



计算机行业研究

买入（维持评级）

行业深度研究

证券研究报告

计算机组

分析师：陈奕骄（执业 S1130523020001） 分析师：孟灿（执业 S1130522050001）

chenyijiao@gjzq.com.cn

mengcan@gjzq.com.cn

算力深度报告一：算力研究框架-产业链全梳理

行业观点

■ **算力，AI 大模型时代的能源动力。**通用人工智能技术有望启动以算力为主体能源的第四次工业革命，算力需求有望高增。技术革命是以新一代生产力工具的大范围应用为核心，以主体能源为基础动力的全局性经济范式转型。历史上共发生过三次技术革命，我们认为，本轮通用型人工智能发展浪潮有望启动第四次技术革命：以 AGI 为新一代生产力工具，以算力为主体能源。历史上历次技术革命中主体能源都会随新一代生产力工具的广泛应用而需求高增，我们预计 AI 算力需求有望在模型规模、算法创新、海量数据、应用落地等多端推动下高增。

■ **AI 模型训练规模增速超过摩尔定律，进入大模型时代。**在 2010 年以前（前深度学习时代），AI 训练计算的增长与摩尔定律一致，大约每 20 个月翻一倍。2010~2015 年（深度学习时代），AI 深度学习训练所需 FLOPs 相比之前发生斜率突变，每 6 个月翻一倍，远超摩尔定律。2016 年至今（大模型时代），AI 大模型训练所需 FLOPs 出现断点跳跃，随着能够处理更长序列和更复杂任务的 Transformer 等模型问世，算法的持续创新推动了对更高计算能力的需求，2022 年训练大模型的算力需求达到了 8×10^{23} FLOPs。

■ **透视算力产业链——上游：芯片/元器件→中游：服务器/网络设施→下游：IDC/云服务。**

1) **上游-芯片/元器件——异构算力蓬勃发展，国产替代空间广阔。**随着业界对算力需求的不断提高，越来越多的计算平台开始引入 CPU、GPU、ASIC、FPGA 等多种不同计算单元来进行加速计算。2020 年我国 MPU 市场规模达到 2,336 亿元，约占全球 39.7%；2022 年全球 GPU 市场规模达 422 亿美元，NV 占全球 GPU 市场份额 80%以上；2022 年全球 ASIC 市场规模在千亿体量，我国约占全球市场 33.4%；2023 年全球 FPGA 规模近百亿美元，中国市场约占 37.9%。

2) **中游-服务器/网络设备——国内服务器厂商领军全球，光模块市场份额迅速提升。**2022 年全球服务器市场高达千亿美元，我国约占全球市场四分之一，国产厂商浪潮信息在 2023 年 Q3 全球服务器市场份额排名第二，仅次于戴尔；2022 年光模块全球前 10 名中国占据 6 家，中际旭创与 Coherent 市占率并列全球第一；2022 年全球交换机市场规模达 400 亿美元，中国市场约占 17%；而全球存储市场则基本被国外厂商垄断，不过随着 AI Server、AI PC、AI Mobile 持续放量，国内存储市场有望迎来新发展机遇。

3) **下游-数据中心/云服务——东数西算铸基，算力租赁牵引。**受新基建、数字化转型及数字中国远景目标等国家政策促进及企业降本增效需求的驱动，我国数据中心市场规模持续高速增长。2021 年我国数据中心市场规模达到 1,500 亿元，近五年 CAGR 达到 29.96%；中国云计算产业正处蓬勃发展阶段，2022 年中国云计算市场规模达到 4,552.4 亿元，同比增长 33.5%。

投资建议

■ 推荐关注 AI 算力产业相关计算机标的，浪潮信息、中科曙光、宝信软件、神州数码、高新发展。

风险提示

■ 国产替代进程不及预期的风险；数据中心能耗政策推进不及预期的风险；技术变化较快，技术路线调整的风险；行业竞争加剧的风险。



内容目录

1. 算力——AI大模型时代的能源动力	6
1.1 技术革命意味着什么？——新一代生产力工具（AGI）问世推动主体能源（算力）需求快速上行	6
1.1.1 AGI曙光：通用人工智能技术有望启动第四次工业革命，算力为主体能源	6
1.1.2 押注能源：历次技术革命的主体能源需求量受益于生产力解放快速上行	6
1.2 AI算力需求有望在模型规模、算法创新、海量数据、应用落地等多端推动下高增	7
1.2.1 AI三要素：算法是车间工艺，算力是能源动力，数据是生产资料	7
1.2.2 算力需求跃迁：AI模型训练规模增速超过摩尔定律，引发算力需求高增	7
1.2.3 算力分类：训练算力+推理算力，算力体系逐步完善	8
1.2.4 算力规模：训练/推理对AI芯片提出不同要求，推理端算力需求有望远超训练端	9
2. 透视算力产业链——上游：芯片/元器件→中游：服务器/网络设施→下游：IDC/云服务	11
3. 上游-芯片/元器件——异构算力蓬勃发展，国产替代空间广阔	12
3.1 CPU：中央处理器，Intel与AMD两大巨头占据x86 CPU市场	12
3.1.1 基本定义：CPU负责解释计算机指令以及处理计算机软件中的数据	12
3.1.2 CPU分类：基于指令集的CISC与RISC，基于应用领域的MPU/MCU/DSP	13
3.1.3 产业链条：上游-EDA/IP/半导体设备/测试，下游-整机厂商/行业解决方案/应用系统	15
3.1.4 市场规模：全球微处理器市场规模高达千亿美元，我国约占全球市场规模40%	15
3.1.5 竞争格局：Intel约占全球MPU市场半壁江山，Intel与AMD独占x86全球市场	16
3.2 GPU：CUDA大幅降低GPU并行计算的编程难度，NV约占全球市场九成份额	18
3.2.1 基础定义：CPU侧重低延迟计算，GPU更侧重并行计算	18
3.2.2 CUDA：大幅降低GPU并行计算的编程难度，实现GPU的通用化	19
3.2.3 市场规模：2022年全球GPU市场规模达422亿美元，未来十年有望高速增长	20
3.2.4 竞争格局：NV约占全球GPU市场份额近九成，AMD与Intel约占10%	20
3.3 ASIC：算力由通用走向定制，2022年我国ASIC市场规模达335亿元	21
3.3.1 基础定义：专用集成电路，算力从通用走向定制	21
3.3.2 市场规模：2022年全球ASIC市场规模在千亿体量，我国约占全球市场33.4%	21
3.3.3 竞争格局：竞争格局开放，国产替代正当时	22
3.4 FPGA：现场可编程芯片，Xilinx约占全球市场份额52%	22
3.4.1 基础定义：现场可编程的芯片，FPGA=LC+IO+SB	22
3.4.2 功能实现：编程LC与SB改变逻辑电路进而实现不同的功能	23
3.4.3 市场规模：全球FPGA规模或近百亿美元，中国市场约占37.9%	23
3.4.4 竞争格局：Xilinx约占全球FPGA市场份额52%，安路科技为国产龙头厂商	24
4. 中游-服务器/网络设备——国内服务器厂商领军全球，光模块市场份额迅速提升	24
4.1 服务器：管理计算资源的大型计算机，全球市场规模超千亿美元	24
4.1.1 基本定义：服务器在网络中为其它客户机提供计算或应用服务	24
4.1.2 产品分类：基于体系架构方式/基于应用层次/基于用途/基于机箱结构	24
4.1.3 产业链条：上游-硬件设备/软件产品，下游-应用市场	25
4.1.4 市场规模：2022年全球服务器市场高达千亿美元，我国约占全球市场四分之一	25
4.1.5 竞争格局：戴尔/浪潮/HPE全球市场CR3，浪潮/新华三/超聚变引领国内市场	26



4.2 光模块：实现光电信号互相转换，国产厂商占全球市场份额 50%以上	27
4.2.1 基本定义：光模块的功能是实现光电转换，是光通信系统的核心器件.....	27
4.2.2 产品分类：分类方式多样，小型化、高速率、低功耗、低成本是发展趋势.....	27
4.2.3 产业链条：上游-元器件等，下游-通信设备制造/数据中心/电信运营商.....	28
4.2.4 市场规模：全球光模块市场规模突破百亿美元，中国市场约占四分之一.....	28
4.2.5 竞争格局：国产厂商份额不断提升，占据全球光模块市场份额超过 50%	29
4.3 交换机：提供网络连接端口，思科独占全球市场份额 40%.....	29
4.3.1 基本定义：交换机为接入它的任意两个网络节点提供独享的电信号通路.....	29
4.3.2 产品分类：基于网络覆盖范围/基于传输速率和介质/基于应用规模等	29
4.3.3 产业链条：上游-电子元器件，下游-数据中心等领域	30
4.3.4 市场规模：全球交换机市场规模达 400 亿美元，我国约占全球市场规模 17%.....	30
4.3.5 竞争格局：思科是全球交换机市场绝对龙头，华为、新华三、锐捷合计占国内市场超过八成.....	31
4.4 存储设备：国外厂商垄断全球存储市场，三星、海力士、美光合计占比超过 80%.....	31
4.4.1 基本定义：存储器是用来存储程序和各种数据信息的记忆部件	31
4.4.2 产品分类：NAND 和 DRAM 市场是主要细分市场	31
4.4.3 市场规模：全球 NAND 市场规模达 600 亿美元，全球 DRAM 市场规模近千亿美元	32
4.4.4 竞争格局：国外厂商垄断全球存储市场，三星据 NAND、DRAM 市场份额首位	33
5. 下游-数据中心/云服务——东数西算铸基，算力租赁牵引	33
5.1 数据中心：数据存储和交换的中心，Equinix 是全球 IDC 龙头	33
5.1.1 基本定义：拥有完善的设备、专业化的管理、完善的应用级服务的服务平台	33
5.1.2 产业链条：上游-基础设施，中游-数据中心运营服务商，下游-应用行业	34
5.1.3 市场规模：全球 IDC 市场达 800 亿美元，我国 IDC 市场规模在 2,500 亿元左右	34
5.1.4 竞争格局：Equinix 是全球 IDC 龙头，国内市场三大运营商合计占比超 50%.....	35
5.2 云服务：分布式计算，AWS 独占全球市场份额 40%.....	35
5.2.1 基本定义：通过网络以按需、易扩展的方式获得所需服务	35
5.2.2 产品分类：基础设施即服务(IaaS)/ 平台即服务(PaaS)/ 软件即服务(SaaS).....	35
5.2.3 产业链条：上游-芯片/基础设备，中游-IDC 制造商/云服务提供商，下游-客户端	36
5.2.4 市场规模：全球云计算规模达 4,000 亿美元，中国市场规模 4,500 亿元	36
5.2.5 竞争格局：AWS 约占全球云计算市场份额 40%，国内四大云计算厂商占比 79%	36
6. 投资建议	37
7. 风险提示	38



图表目录

图表 1: 通用人工智能有望开启第四次工业革命	6
图表 2: 1860 年起英国煤炭消耗量上行约一个世纪	7
图表 3: 1920 年起美国原油消耗量大幅上行	7
图表 4: AI 三要素: 算法是车间工艺, 算力是能源动力, 数据是生产资料	7
图表 6: 2016 年起 AI 大模型训练所需 FLOPs 出现断点跳跃	8
图表 7: AI 训练与 AI 推理的对比	9
图表 8: 训练与推理芯片对比	9
图表 9: 2022-2027 年人工智能服务器推理和训练工作负载预测	10
图表 10: 训练端算力需求预测	10
图表 11: 推理端算力需求预测	10
图表 12: 算力产业链: 上游-芯片/元器件&中游-服务器/网络设备&下游-数据中心/云服务	11
图表 13: 算力产业图谱: 上游-芯片/元器件&中游-服务器/网络设备&下游-数据中心/云服务	12
图表 14: CPU 基本架构图: 运算器+控制器	13
图表 15: CPU 主流架构对比	14
图表 16: CPU 产品分类: 基于指令集/基于应用领域	15
图表 17: CPU 产业链: 上游连接 EDA/IP/半导体设备/测试, 下游连接整机/设备/应用	15
图表 18: 2022 年全球 CPU 市场规模达 650 亿美元	16
图表 19: 2021 年全球微处理器市场规模高达千亿美元	16
图表 20: 2023 年我国微处理器市场规模预计在 3,000 亿元左右	16
图表 21: 2022 年 Intel 与 AMD 约占全球服务器 CPU 市场份额 90% 以上	17
图表 22: 2021 年 Intel 约占全球 MPU 市场规模半壁江山	17
图表 23: 2020 年日韩系厂商占比全球 MCU 市场主要市场份额	17
图表 24: 2022 年 x86 架构占全球 PC 市场主导地位	18
图表 25: 2023 年全球 x86 CPU 市场份额由 Intel 与 AMD 两巨头占据	18
图表 26: GPU 与 CPU 对比	18
图表 27: GPU 相比 CPU 架构更重视并行计算	19
图表 28: CUDA 支持多种语言 and 应用程序编程接口	20
图表 29: 2022 年全球 GPU 达 422 亿美元, 未来十年有望保持 30% 以上迅猛增长	20
图表 30: 2023 年 NV 约占全球 GPU 市场份额近九成	21
图表 31: 算力由通用走向定制	21
图表 32: 2022 年全球 ASIC 市场规模在千亿体量, 我国约占全球市场 33.4%	22
图表 33: ASIC 市场格局及代表厂商	22
图表 34: FPGA=LC+IO+SB	23
图表 35: 编程 LC 与 SB 改变逻辑电路进而实现不同的功能	23
图表 36: 2023 年全球 FPGA 市场规模有望达 94 亿美元	24
图表 37: 2023 年我国 FPGA 市场规模有望达 250 亿元	24
图表 38: 2022 年 Xilinx 约占全球 FPGA 市场份额 52%	24
图表 39: 2019 年 Xilinx 约占我国 FPGA 市场份额 37%	24



