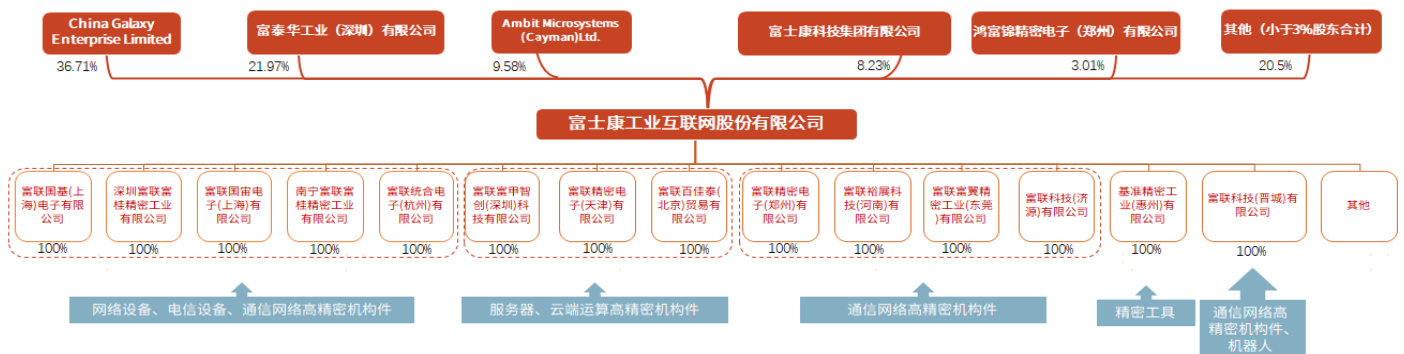
图 25：公司发展历程



资料来源：公司官网，光大证券研究所整理

股权结构：公司股权结构稳定，子公司数量众多涉及业务广泛。公司第一大股东为由鸿海精密间接持有 100% 股权的 China Galaxy Enterprise Limited（中坚企业有限公司）。工业富联子公司中富联国基(上海)电子有限公司、深圳富联富桂精密工业有限公司等主要负责网络设备、电信设备、通信网络高精密机构件业务；富联富甲智创(深圳)科技有限公司、富联精密电子(天津)有限公司、富联百佳泰(北京)贸易有限公司主要负责服务器、云端运算高精密机构件业务；基准精密工业(惠州)有限公司主要负责精密工具业务；富联科技(晋城)有限公司主要负责通信网络高精密机构件、机器人业务。

图 26：公司股权结构（截至 2023 年 11 月）



资料来源：Wind，光大证券研究所整理

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/176025045100010101>