图表 40: 服务器的分类结构.....	25
图表 41: 服务器产业链: 上游连接硬件设备/软件产品, 下游连接应用市场.....	25
图表 42: 2022 年全球服务器市场规模高达千亿美元, 我国约占全球市场规模四分之一.....	26
图表 43: X86 是全球服务器主流架构.....	26
图表 44: 2023 年中国 X86 服务器出货量或达 445 万台.....	26
图表 45: 2023 年戴尔、浪潮、HPE 占据全球服务器市场 R3.....	27
图表 46: 2022 年浪潮、新华三、超聚变引领国内服务器市场.....	27
图表 47: 光模块结构示意图.....	27
图表 48: 光模块分类.....	28
图表 49: 光模块产业链: 上游连接元器件供应商, 下游连接通信设备制造等领域.....	28
图表 50: 2023 年全球光模块市场规模突破百亿美元.....	29
图表 51: 2022 年我国光模块市场规模近 500 亿元.....	29
图表 52: 2021 年全球光模块市场份额.....	29
图表 53: 2022 年全球光模块市占率前 10 中国占据 7 家.....	29
图表 54: 交换机的分类.....	30
图表 55: 交换机产业链: 上游连接电子元器件, 下游连接数据中心等领域.....	30
图表 56: 2022 年全球交换机市场规模达 400 亿美元.....	31
图表 57: 2022 年我国交换机市场规模达 70 亿美元.....	31
图表 58: 2022 年思科是全球交换机市场绝对龙头.....	31
图表 59: 2022 年华为、新华三、锐捷合计占国内市场超过八成.....	31
图表 60: 存储器主要可分为外部存储器和内部存储器.....	32
图表 61: NAND 和 DRAM 对比.....	32
图表 62: 2022 年全球 NAND 市场规模近 700 亿美元.....	33
图表 63: 2021 年全球 DRAM 市场规模近千亿美元.....	33
图表 64: 2022 年三星占全球 NAND 市场三分之一份额.....	33
图表 65: 2021 年三星、SK 海力士、美光垄断全球 DRAM 市场.....	33
图表 66: IDC 产业链: 上游连接基础设施, 下游连接应用行业.....	34
图表 67: 全国数据中心机架需求预计将在 2025 年达到 1400 万架.....	34
图表 68: 2022 年全球数据中心市场规模达 1300 亿美元.....	35
图表 69: 2023 年我国数据中心市场规模在 2,500 亿元左右.....	35
图表 70: 2022 年 Equinix 是全球 IDC 龙头.....	35
图表 71: 2020 年三大运营商占据国内市场超 50% 份额.....	35
图表 72: 云计算产业链: 上游连接芯片/基础设备, 下游连接客户端.....	36
图表 73: 2022 年全球云计算市场规模达 4,000 亿美元.....	36
图表 74: 2022 年我国云计算市场规模达 4,500 亿元.....	36
图表 75: 2022 年 AWS 约占全球云计算市场份额 40%.....	37
图表 76: 2022 年国内四大云计算厂商共占市场份额 79%.....	37
图表 77: 推荐 AI 算力产业相关的计算机赛道 22 家主要标的.....	37



1. 算力——AI 大模型时代的能源动力

1.1 技术革命意味着什么？——新一代生产力工具（AGI）问世推动主体能源（算力）需求快速上行

1.1.1 AGI 曙光：通用人工智能技术有望启动第四次工业革命，算力为主体能源

工业革命的本质是以新一代生产力工具的大范围应用为核心，以主体能源为基础动力的全局性经济范式转型。历史上共发生过三次工业革命，我们认为，本轮通用型人工智能发展浪潮有望启动第四次工业革命：

- 第一次技术革命：生产力工具：蒸汽机；主体能源类型：煤炭。
- 第二次技术革命：生产力工具：内燃机；主体能源类型：石油。
- 第三次技术革命：生产力工具：计算机；主体能源类型：电力。
- 第四次技术革命：生产力工具：通用人工智能；主体能源类型：算力。

图表1：通用人工智能有望开启第四次工业革命

第 X 次工业革命	生产力变革	能源动力变革	关键事件驱动
第一次	蒸汽机	煤炭	1763-1775 年，瓦特在马修·博尔顿的支持下着手改良蒸汽机；1776 年，瓦特设计的蒸汽机进行商业化应用。1829 年，蒸汽动力机车“火箭号”在利物浦到曼彻斯特的铁路上试验成功。全国性铁路网络的建设促进了物流效率的提升与社会分工的加速，进而推动了标准零部件生产的规模经济雏形诞生。
第二次	内燃机	石油	1860 年，比利时工程师艾蒂安·勒努瓦以蒸汽机为蓝本，制成了首台以天然气为燃料的燃气发动机。1862-1876 年，德国工程师尼古拉斯·奥托将内燃机的热效率提升至 10%；1876 年，奥托制成了四冲程循环的内燃机。1885 年，戈特利布·戴姆勒制成了第一台汽油机，并于次年造出第一辆用汽油机驱动的汽车。1908 年，第一辆 T 型车从密歇根州底特律的福特工厂下线，福特汽车开启了大规模流水线装配的生产时代，为规模经济的工厂组织形态奠基。1924 年，意大利建成了世界上第一条高速公路 A8，连接米兰和瓦雷泽，全长约 40 公里，标志着高速公路时代的开始。1956 年，美国通过《联邦援助公路法案》，开工建设彼时世界上最大的公路网络。
第三次	计算机	电力	1943-1945 年，美国的 ENIAC（电子数值积分计算机）完成，这是第一台完全电子的通用计算机，能够重新编程以执行各种任务。1971 年，全球第一款商用微处理器英特尔微处理器 4004 问世。1974 年，罗伯特·卡恩和文顿·瑟夫提出 TCP/IP，定义了电脑网络之间传送报文的方法。1983 年，苹果电脑公司推出全球第一款搭载图形界面的个人电脑 Apple Lisa。1990 年，蒂姆·伯纳斯-李创建了第一个网页浏览器、第一个网页服务器和第一个网站，开启 PC 互联网时代。2007 年，初代 iPhone 面世，开启移动互联网时代。两轮互联网革命将实现了全球范围内的信息交流与远程通讯，降低沟通成本，促使全球人员协作成为可能。
第四次	通用人工智能	算力	1997 年，IBM 深蓝战胜国际象棋世界冠军加里·卡斯帕罗夫。2016 年，AlphaGo 战胜围棋世界冠军李世石。2017 年，《Attention Is All You Need》发表，Transformer 模型问世，后续成为 AI 大模型的主流选择。2022 年，OpenAI 旗下大语言模型产品 ChatGPT 问世。2023 年，OpenAI 旗下多模态大模型 GPT4 与 Google 旗下多模态大模型 Gemini 相继问世。2024 年，OpenAI 旗下 Text to Video/Image 大模型 Sora 问世。大模型时代动辄千亿参数的训练需要巨量的算力资源支撑，以 GPU 为代表的芯片或将成为通用人工智能时代的新能源。AI 与软件结合有望替换文书、设计与管理劳动工作，AI 与硬件结合以机器人形态有望取代体力劳动工作。

来源：Wikipedia, Carlota Perez《技术革命与金融资本——泡沫与黄金时代的动力学》，中国科学院《人工智能的历史、现状和未来》，国金证券研究所

1.1.2 押注能源：历次技术革命的主体能源需求量受益于生产力解放快速上行

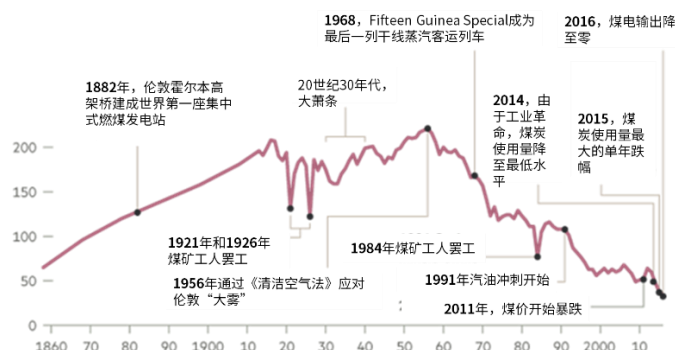
历史上历次技术革命所涉及的主体能源都会随新一代生产力工具广泛应用而呈现能源需求持续上行的阶段，以煤炭和石油两种一次化石能源为例说明：

- 以瓦特蒸汽机的大范围应用为代表的第二次技术革命最早在英国展开，后续逐渐扩散至欧洲大陆和美国。1860 年起，英国煤炭消耗量上行约一个世纪，达峰之后逐渐回落，1860-1920 年英国煤炭消耗量的快速爬升主要受益于蒸汽机在铁路运输、纺织、制造、采矿等各行业的应用（1882 年以后同时受益于煤电厂的兴建）。
- 以内燃机的大范围应用为代表的第三次技术革命最早在美国和德国展开，后续逐渐扩散至欧陆各国。1920 年起，美国原油生产量持续保持高速爬升，1970s，美国国内石油开采量放缓，对海外原油进口飙升，国内石油需求持续增长。



图表2：1860年起英国煤炭消耗量上行约一个世纪

英国历史煤炭使用情况
单位（百万吨）



来源：FT《Britain passes historic milestone with first days of coal-free power》，国金证券研究所

图表3：1920年起美国原油消耗量大幅上行



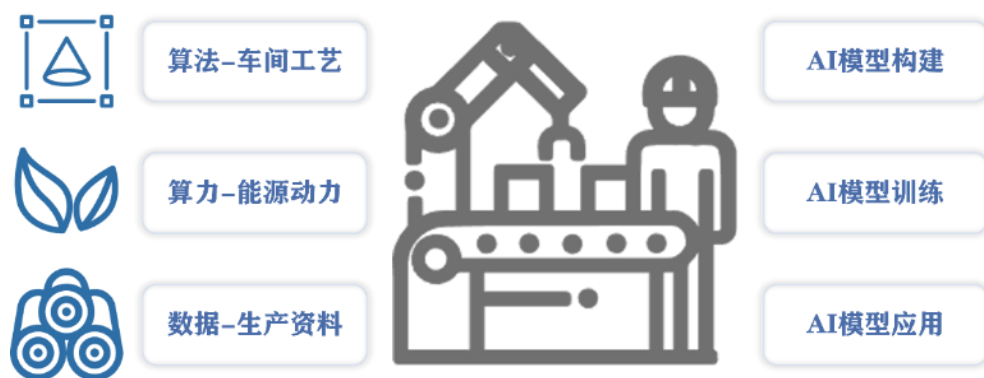
来源：U.S. Energy Information Administration, Annual Energy Review 2009, 国金证券研究所

1.2 AI 算力需求有望在模型规模、算法创新、海量数据、应用落地等多端推动下高增

1.2.1 AI 三要素：算法是车间工艺，算力是能源动力，数据是生产资料

- 算法：算法是 AI 的车间工艺，它定义了如何处理和解释数据。随着深度学习等先进技术的兴起，算法变得越来越复杂，对计算资源的需求也随之增加。
- 算力：算力是 AI 的能源动力，是执行算法所需的硬件资源。它包括 CPU、GPU、TPU 等处理器。算力的提升直接关系到 AI 模型训练的速度和效率，是实现模型快速迭代和优化的关键。
- 数据：数据是 AI 的生产资料，高质量的数据集对于训练出准确可靠的 AI 模型至关重要。数据的收集、处理和分析是 AI 项目成功的关键环节。

图表4：AI 三要素：算法是车间工艺，算力是能源动力，数据是生产资料



来源：中华读书报《智能时代的三要素——数据、算法和算力》，国金证券研究所

1.2.2 算力需求跃迁：AI 模型训练所需 FLOPs 增速超过摩尔定律，引发算力需求高增

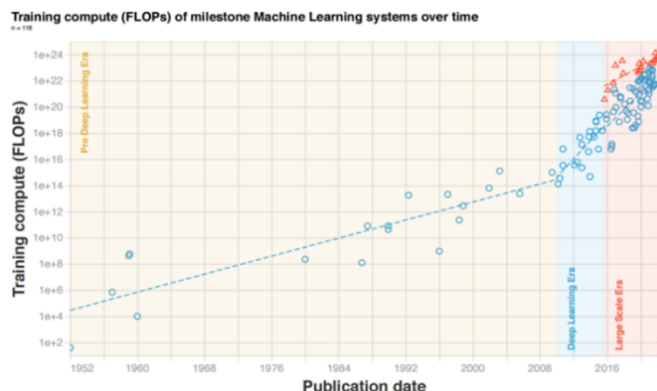
- 2010 年以前（前深度学习时代）：2010 年以前，AI 模型训练所需 FLOPs（Floating Point Operations Per Second，衡量计算性能的指标，特别是在深度学习和神经网络领域中，用来描述硬件设备在单位时间内能够完成的浮点运算次数。是评估深度学习模型的计算复杂度和硬件性能的重要指标）始终保持稳定（指数级）线性增长。这一阶段 AI 模型训练所需 FLOPs 增长与摩尔定律保持一致，大约每 20 个月翻一倍，算力需求从 3×10^4 FLOPs 增长到 2×10^{14} FLOPs。
- 2010~2015 年（深度学习时代）：2010 年起，AI 深度学习模型训练所需 FLOPs 相对前深度学习时代出现斜率突变。自从深度学习在 2010 年代初兴起以来，训练计算的规模加速增长，大约每 6 个月翻倍，远超摩尔定律每 20 个月翻一倍的算力供给。到 2015 年底，随着大型机器学习模型的开发，训练计算的需求增长了 10 到 100 倍，算力需求总规模达到 4×10^{21} FLOPs。



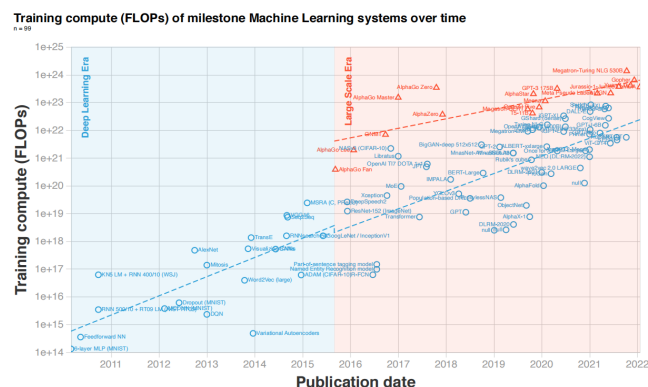
- 2016 年至今（大模型时代）：2016 年起，AI 大模型训练所需 FLOPs 出现断点跳跃。2017 年，《Attention Is All You Need》发表，Transformer 模型问世。Transformer 架构的出现使得模型能够处理更长的序列和更复杂的任务（2017 年之后陆续诞生众多千亿参数模型），但同时也需要更多的计算资源，算法上的创新推动了对更高计算能力的需求。同时，尽管存在如分布式训练和混合精度训练等提高训练效率的技术，但这些技术的应用往往需要更多的硬件资源。为了实现大规模并行处理，需要大量的 GPU 或 TPU 资源，这进一步加剧了对算力的非线性需求，2022 年训练大模型的算力需求达到了 8×10^{23} FLOPs。

图表5：2010 年起 AI 深度学习模型训练所需 FLOPs 相比前深度学习时代出现斜率突变

图表6：2016 年起 AI 大模型训练所需 FLOPs 出现断点跳跃



来源：《Compute Trends Across Three Eras of Machine Learning》(Jaime Sevilla et al., 2022)，国金证券研究所



来源：《Compute Trends Across Three Eras of Machine Learning》(Jaime Sevilla et al., 2022)，国金证券研究所

备注：蓝色虚线代表深度学习模型训练所需算力随时间的变化，红色虚线代表大模型训练所需算力随时间的变化

1.2.3 算力分类：训练算力+推理算力，算力体系逐步完善

训练算力：指在训练机器学习模型时所需的计算资源。训练过程涉及将模型与训练数据进行多次迭代，调整模型参数以最小化损失函数（损失函数是衡量模型预测值与实际值之间差异的函数，反映模型性能）。所需训练算力的大小与模型复杂度、数据集大小和迭代次数有关。

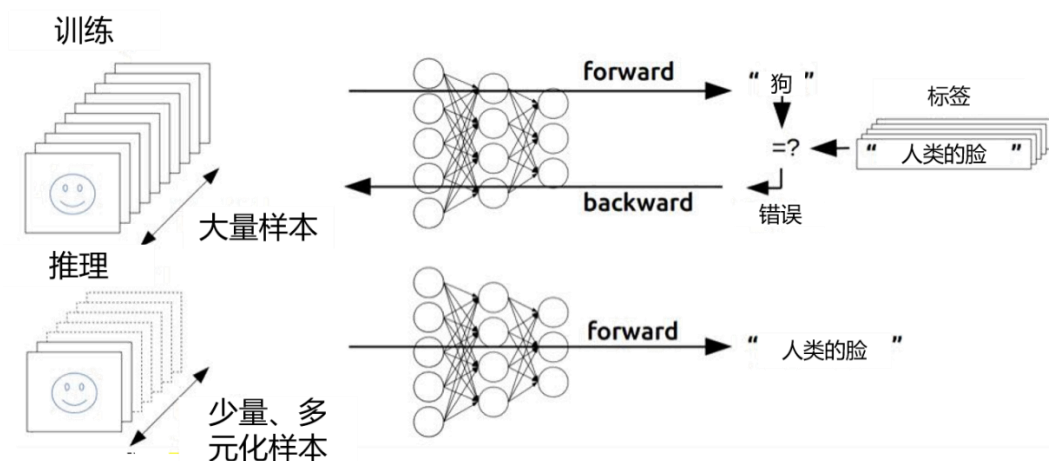
- 越复杂的模型通常需要越多的训练算力，因为它们有更多的参数需要进行调整；
- 使用更大数据集进行训练通常需要更多计算资源，因为需要处理更多的数据量；
- 较大的模型或更复杂的任务可能需要更多的迭代次数才能达到收敛，这会增加训练时间和计算资源的需求。

推理算力：指执行已经训练好的模型进行预测或推理时所需的计算资源。在推理阶段，模型不再更新参数，而是用来对新数据进行预测。这时，算力的需求与模型大小、批处理大小（Batch Size，是指在深度学习中每次输入的数据集大小）、硬件配置有密切联系。

- 越大的模型通常需要越多的推理算力来执行预测，因为它们具有更多的参数需要进行计算；
- 推理时的批处理大小也会影响算力需求，较大的批处理大小可能需要更多的内存和计算资源来处理；
- GPU、TPU 等硬件的类型和数量会影响推理的速度和效率，更强大的硬件通常可以更快地执行推理任务。



图表7: AI 训练与 AI 推理的对比



来源: MichaelAndersch, 国金证券研究所

- 训练芯片: 用于构建神经网络模型, 需要高算力和一定的通用性。
- 推理芯片: 利用神经网络模型进行推理预测, 注重综合指标, 单位能耗算力、时延、成本等。

图表8: 训练与推理芯片对比

场景	算力需求	精度要求	主要产品	应用场景
训练	高, 每个Token 的训练成本通常约为 6N (N 是 LLM 的参数数量)	高, 一般使用 16 位、32 位浮点甚至 64 位双精度数据	英伟达 V100、A100、H100; AMD MI200; 华为昇腾 910; 寒武纪思元 290; 壁仞 BR100 等	对特定的预研模型、神经网络的大量参数进行运算训练, 使模型能够完成各种任务
推理	较低, 每个Token 的推理成本通常约为 2N	较低, 一般使用 8 位或 16 位整型数据	英伟达 L40、T4、A10、A30; AMD Alveo V70; 寒武纪思元 100、思元 370; 燧原邃思 1.0; 阿里平头哥含光 800 等	在已经完成训练的模型基础上, 利用输入数据获得所需要输出的结果

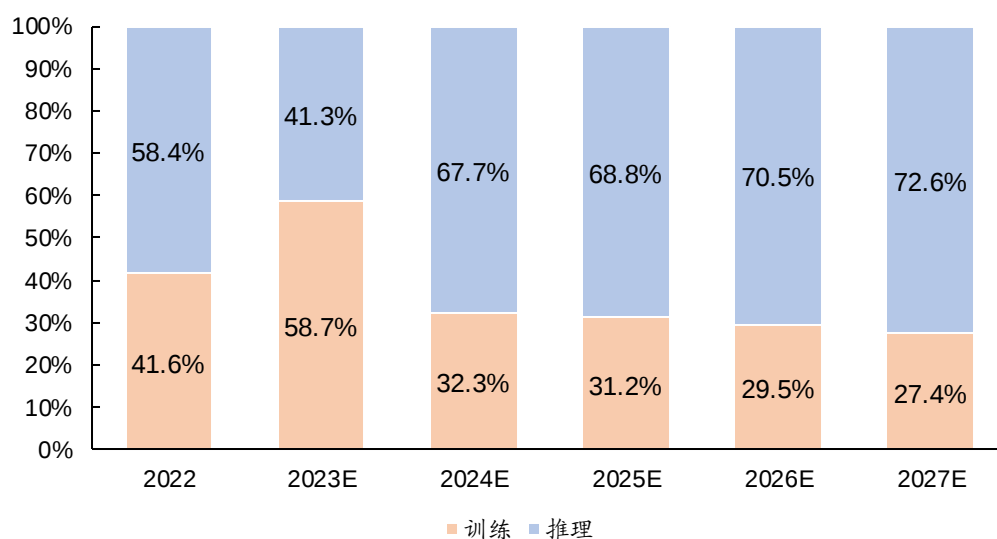
来源: Nvidia 官网, 国金证券研究所

1.2.4 算力规模: 训练/推理对 AI 芯片提出不同要求, 推理端算力需求有望远超训练端

据 IDC, 2023 上半年国内训练工作负载的服务器占比达到 49.4%, 预计全年的占比将达到 58.7%。随着训练模型的完善与成熟, 模型和应用产品逐步进入投产模式, 处理推理工作负载的人工智能服务器占比将随之攀升, 到 2027 年, 国内用于推理的工作负载预计将达到 72.6%。



图表9: 2022-2027 年人工智能服务器推理和训练工作负载预测



来源: IDC《2023-2024 年中国人工智能算力发展评估报告》, 国金证券研究所

AI 大模型在训练端和推理端都将产生巨量的算力需求:

- 训练端算力需求的测算: 从模型的参数规模入手, 根据训练大模型所需的 Token 数量和每 Token 训练成本与模型参数数量的关系分别估算总算力需求, 再考虑单张 GPU 算力和 GPU 集群的算力利用率推导出 GPU 总需求。其中, 参考 OpenAI 发布的论文《Scaling Laws for Neural Language Models》中的分析, 每个 Token 的训练成本通常约为 $6N$, 其中 N 是 LLM 的参数数量。以 GPT-3 为例, 根据《Training Compute-Optimal Language Models》, GPT3 的参数数量为 1,750 亿, 训练 Tokens 为 3,000 亿, 该模型训练算力总需求为 $6 \times 1,750 \text{ 亿} \times 3,000 \text{ 亿} = 3.15 \times 10^{23}$ FLOPs, 对应图表 10 公式 (1)。假如目前要同时 3 个 GPT-3 大模型, 且需要在一周以内训练完成, 则训练端峰值算力需求为 $3.15 \times 10^{23} \text{ FLOPs} \times 2 / (7 \times 24 \times 3600) = 1.04 \times 10^{18}$, 对应图表 10 公式 (2)。

图表10: 训练端算力需求预测

单个大模型所需算力 = $6N \times$ 训练 Tokens 数量 (1)

训练端峰值算力需求 = 单模型所需算力 $\times$ 模型数量 $\div$ 单次训练秒数 (2)

来源: OpenAI《Scaling Laws for Neural Language Models》(Jared Kaplan et al., 2020), 国金证券研究所

- 推理端算力需求的测算: 根据大模型日活用户人数、每人平均查询 Token 数量、每 Token 推理成本与模型参数数量的关系估算推理端总算力需求, 同时考虑最大并发峰值的算力乘数为高并发提供冗余。其中, 参考 OpenAI 发布的论文《Scaling Laws for Neural Language Models》中的分析, 每个 Token 的推理成本通常约为 $2N$, 其中 N 是 LLM 的参数数量。以 ChatGPT 为例, 据 SimilarWeb, 2024 年 2 月 ChatGPT 总访问次数约为 16.3 亿次, 平均每天访问次数为 0.56 亿次, 假设每次访问查询 token 数为 1000 个 (对应 750 个单词), 假设目前 ChatGPT 的参数规模为 2,000 亿, 则全天计算次数合计为 $2 \times 2000 \text{ 亿} \times 1,000 \times 0.56 \text{ 亿} = 2.24 \times 10^{22}$ 次, 对应图表 11 公式 (1)。平均每秒所需峰值算力为 $2.24 \times 10^{22} / (24 \times 60 \times 60) = 2.60 \times 10^{17}$ FLOPs, 对应图表 11 公式 (2)。假设最大并发峰值算力乘数取 5, 则最大并发峰值算力为 $2.60 \times 10^{17} \text{ FLOPs} \times 5 = 1.3 \times 10^{18} \text{ FLOPs}$, 对应图表 11 公式 (3)。

图表11: 推理端算力需求预测

全天计算次数合计 = $2N \times$ Tokens 数量 $\times$ 查询次数 $\times$ 日活人数 (1)

平均每秒所需峰值算力 = 全天计算次数合计 $\div 24 \div 60 \div 60$ (2)

最大并发峰值算力 = 平均每秒所需峰值算力 $\times$ 最大并发峰值算力乘数 (3)

来源: OpenAI《Scaling Laws for Neural Language Models》(Jared Kaplan et al., 2020), 国金证券研究所



2. 透视算力产业链——上游：芯片/元器件→中游：服务器/网络设施→下游：IDC/云服务

算力产业包括上游的芯片及元器件，中游的服务器与网络设备，下游的数据中心及云服务三部分构成，在其中的各个板块，均已有国产厂商深度布局。

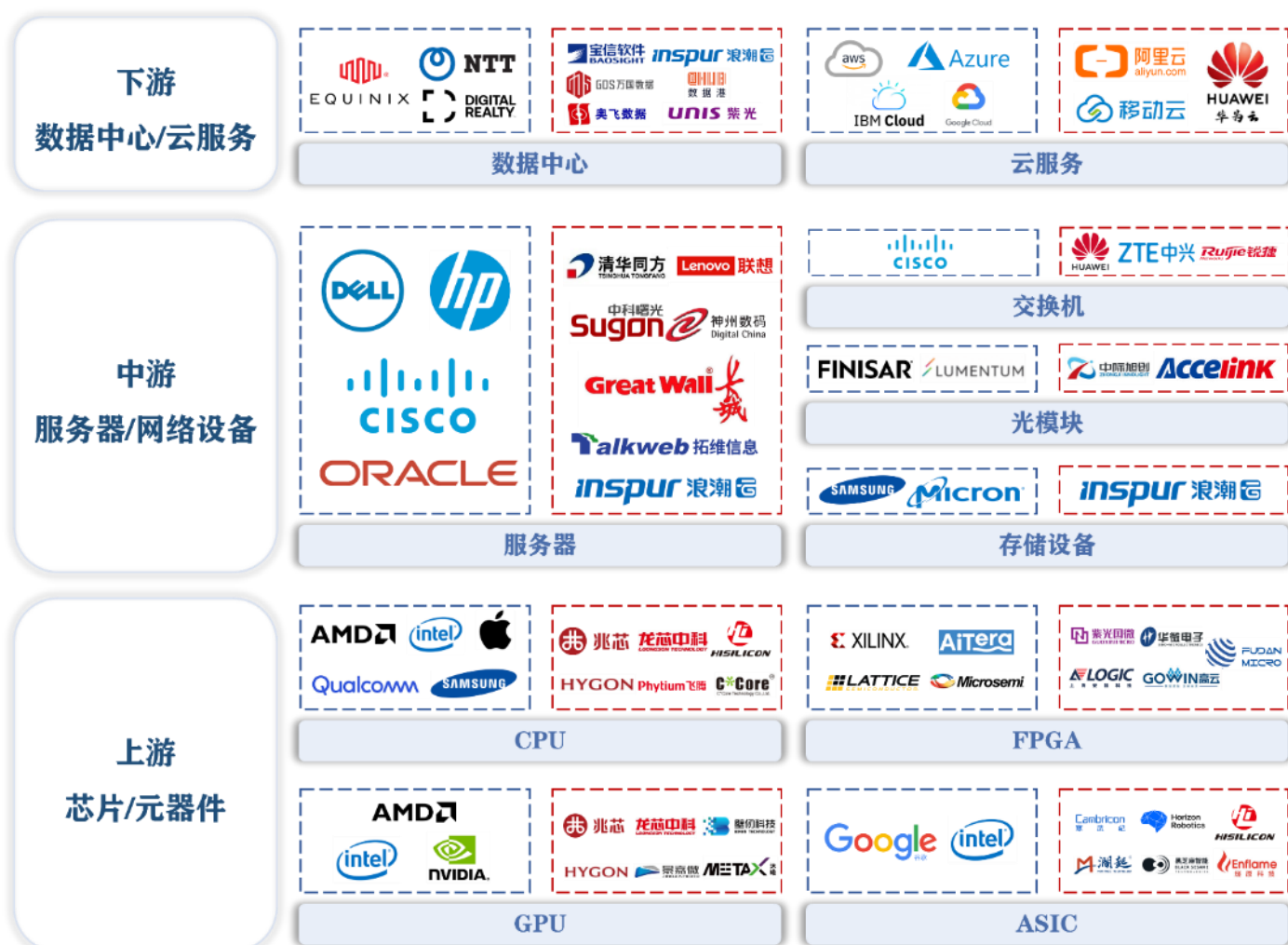
图表12：算力产业链：上游芯片/元器件&中游服务器/网络设备&下游数据中心/云服务

产业链模块	具体产品	国内外代表厂商
上游：芯片及元器件	CPU	国外代表厂商：超微半导体、英特尔、苹果、高通、三星 国内代表厂商：上海兆芯、龙芯中科、海思（华为）、平头哥（阿里）、海光信息、飞腾公司、国芯科技
	GPU	国外代表厂商：超微半导体、英特尔、英伟达 国内代表厂商：上海兆芯、龙芯中科、壁仞科技、海光信息、景嘉微、沐曦
	FPGA	国外代表厂商：赛灵思、阿尔特拉、莱迪思、美高森美 国内代表厂商：紫光股份、华微电子、安路科技、高云、复旦微电、昆仑芯（百度）。
	ASIC	国外代表厂商：谷歌、英特尔 国内代表厂商：寒武纪、地平线、海思（华为）、澜起科技、黑芝麻智能、燧原科技
中游：服务器与网络设备	服务器	国外代表厂商：戴尔、惠普、思科、甲骨文 国内代表厂商：同方股份、联想、中科曙光、神州数码、中国长城、拓维信息、浪潮信息
	交换机	国外代表厂商：思科 国内代表厂商：华为、中兴、锐捷
	光模块	国外代表厂商：菲尼萨、鲁门特姆 国内代表厂商：中际旭创、光迅科技
	存储设备	国外代表厂商：三星、美光科技 国内代表厂商：浪潮信息
下游：数据中心及云服务	数据中心	国外代表厂商：易昆尼克斯、日本电信、数字地产 国内代表厂商：宝信软件、万国数据、奥飞数据、浪潮信息、数据港、紫光股份
	云服务	国外代表厂商：亚马逊云、微软云、IBM云、谷歌云 国内代表厂商：阿里云、华为云、运营商云、润泽科技、润健股份、科华数据。

来源：中国信通院《数据中心产业图谱研究报告(2022)》，国金证券研究所



图表13：算力产业图谱：上游-芯片/元器件&中游-服务器/网络设备&下游-数据中心/云服务



来源：中国信通院《数据中心产业图谱研究报告(2022)》，国金证券研究所（注：图中蓝色虚线方框内为境外公司，红色虚线方框内为中国公司）

3. 上游-芯片/元器件——异构算力蓬勃发展，国产替代空间广阔

3.1 CPU：中央处理器，Intel 与 AMD 两大巨头占据 x86 CPU 市场

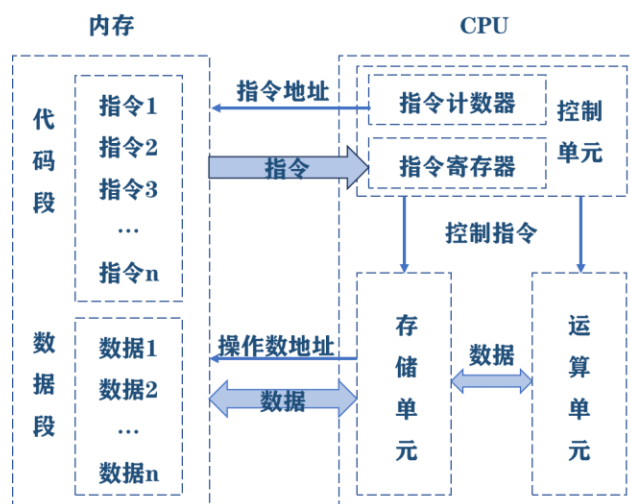
3.1.1 基本定义：CPU 负责解释计算机指令以及处理计算机软件中的数据

CPU 是中央处理器（Central Processing Unit）的简称，采用超大规模的集成电路组成制造，是计算机的运算核心和控制核心。CPU 包括运算器（算术逻辑单元、累加寄存器、数据缓冲寄存器、状态条件寄存器）、控制器（指令寄存器、程序计数器、地址寄存器、指令译码器、时序、总线、中断逻辑控制）、高速缓冲存储器、内部数据总线、控制总线及状态总线输入/输出接口等模块。

CPU 的主要功能是解释计算机指令以及处理计算机软件中的数据。指令寄存器（IR）从存储器或高速缓冲存储器中取出指令，指令译码器（ID）对指令进行译码，将指令分解成一系列的微操作，向运算器中能够进行所需运算的模块发送控制命令，从而完成指令的执行。



图表14: CPU基本架构图: 运算器+控制器



来源: 中国软件评测中心·集成电路测评工程技术中心《CPU技术与产业白皮书(2020年)》, 国金证券研究所

3.1.2 CPU 分类: 基于指令集的 CISC 与 RISC, 基于应用领域的 MPU/MCU/DSP

指令集是 CPU 所执行指令的二进制编码方法, 是软件和硬件的接口规范。CPU 按照指令集可分为 CISC (Complex Instruction Set Computer, 复杂指令集) 和 RISC (Reduced Instruction Set Computer, 精简指令集) 两大类, CISC 型 CPU 目前主要是 X86 架构, RISC 型 CPU 主要包括 ARM、RISC-V、MIPS、POWER 架构等:

- X86: 基于 CISC (复杂指令集) 的 X86 架构是一种为了便于编程和提高存储器访问效率的芯片设计体系, 包括两大主要特点: 一是使用微代码, 指令集可以直接在微代码存储器里执行, 新设计的处理器, 只需增加较少的晶体管电路就可以执行同样的指令集, 也可以很快地编写新的指令集程式; 二是拥有庞大的指令集, X86 拥有包括双运算元格式、寄存器到寄存器、寄存器到存储器以及存储器到寄存器的多种指令类型。
- ARM 架构: ARM 架构过去称作进阶精简指令集机器, 是一个 32 位精简指令集处理器架构, 其广泛地使用在许多嵌入式系统设计, 近年来也因其低功耗多核等特点广泛应用在数据中心服务器市场。早期 ARM 指令集架构的主要特点: 一是体积小、低功耗、低成本、高性能; 二是大量使用寄存器, 且大多数数据操作都在寄存器中完成, 指令执行速度更快; 三是寻址方式灵活简单, 执行效率高; 四是指令长度固定, 可通过多流水线方式提高处理效率。
- RISC-V 架构: RISC-V 是加州大学伯克利分校设计并发布的一种开源指令集架构, 其目标是成为指令集架构领域的 Linux, 主要应用于物联网 (IoT) 领域, 但可扩展至高性能计算领域。RISC-V 采用 BSD License 发布, 由于允许衍生设计和开发闭源, 吸引了一大批公司的关注, 目前已有不少公司开发基于 RISC-V 的 IP 核, 如 Si-Five、台湾晶心、阿里平头哥等已可提供基于 RISC-V 的处理器 IP 核, 部分企业如兆易创新、北京君正等已开发出基于 RISC-V 的 MCU 芯片等。但整体上, 由于 RISC-V 产业生态还比较薄弱, 未来的发展仍有较长一段路要走。
- MIPS 架构: MIPS 是高效精简指令集计算机体系结构中的一种, MIPS 的优势主要有三点: 一是发展历史早, MIPS 在 1990 年代已经广泛使用在服务器、工作站设备上。二是在学术界影响广泛, 计算机体系结构教材都是以 MIPS 为实际例子。三是 MIPS 在架构授权方面更为开放, 授权门槛远低于 X86、ARM, 在 2019 年曾经有开放授权的实际行动, 并且 MIPS 允许授权商自行更改设计、扩展指令, 允许二次授权。
- POWER 架构: POWER 架构是由 IBM 设计的一种 RISC 处理器架构, POWER 在大型机领域独具优势。POWER3 是全球首款 64 位架构处理器, 开始应用铜互联和 SOI (绝缘体上硅) 技术。直至 POWER 9 依然追求最高性能, 不仅具备乱序执行、智能线程等技术, 还实现了 SMP (对称多处理技术) 的硬件一致性处理。


图表15: CPU 主流架构对比

指令集架构	X86	ARM	RISC-V	MIPS	POWER
指令集	CISC	RISC	RISC	RISC	RISC
发布时间	1978 年	1985 年	2010 年	1981 年	1990 年
设计方	英特尔	ARM 计算机公司	加州大学伯克利分校	美普思科技	IBM
核心应用领域	主要应用于 Intel, AMD 等 PC 机, X86 服务器中	主要应用于移动终端之中, 类如手机, 平板等	主要应用于物联网领域, 可扩展至高性能计算领域	主要应用于数字消费、宽带、无线、网络和便携式媒体领域	大型机领域
市占率	在 PC 市场独霸多年	在移动领域占主流	起步相对较晚, 但发展很快, 在智能穿戴产品上应用广泛	MIPS 两度易主后, 新公司已转向 RISC-V, MIPS 基本退出国际主流应用	/
开/闭源	闭源	闭源	开源	闭源	闭源
显著特征	面向家用, 商用领域, 性能和兼容性上做得更好	面向移动, 低功耗领域, 体积小、功耗低、低成本、高性能	低功耗、低成本、灵活可扩展、安全可靠	追求高性能	追求最高性能

来源: Wikipedia, 国金证券研究所

CPU 按照下游应用领域可分为通用微处理器 (MPU, Micro Processor Unit)、微控制器 (MCU, Micro Controller Unit) 和专用处理器:

- 通用微处理器: MPU 通常按照面向的市场分为用于服务器、桌面 (台式机/笔记本)、超级计算机等。
- 微控制器: MCU 微控制单元, 是用于控制类应用的低性能、低功耗 CPU。MCU 的主频一般低于 100MHz, 是把 CPU 的频率与规格做适当缩减, 并将内存、计数器、USB、A/D 转换、UART、PLC、DMA 等周边接口, 甚至 LCD 驱动电路都整合在单一芯片上, 形成的芯片级计算机。MCU 的主频、功耗都很低, 为不同应用场合做不同组合控制。诸如在智能制造、工业控制、智能家居、遥控器消费领域, 以及汽车电子、工业上的步进马达、机器手臂的控制等, 都大量使用 MCU。
- 专用处理器: 专用处理器实现面向某一领域的特定功能。例如数字信号处理器 (DSP, Digital Signal Processor), DSP 芯片也称数字信号处理器, 是一种专用于进行数字信号处理运算的微处理器, 其主要应用是实时快速地实现各种数字信号处理算法。DSP 广泛应用于数字控制、运动控制等方面。其它专用处理器还有深度学习处理器、数据库加速处理器、安全处理器、类脑计算芯片等。



图表16: CPU产品分类: 基于指令集/基于应用领域

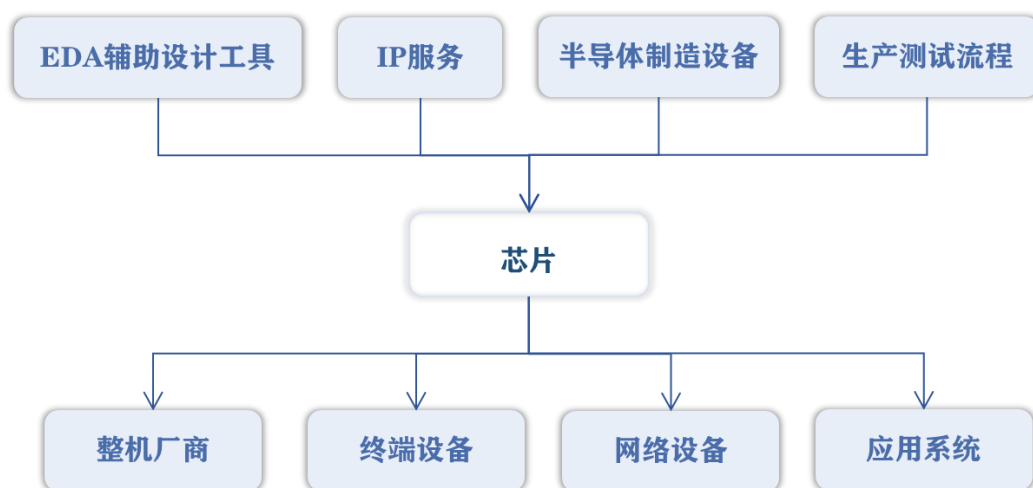


来源: 中国软件评测中心·集成电路测评工程技术中心《CPU技术与产业白皮书(2020年)》, 国金证券研究所

3.1.3 产业链条: 上游-EDA/IP/半导体设备/测试, 下游-整机厂商/行业解决方案/应用系统

CPU的产业链上游包括支撑集成电路设计和制造的EDA辅助设计工具和IP服务, 半导体制造设备、芯片生产测试流程。目前CPU的产业链上游企业多为国外知名厂商, 具有垄断优势, 国产化程度较低。CPU的产业链下游包括各类整机厂商、终端设备、网络设备和应用系统等, 其中最重要的是服务器、桌面和嵌入式系统等硬件设备厂商。

图表17: CPU产业链: 上游连接EDA/IP/半导体设备/测试, 下游连接整机/设备/应用



来源: 中国软件评测中心·集成电路测评工程技术中心《CPU技术与产业白皮书(2020年)》, 国金证券研究所

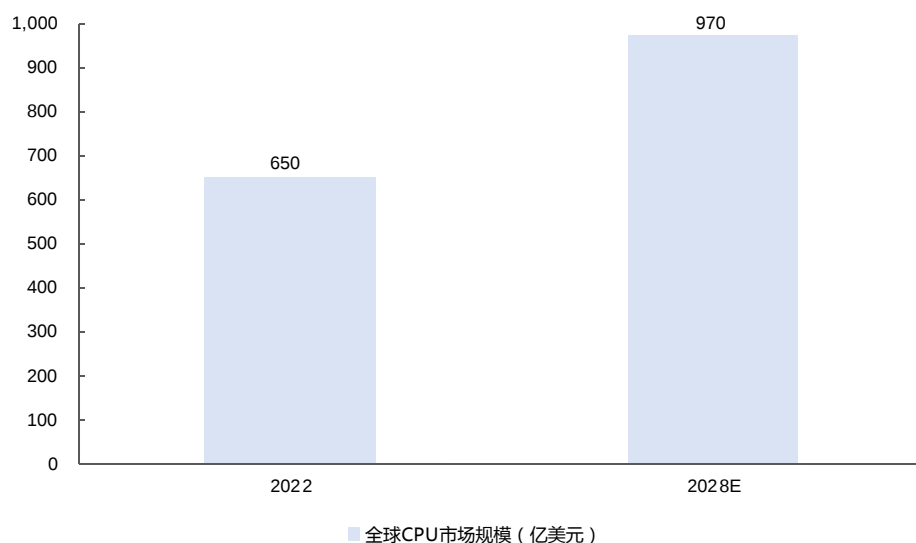
备注: 应用系统指的是一种能够满足特定业务需求和用户需求的软件系统, 它通过与硬件设备交互, 完成预先定义的任务, 比如智能家居应用系统、智慧农业应用系统。

3.1.4 市场规模: 全球微处理器市场规模高达千亿美元, 我国约占全球市场规模40%

根Yole数据, 2022年全球CPU市场规模为650亿美元, 预计2028年增长到970亿美元, 年复合增长率为7%。



图表18: 2022 年全球 CPU 市场规模达 650 亿美元

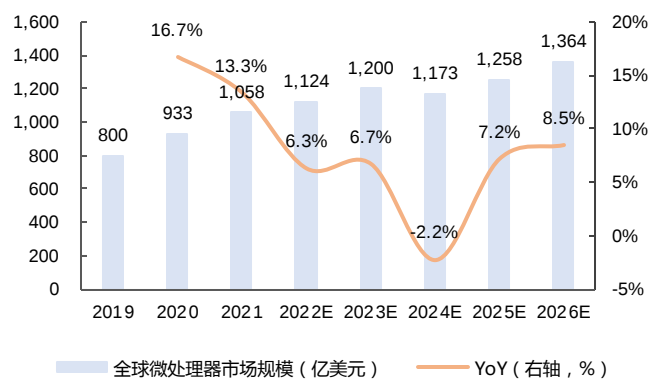


来源: Yole, 国金证券研究所

据 IC Insights, 2021 年, 得益于公共卫生事件期间的封控驱动了市场对 PC 和智能手机的需求, 全球微处理器总销售额保持 13.3% 的强劲增长, 高达 1,058 亿美元。2022 年全球 MPU (微处理器) 总销售额有望增长 6.3%。

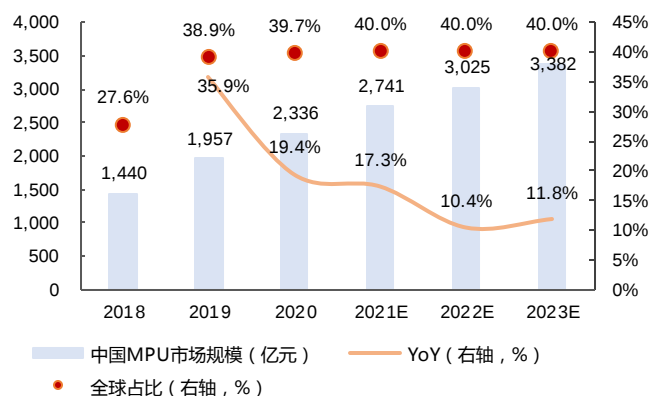
据 IC Insights, 2020 年, 我国 MPU 市场规模达到 2,336 亿元, 约占全球规模的 36.3%。假设我国占全球市场比重保持不变, 2023 年我国 MPU 市场规模有望达 3,382 亿元。

图表19: 2021 年全球微处理器市场规模高达千亿美元



来源: IC Insights, 国金证券研究所

图表20: 2023 年我国微处理器市场规模预计在 3,000 亿元左右



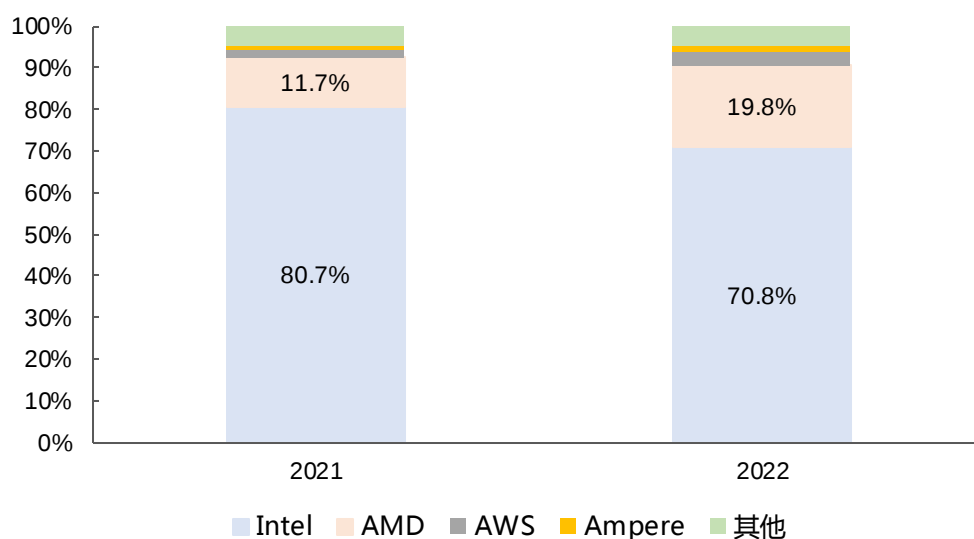
来源: IC Insights, 国金证券研究所

3.1.5 竞争格局: Intel 约占全球 MPU 市场半壁江山, Intel 与 AMD 独占 x86 全球市场

据 Counterpoint Analysis, 2022 年, Intel 约占全球服务器 CPU 市场份额 70.8%, AMD 约占全球服务器 CPU 市场份额 19.8%, 二者合计约占 90.6%; 2021 年, 二者合计约占 92.5%。



图表21: 2022 年 Intel 与 AMD 约占全球服务器 CPU 市场份额 90% 以上

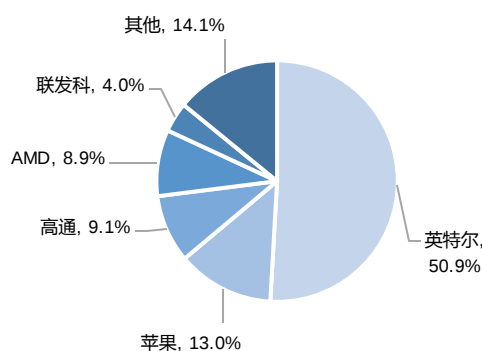


来源: CounterpointAnalysis, 国金证券研究所

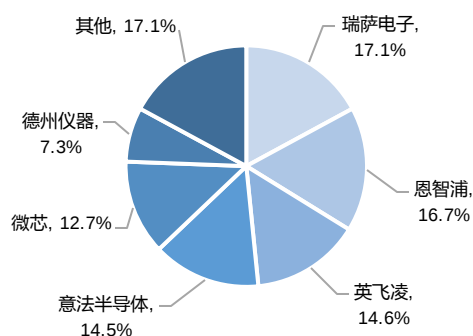
据艾瑞咨询, 2021 年, 全球 MPU (微处理器) 市场多被英特尔、苹果和高通等美系厂商占据。中国厂商, 如紫光展锐与华为海思已逐步走进国际视野, 在 21 年 IC Insight 竞争格局中分别位列第八、九位, 未来份额有待进一步提升。

据艾瑞咨询, 2020 年, 欧美及日韩系厂商在全球 MCU (微控制器) 市场占据绝对优势, 尤其在汽车/车规级与工控领域的中高端产品线。中颖电子、兆易创新等中国厂商多由消费电子切入, 正发力转型突破以进入中高端市场。

图表22: 2021 年 Intel 约占全球 MPU 市场规模半壁江山



图表23: 2020 年日韩系厂商占比全球 MCU 市场主要市场份额



来源: 艾瑞咨询《中国半导体 IC 产业研究报告(2022 年)》, 国金证券研究所

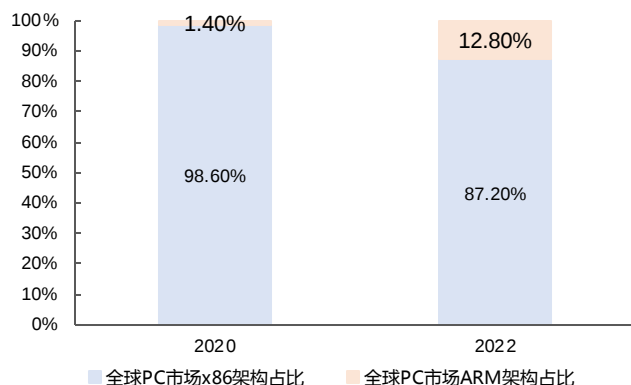
来源: 艾瑞咨询《中国半导体 IC 产业研究报告(2022 年)》, 国金证券研究所

根据指令集架构区分, 全球 PC 市场中 x86 架构占绝对主导地位, 市占率在 90% 左右。但得益于 ARM 架构的低成本、低功耗、系统通用性、日益加强的性能以及海量终端设备生态, 未来有巨大的市场增长潜力。在过去 2 年, Arm 架构在全球 PC 市场的份额从 2020 年的 1.4% 增长至 2022 年的 12.8%, 蚕食 x86 PC 的部分市场。

据 Mercury Research, 全球 x86 CPU 市场份额由 Intel 与 AMD 两巨头占据。2023Q2, Intel 约占全球 x86 CPU 市场份额 68.4%, AMD 约占 31.6%。

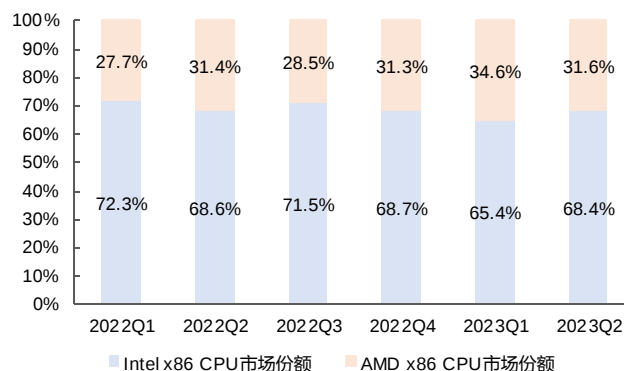


图表24: 2022 年 x86 架构占全球 PC 市场主导地位



来源: 亿欧智库, 国金证券研究所

图表25: 2023 年全球 x86 CPU 市场份额由 Intel 与 AMD 两巨头占据



来源: Mercury Research, 国金证券研究所

3.2 GPU: CUDA 大幅降低 GPU 并行计算的编程难度, NV 约占全球市场九成份额

3.2.1 基础定义: CPU 侧重低延迟计算, GPU 更侧重并行计算

GPU 一般指图形处理器 (Graphics Processing Unit, 缩写 GPU), 又称显示核心、视觉处理器、显示芯片, 是一种专门在个人电脑、工作站、游戏机和一些移动设备 (如平板电脑、智能手机等) 上做图像和图形相关运算工作的微处理器。

GPU 与 CPU 设计处理的计算任务的目标是不一样的, CPU 是一种低延迟的设计, GPU 更侧重并行计算。通俗地讲, CPU 可单独处理复杂的逻辑运算和不同的数据类型, 如同一位教授单独去解一道复杂的高数题, 而 GPU 是用很多简单的计算单元去完成大量的计算任务, 如同一百名小学生分别独立计算 10 以内的加减法。因此, CPU 和 GPU 整体结构有很大的区别。

图表26: GPU 与 CPU 对比

比对角度	CPU	GPU
计算能力	具有强大且较少的 ALU (Arithmetic Logic Unit, 算术逻辑单元), 时钟频率很高	有大量的 ALU, 可以支持并发的计算线程
存储容量	具有容量较大的 Cache (高速缓存), 一般包括 L1、L2 和 L3 三级高速缓存; L3 可以达到 8MB	Cache 很小, 缓存的目的不是保存后面需要访问的数据, 而是为线程提高服务的
控制逻辑	有复杂的控制逻辑, 比如复杂的流水线 (Pipeline)、分支预测 (Branch prediction)、乱序执行 (Out-of-order execution) 等	没有复杂的控制逻辑, 没有分支预测等组件
应用领域	各种计算任务, 例如操作系统、应用程序和各种算法	图形和图像处理任务, 例如游戏、视频编辑和计算机视觉; 非图像显示并涉及大量并行运算的领域, 例如 AI 训练、加密解密、科学计算

来源: 英特尔官网, 腾讯云, 国金证券研究所



图表27: GPU 相比 CPU 架构更重视并行计算



来源: Github, NVIDIA, 国金证券研究所

3.2.2 CUDA: 大幅降低 GPU 并行计算的编程难度, 实现 GPU 的通用化

CUDA (Compute Unified Device Architecture, 统一计算设备架构), 由 NVIDIA 于 2007 年开发, CUDA 可以让开发者能够用类似 C 语言的方式编写程序, 让 GPU 来处理计算密集型任务。因此, 从简单的角度, 可以理解为这是一套英伟达提供给开发人员的编程工具, 运用 CUDA 能省下大量撰写低阶语法的时间, 进而直接使用高阶语法诸如 C++或 Java 等来编写应用于通用 GPU 上的演算法, 解决平行运算中复杂的问题。

CUDA 平台包含了一系列工具函数, 有各种功能, 如:

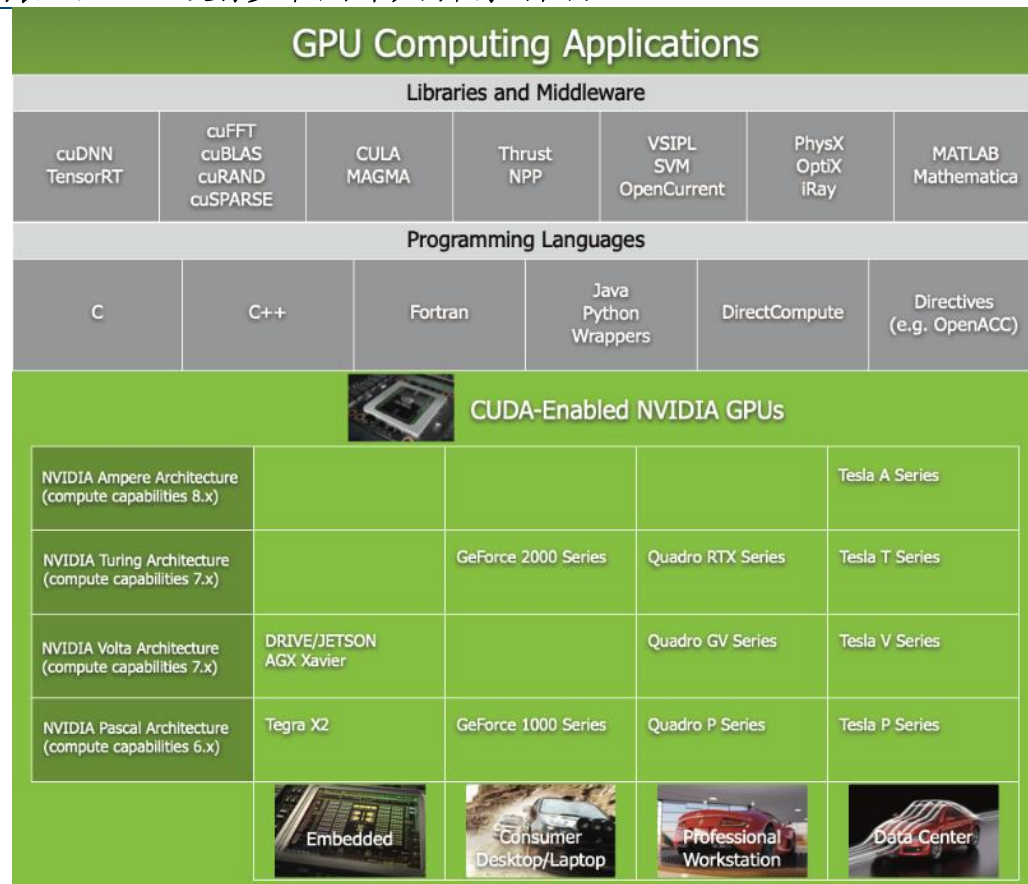
- 数据索引: 从主机端(CPU)存储器获取数据, 到相应的设备端(GPU)存储器
- 内核函数: GPU 上的计算函数代码
- 线程分配: 把计算任务分配到各个线程 (对应硬件单元里的 SP)
- 定义数据存储的缓存层次: 哪些数据在离计算单元较近的缓存, 哪些在比较远的全局存储器
- 设置流与事件: 划定某些计算节点, 以实现复杂的任务图

CUDA 具备三点核心优势:

- 高效的并行计算能力: CUDA 能够让成千上万的 GPU 核心同时工作, 大幅提高计算速度。
- 广泛的应用场景: CUDA 的应用几乎遍及所有需要大量计算的领域。如: 科学研究 (在物理、化学、生物等领域, CUDA 能够加速复杂的模拟和计算过程。比如, 模拟宇宙的起源、分析大量的基因序列等。)、深度学习 (深度学习需要处理大量的数据和复杂的计算。使用 CUDA, 训练神经网络的时间可以从几周缩短到几小时。)、图像处理 (从电影的特效制作到医学图像的分析, CUDA 能够加速图像处理的过程, 让复杂的图像分析变得更加快速和准确。)、金融分析 (在金融领域, CUDA 被用来加速风险分析、市场模拟等计算密集型任务, 帮助分析师更快地做出决策。)等。
- 强大的生态支持: NVIDIA 提供了丰富的文档、教程和工具, 让开发者更容易地开发基于 CUDA 的应用。



图表28: CUDA 支持多种语言 and 应用程序编程接口



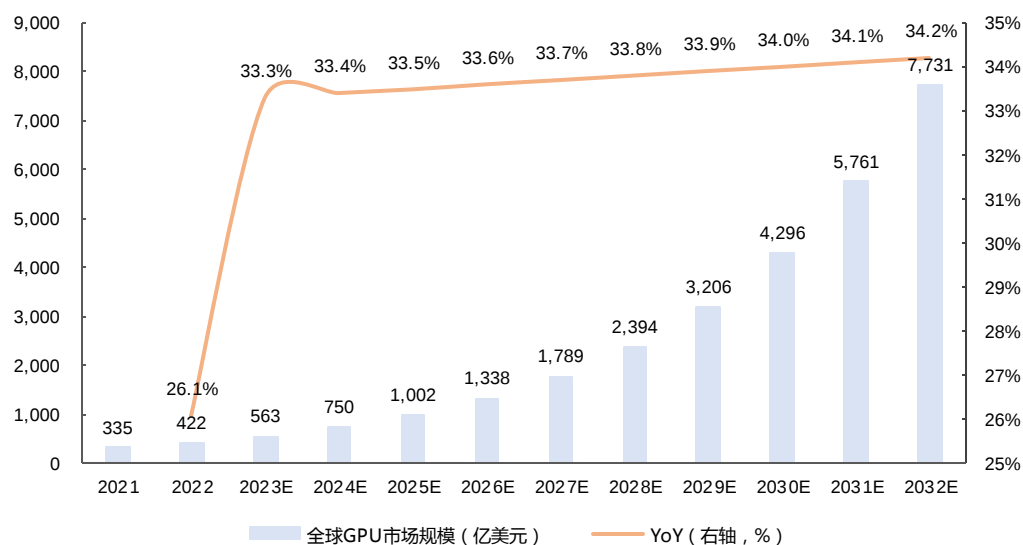
来源: NVIDIA, 国金证券研究所

3.2.3 市场规模: 2022 年全球 GPU 市场规模达 422 亿美元, 未来十年有望高速增长

据 Verified Market Research, 2021 年, 全球 GPU 市场规模高达 334.7 亿美元。

据 Precedence Research, 2022 年, 全球 GPU 市场规模未来 10 年 CAGR 有望保持 30% 以上, 2032 年全球 GPU 市场规模有望达到 7,731 亿美元。

图表29: 2022 年全球 GPU 达 422 亿美元, 未来十年有望保持 30% 以上迅猛增长



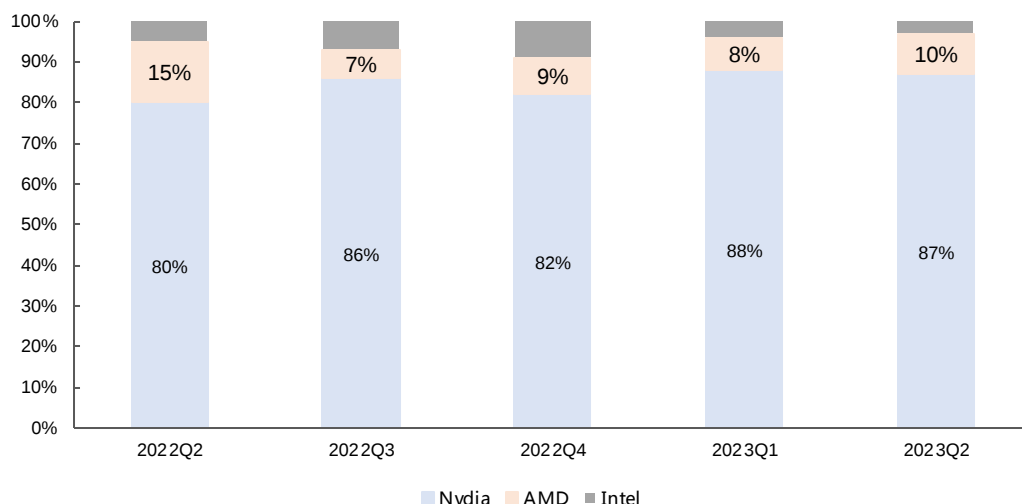
来源: Verified Market Research, Precedence Research, 国金证券研究所

3.2.4 竞争格局: NV 约占全球 GPU 市场份额近九成, AMD 与 Intel 约占 10%

据 JCR, 2022Q2 开始, NV 约占全球 GPU 市场份额 80% 以上。2023Q2, NV 约占全球 GPU 市场份额高达 87%, AMD 约占 10%, Intel 约占 3%。



图表30: 2023 年 NV 约占全球 GPU 市场份额近九成



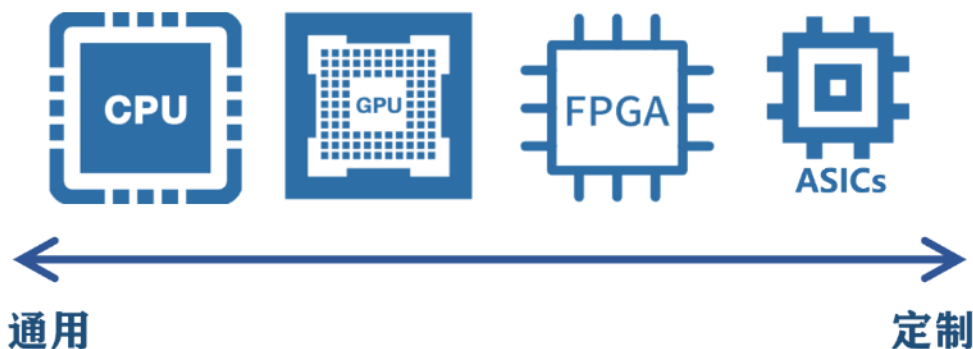
来源: JCR, 国金证券研究所

3.3 ASIC: 算力由通用走向定制, 2022 年我国 ASIC 市场规模达 335 亿元

3.3.1 基础定义: 专用集成电路, 算力从通用走向定制

ASIC (Application Specific Integrated Circuit, 专用集成电路) 是一种专用芯片, 与传统的通用芯片有一定的差异, 是为了某种特定的需求而专门定制的芯片。ASIC 芯片的计算能力和计算效率都可以根据算法需要进行定制, 所以 ASIC 与通用芯片相比, 具有以下几个方面的优越性: 体积小、功耗低、计算性能高、计算效率高、芯片出货量越大成本越低。但是缺点也很明显, 只能针对特定的某个或某几个应用场景, 一旦算法和流程变更更可能导致 ASIC 无法满足业务需求。

图表31: 算力由通用走向定制



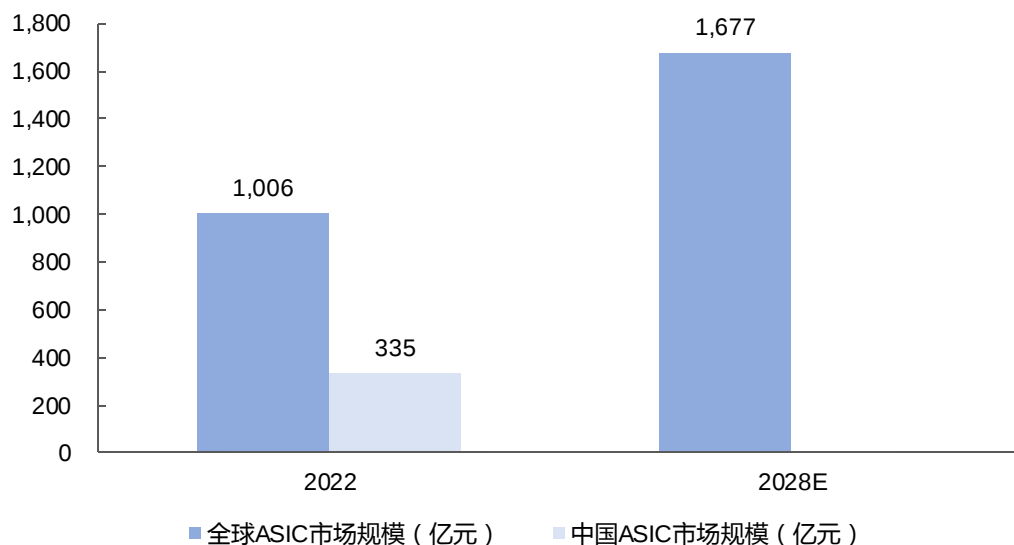
来源: 中国联通算力网络产业技术联盟《异构算力统一标识与服务白皮书(2021 年)》, 国金证券研究所

3.3.2 市场规模: 2022 年全球 ASIC 市场规模在千亿体量, 我国约占全球市场 33.4%

据 Market Monitor, 2022 年, 全球 ASIC 市场规模在千亿体量, 我国市场规模约在 335 亿元, 约占全球市场的 33.4%。预计 2028 年全球市场有望达 1,677 亿元, 22~28 年 CAGR 或达 8.9%。



图表32: 2022 年全球 ASIC 市场规模在千亿体量, 我国约占全球市场 33.4%



来源: Market Monitor, 国金证券研究所

3.3.3 竞争格局: 竞争格局开放, 国产替代正当时

ASIC 不同于 CPU、GPU、FPGA, 目前全球 ASIC 市场并未形成明显的头部厂商。由于 ASIC 开发周期长, 仅有大厂有资金与实力进行研发。同时, ASIC 是全定制芯片, 在某些特定场景下运行效率最高, 故某些场景下游市场空间足够大时, 量产 ASIC 芯片可以实现丰厚的利润。目前市场上主流 ASIC 有 TPU 芯片、NPU 芯片、VPU 芯片以及 BPU 芯片, 它们分别是由 Google、寒武纪、Intel 以及地平线公司设计生产。在海外, 谷歌 TPU 是主导者, 国内初创芯片企业 (如寒武纪、比特大陆和地平线), 互联网巨头 (如百度、华为海思和阿里) 在细分领域也有所建树, 目前已经取得了一定的成果。

图表33: ASIC 市场格局及代表厂商

名称	提供商	简介
ASIC	TPU	Google
	NPU	寒武纪
	VPU	Intel
	BPU	地平线
	BM1680	比特大陆

来源: 英特尔官网, 智能计算芯世界, 腾讯网, 国金证券研究所

3.4 FPGA: 现场可编程芯片, Xilinx 约占全球市场份额 52%

3.4.1 基础定义: 现场可编程的芯片, FPGA=LC+IO+SB

FPGA (Field-Programmable Gate Array, 现场可编程门阵列) 芯片的最大特点是现场可编程性。无论是 CPU、GPU、DSP、Memory 还是各类 ASIC 芯片, 在芯片被制造完成之后, 其芯片的功能就已被固定, 用户无法对其硬件功能进行任何修改, 而 FPGA 芯片在制造完成后, 其功能并未固定, 用户可以根据自己的实际需要, 将自己设计的电路通过 FPGA 芯片公司提供的专用 EDA 软件对 FPGA 芯片进行功能配置, 从而将空白的 FPGA 芯片转化为具有特定功能的集成电路芯片。

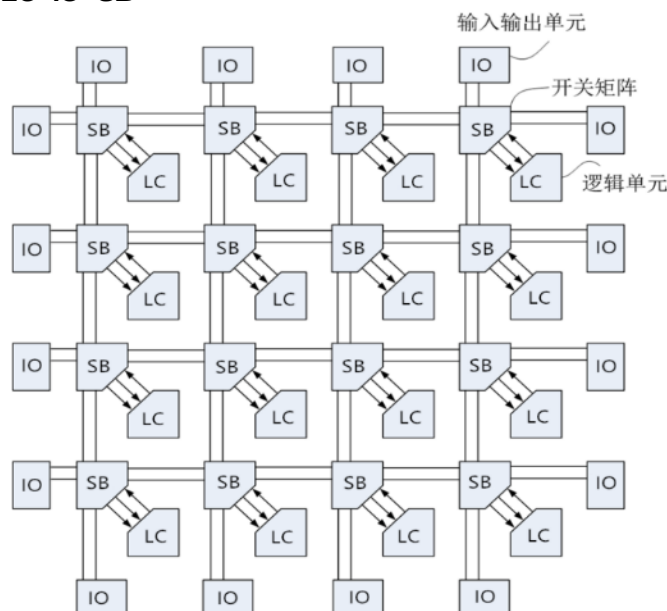
FPGA 的优势有以下四点: 1、可重构性: 可以被重新编程来实现不同的逻辑功能, 灵活度高, 能适应不同的应用需求。2、高性能: 由于 FPGA 可以被定制设计, 因此可以实现非常高效的逻辑运算。此外, FPGA 通常具有并行计算能力, 可以处理大量数据, 在某些应用中比传统处理器更快。3、低功耗: 由于 FPGA 可以被编程来执行特定的任务,



因此能更有效地利用能量，减少功耗。4、实时性：FPGA 可以实时处理输入数据，在需要实时响应的应用中具有很大的优势。

FPGA 芯片由可编程的逻辑单元（LogicCell, LC）、输入输出单元（InputOutputBlock, IO）和开关连线阵列（SwitchBox, SB）三个部分构成。

图表34: $FPGA=LC+IO+SB$



来源：安路科技招股说明书，国金证券研究所

3.4.2 功能实现：编程 LC 与 SB 改变逻辑电路进而实现不同的功能

用户通过对逻辑单元（LC）和开关阵列（SB）的编程，使 FPGA 内部形成不同的逻辑电路进而实现用户所需的功能。具体而言，用户将需要实现的电路功能用硬件描述语言进行描述，配套的 FPGA 专用 EDA 软件接受用户的功能描述和目标要求后，通过逻辑综合、布局布线和物理优化，最终编译生成二进制位流数据（Bitstream）。用户将位流数据存放在 FPGA 芯片上或芯片外的特定存储空间，FPGA 芯片启动后将位流加载到逻辑单元的 SRAM 数据存储空间和开关连线的 SRAM 控制存储空间，FPGA 芯片就实现了用户期望的特定功能。

图表35: 编程 LC 与 SB 改变逻辑电路进而实现不同的功能



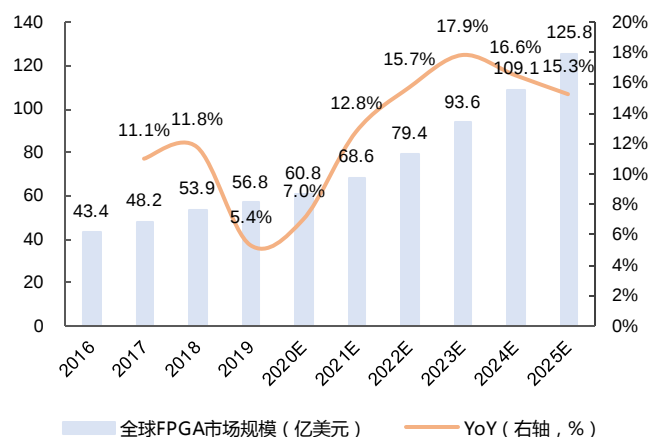
来源：安路科技招股说明书，国金证券研究所

3.4.3 市场规模：全球 FPGA 规模或近百亿美元，中国市场约占 37.9%

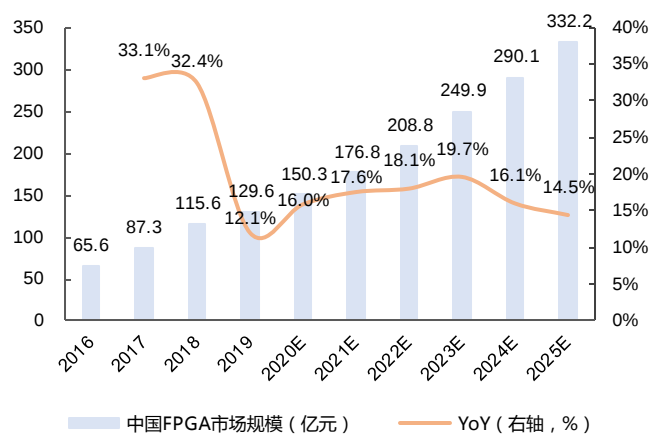
据 Frost&Sullivan，2023 年全球 FPGA 市场规模有望达 94 亿美元，18~23 年 CAGR 有望达 11.7%；同年，我国 FPGA 市场规模有望达 250 亿元，约占全球市场规模的 37.9%，18~23 年 CAGR 有望达 16.7%。



图表36: 2023 年全球 FPGA 市场规模有望达 94 亿美元



图表37: 2023 年我国 FPGA 市场规模有望达 250 亿元



来源: Frost&Sullivan, 国金证券研究所

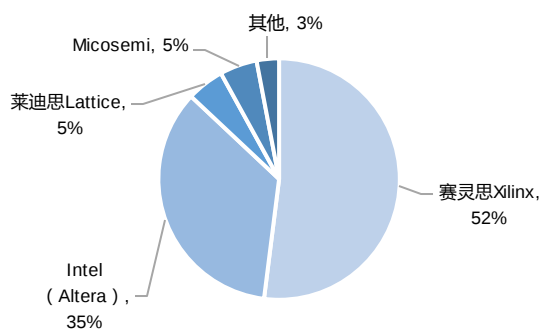
来源: Frost&Sullivan, 国金证券研究所

3.4.4 竞争格局: Xilinx 约占全球 FPGA 市场份额 52%, 安路科技为国产龙头厂商

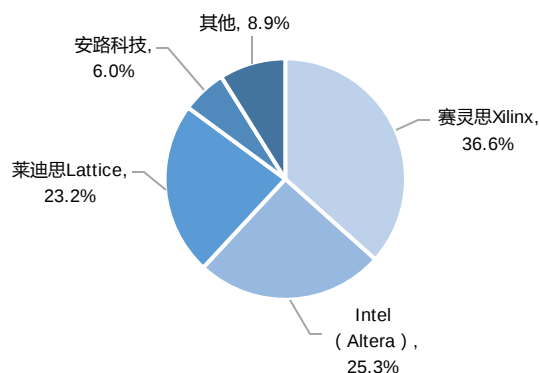
据中商产业研究院, 2022 年, Xilinx 约占全球 FPGA 市场份额 52%, Intel 旗下 Altera 约占 35%, Lattice 约占 5%, Microsemi 约占 5%, 全球市场仍由国外厂商主导。

据 Frost&Sullivan, 2019 年, Xilinx 约占我国 FPGA 市场份额 36.6%, 国产 FPGA 龙头安路科技约占 6%, 国产替代空间广阔。

图表38: 2022 年 Xilinx 约占全球 FPGA 市场份额 52%



图表39: 2019 年 Xilinx 约占我国 FPGA 市场份额 37%



来源: 中商产业研究院, 国金证券研究所

来源: Frost&Sullivan, 国金证券研究所

4. 中游-服务器/网络设备——国内服务器厂商领军全球, 光模块市场份额迅速提升

4.1 服务器: 管理计算资源的大型计算机, 全球市场规模超千亿美元

4.1.1 基本定义: 服务器在网络中为其它客户机提供计算或应用服务

服务器是为各类互联网用户提供综合业务的服务平台, 其功能是为网络中的客户机 (如 PC 机、智能手机、大型系统设备等终端) 提供计算及应用服务。

服务器由处理器、硬盘、内存、系统总线等软硬件构成, 和通用的计算机架构类似, 服务器需为计算机提供高可靠的服务, 因此在处理能力、稳定性、可靠性、安全性、可扩展性、可管理性等方面要求较高。

4.1.2 产品分类: 基于体系架构方式/基于应用层次/基于用途/基于机箱结构

服务器按照体系架构方式划分为 x86 服务器及 non-x86 服务器; 按应用层次划分为入门级服务器、工作组服务器、部门级服务器及企业级服务器; 按用途划分为通用型服务器、专用型服务器; 按服务器机箱结构划分为台式服务器、机架式服务器、机柜式服务器及刀片式服务器。



图表40: 服务器的分类结构

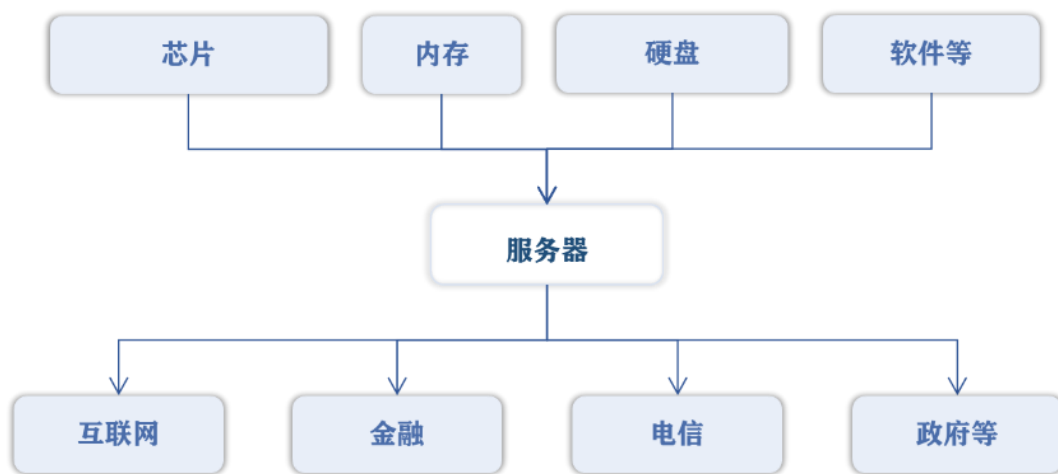


来源: 智研咨询, 国金证券研究所

4.1.3 产业链条: 上游-硬件设备/软件产品, 下游-应用市场

从产业链来看, 服务器行业上游为原材料及设备市场, 主要包括芯片、硬盘、内存等硬件设备和软件产品。服务器行业下游为应用市场, 主要有互联网、电信、金融、政府、能源和交通等领域。

图表41: 服务器产业链: 上游连接硬件设备/软件产品, 下游连接应用市场



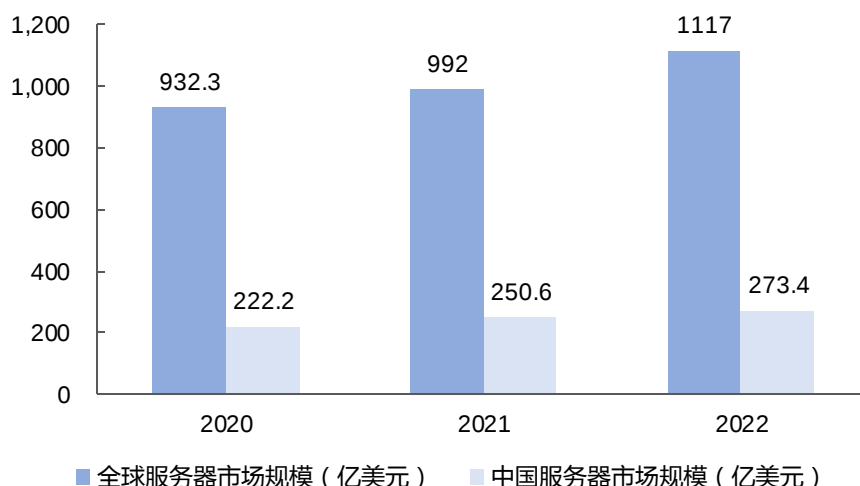
来源: 智研咨询, 国金证券研究所

4.1.4 市场规模: 2022 年全球服务器市场高达千亿美元, 我国约占全球市场四分之一

据 Counterpoint, 2022 年全球服务器市场将达到 1,117 亿美元, 同比增长 17%。在公共卫生事件后, 企业纷纷向混合云转型, 混合云模式推动组织采用 AIOps、基础设施自动化和边缘计算等新兴技术, 这促使现有基础设施升级, 为增长提供了动力。据 IDC, 2022 年我国服务器市场规模约在 273 亿元, 同比增长 9.1%, 约占全球市场的四分之一。



图表42: 2022 全球服务器市场规模高达千亿美元, 我国约占全球市场规模四分之一

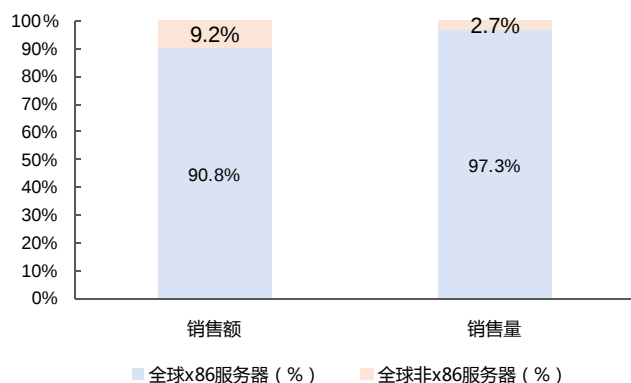


来源: Counterpoint, IDC, 观研天下, 国金证券研究所

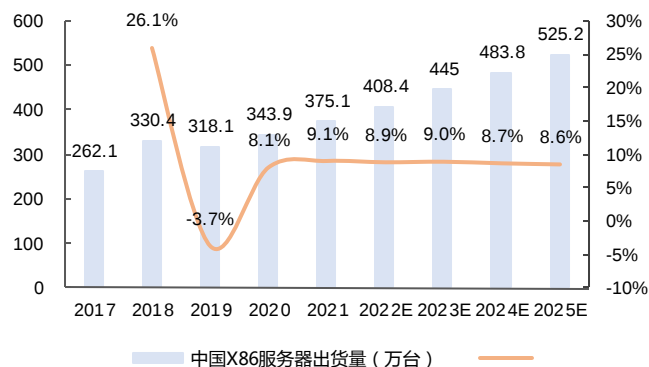
由于 X86 架构起步早, 具备成熟生态优势, 目前 X86 仍是全球服务器主流架构。据华经产业研究院, 2020 年全球 X86 服务器销售额 826.5 亿美元, 占比达到 90.8%, 销售量占比达到 97.3%。

据 IDC, 2021 年, 中国服务器出货量达到 391.1 万台, 同比增长 8.4%, 其中 X86 服务器出货量达到 375.1 万台。预计 2025 年中国 X86 服务器出货量将达到 525.2 万台。

图表43: X86 是全球服务器主流架构



图表44: 2023 年中国 X86 服务器出货量或达 445 万台



来源: 华经产业研究院, 国金证券研究所

来源: IDC, 国金证券研究所

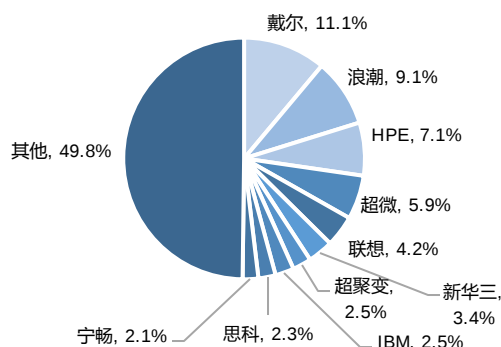
4.1.5 竞争格局: 戴尔/浪潮/HPE 全球市场 CR3, 浪潮/新华三/超聚变引领国内市场

据 IDC, 2023 年第三季度, 戴尔在全球服务器市场上位居榜首, 市场占比达到 11.1%, 同比减少 5.0%; 浪潮信息市场占比 9.1%, 同比增加 1.6%; HPE 市场份额 7.1%, 同比减少 3.3%, 排名第三。

据 IDC, 2022 年, 中国服务器市场排名前三的厂商依次为浪潮, 市场占比 28.1%; 新华三, 市场占比 17.2%; 超聚变, 市场占比 10.1%, 三者合计占比达到 55%, 占据国内服务器市场半壁江山。第二阵营品牌处于贴身缠斗之势。具体来看, 宁畅、中兴、戴尔、联想等品牌市场占比均在 5% 左右。

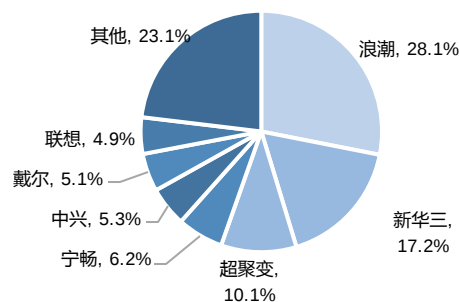


图表45: 2023 年戴尔、浪潮、HPE 占据全球服务器市场 R3



来源: IDC《2023 年第三季度全球服务器市场报告》, 国金证券研究所

图表46: 2022 年浪潮、新华三、超聚变引领国内服务器市场



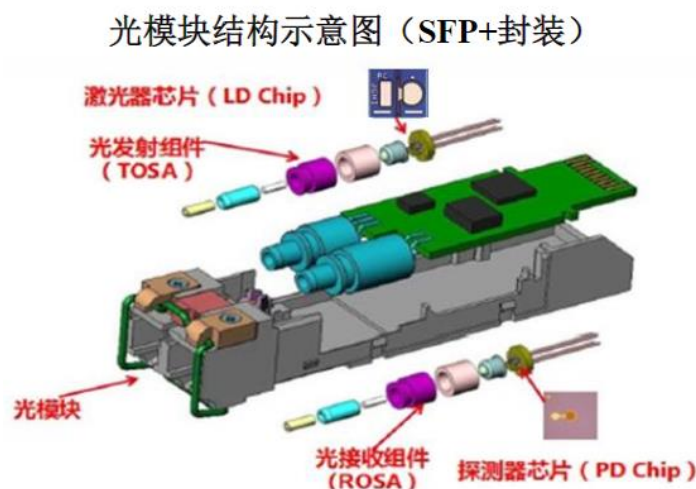
来源: IDC《2022 年中国服务器市场跟踪报告》, 国金证券研究所

4.2 光模块: 实现光电信号互相转换, 国产厂商占全球市场份额 50%以上

4.2.1 基本定义: 光模块的功能是实现光电转换, 是光通信系统的核心器件

光模块的功能是实现光电转换, 是光通信系统的核心器件。光模块在发射端 (TOSA) 通过激光器芯片将电信号转换为光信号, 经过光纤传输至接收端 (ROSA), 在接收端通过探测器芯片将光信号转换为电信号, 实现信号传输。其中光芯片是决定光通信系统信号传输效率和网络可靠性的关键。按照功能光芯片可以分为激光器芯片 (LD) 和探测器芯片 (PD)。

图表47: 光模块结构示意图



来源: 源杰科技招股说明书, 国金证券研究所

4.2.2 产品分类: 分类方式多样, 小型化、高速率、低功耗、低成本是发展趋势

从光模块分类角度, 由于应用场景较多需求各异, 因而分类方式多样, 命名复杂。光模块常见的分类方式包括了封装类型、速率、距离、激光器类型、探测器类型等。整体而言, 小型化、高速率、低功耗、低成本是光模块整体的发展趋势。



图表48：光模块分类

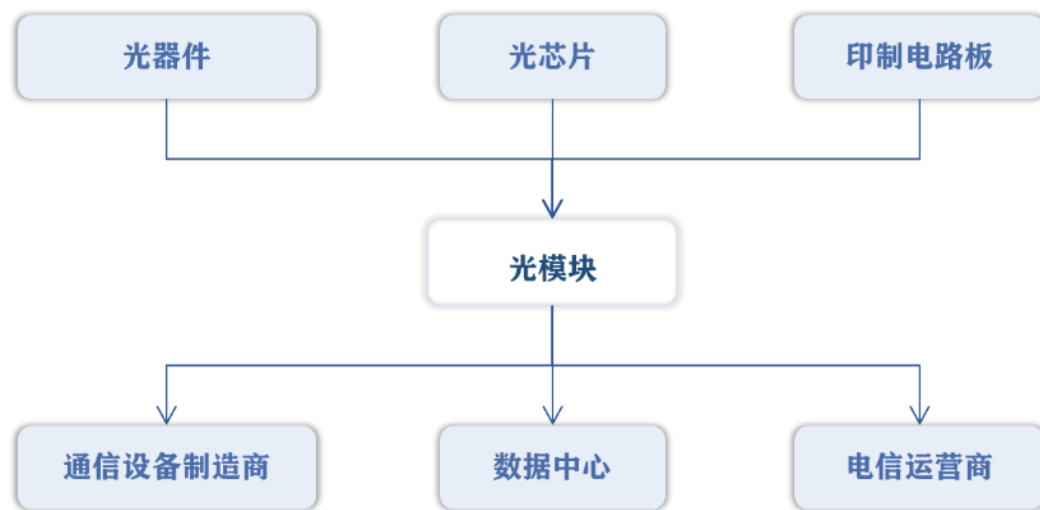
分类方式	类型
封装类型	GBIC、XENPAK、XFP、SFP、SFP+、SFP28、QSFP、QSFP28、CFP、CFP2、QSFP-DD、OSFP 等
速率	10Mbps、100Mbps、155Mbps、622Mbps、1.25Gbps、2.125Gbps、4.25Gbps、10Gbps、25Gbps、50Gbps、100Gbps、400Gbps 等
波长	850nm、1310nm、1490nm、1550nm、CWDM、DWDM 等
距离	100m、300m、550m、10km、20km、40km、80km、120km、160km 等
调制格式	NRZ、PAM4、DP-QPSK/n-QAM 等
是否支持 WDM	灰光模块（不支持 WDM）、彩光模块（支持 WDM）
激光器类型	垂直腔面发射激光器（VCSEL）、法布里-珀罗激光器（FP）、分布式反馈激光器（DFB）、电吸收调制激光器（EML）等

来源：华经产业研究院，国金证券研究所

4.2.3 产业链条：上游-元器件等，下游-通信设备制造/数据中心/电信运营商

光模块行业产业链上游为光器件、集成电路芯片、印制电路板等元器件供应商。中游是光模块生产企业，近年来，中国光模块企业不断进行并购重组，垂直整合产业链，行业集中度进一步提高，光模块供应商逐渐在全球市场上获得份额，中国的供应商目前在全球以太网光模块市场占主导地位。光模块行业下游主要应用于通信设备制造、数据中心、电信运营商等领域。

图表49：光模块产业链：上游连接元器件供应商，下游连接通信设备制造等领域



来源：智研咨询，国金证券研究所

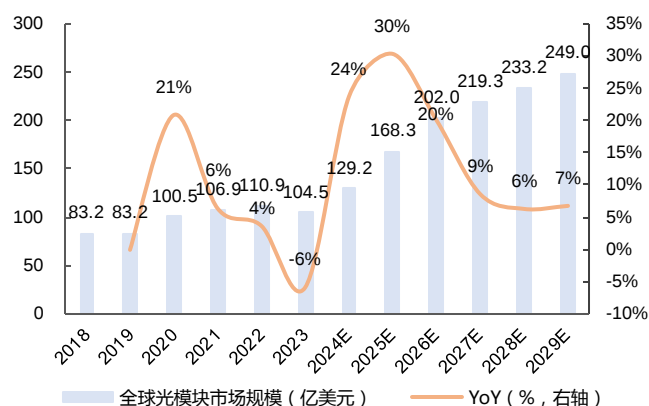
4.2.4 市场规模：全球光模块市场规模突破百亿美元，中国市场约占四分之一

据 Lightcounting，全球光模块市场规模在 16~19 年增速放缓，但在 20 年和 21 年受到居家学习、办公的网络需求驱动，恢复高速增长态势。2020 年增速达到 21%，市场规模突破 100 亿美元；2021 年增速达到 6%，市场规模达到 107 亿美元。但在 2023 年，随着电信运营商资本开支减少、云计算公司在经历了几年两位数的增长后资本开支增长率放缓，2023 年光模块市场小幅下降。未来随着 5G 产业持续渗透和新一轮全球数据中心建设，2024 年全球光模块市场规模预计达到 130 亿美元。

目前国内计算和存储能力的需求不断增加，同时在当前 5G 时代的发展背景之下，整体通信行业呈现高景气运行，以光模块等为代表的核心市场正在迅速发展，市场规模迎来不断提升。据华经产业研究院，2021 年我国光模块行业市场规模从 2016 年的 250.9 亿元增长至 415.22 亿元，2022 年市场规模约为 476.82 亿元，预计 2023 年中国光模块市场规模有望达到 554.5 亿元。

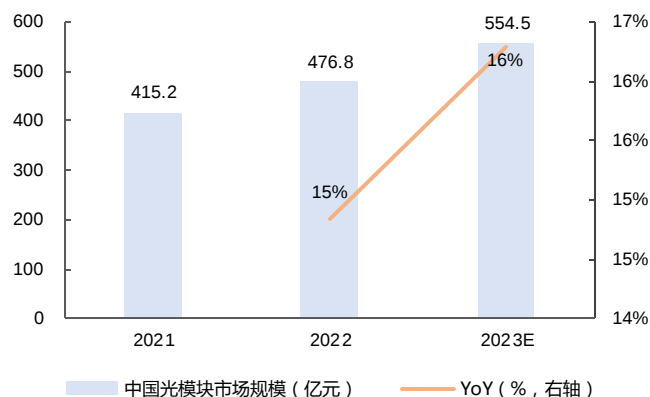


图表50: 2023 年全球光模块市场规模突破百亿美元



来源: Lightcounting, 国金证券研究所

图表51: 2022 年我国光模块市场规模近 500 亿元



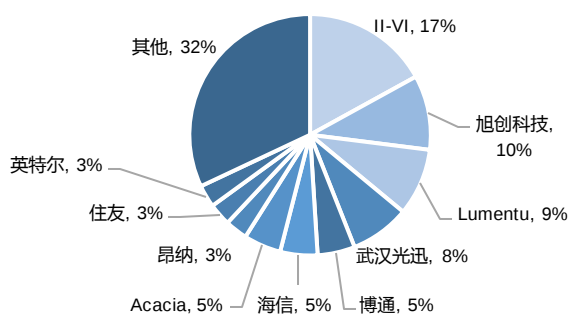
来源: 华经产业研究院, 国金证券研究所

4.2.5 竞争格局: 国产厂商份额不断提升, 占据全球光模块市场份额超过 50%

国产厂商份额不断提升。中国光模块厂商凭借劳动力成本等的优势, 在与海外光模块厂商竞争中不断占据上风。据 Odmia, 2015 年, 全球前十大光模块厂商仅光迅科技一家中国企业, 到 2021 年, 光迅科技、中际旭创、海信宽带、昂纳信息进入全球前十, 合计占据全球 26% 的市场份额。

据 Lightcounting, 2021 年, 中际旭创等五家供应商进入全球前十, 主要厂商相较前一年市场份额均有提升; 2022 年, 光模块全球市前 10 名中国占据 7 家, 其中中际旭创、Coherent、思科、华为四家厂商占据全球光模块市场份额超过 50%, 中际旭创和 Coherent 分别获得近 14 亿美元的收入。

图表52: 2021 年全球光模块市场份额



来源: Odmia, 国金证券研究所

图表53: 2022 年全球光模块市占率前 10 中国占据 7 家

排名	2010	2016	2018	2021	2022
1	Finisar	Finisar	Finisar	中际旭创 & II-VI	中际旭创 & Coherent
2	Opnext	海信宽带	中际旭创		
3	Sumitomo	光迅科技	海信宽带	华为海思	Cisco(Acacia)
4	Avago	Acacia	光迅科技	Cisco(Acacia)	华为海思
5	Source Photonic	FOPT(Avago)	FOPT(Avago)	海信宽带	光迅科技
6	Fujitsu	Oclaro	Lumentum	Broadcom	海信宽带
7	JDSU	中际旭创	Acacia	新易盛	新易盛
8	Emcore	Sumitomo	Intel	光迅科技	华工正源
9	WTD	Lumentum	AOI	Molex	Intel
10	NeoPhotonics	Source Photonic	Sumitomo	Intel	Source Photonic

来源: Lightcounting, 国金证券研究所

4.3 交换机: 提供网络连接端口, 思科独占全球市场份额 40%

4.3.1 基本定义: 交换机为接入它的任意两个网络节点提供独享的电信号通路

交换机是基于以太网进行数据传输的多端口网络设备, 每个端口都可以连接到主机或网络节点, 主要功能就是根据接收到数据帧中的硬件地址, 把数据转发到目的主机或网络节点。交换机相当于一台特殊的计算机, 由硬件和软件组成, 包括中央处理器、存储介质、接口电路及操作系统等。

4.3.2 产品分类: 基于网络覆盖范围/基于传输速率和介质/基于应用规模等

交换机可根据网络覆盖范围、传输速率和介质、应用规模等进行分类。在各类交换机中, 最常见的交换机是以太网交换机。交换机从广义上还可分为广域网交换机和局域网交换



机，广域网交换机主要在电信领域中提供通信的基础平台，局域网交换机则用于连接局域网内的终端设备。从传输介质和传输速度上可分为以太网交换机、快速以太网交换机、千兆以太网交换机、FDDI 交换机、ATM 交换机和令牌环交换机等。从应用规模上可分为企业级交换机、部门级交换机、工作组交换机等。

图表54：交换机的分类

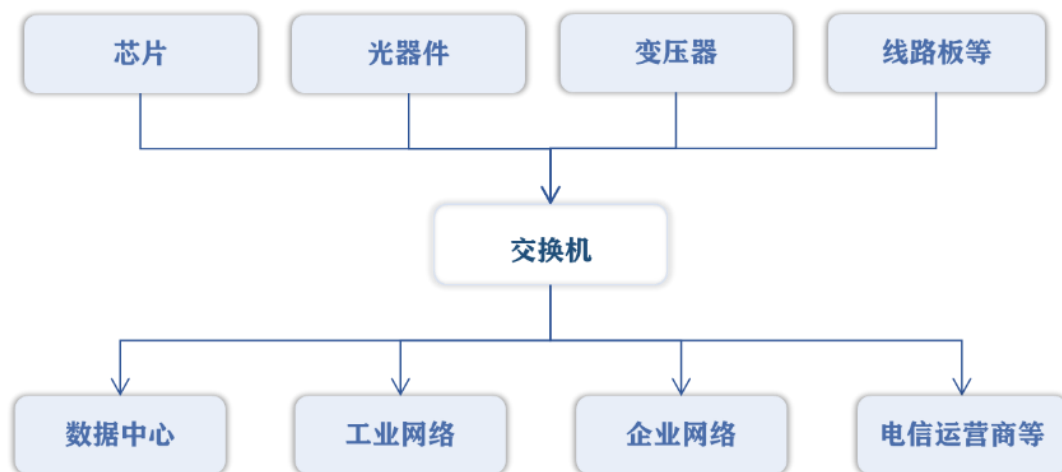
基于网络覆盖范围划分	基于传输介质和传输速度划分	基于端口结构划分	基于应用规模划分	基于是否支持网管功能划分	基于工作协议层划分
广域网交换机	以太网交换机	固定端口交换机	企业级交换机	网管型交换机	第二层交换机
	快速以太网交换机		校园交换机		
	千兆以太网交换机		部门级交换机		第三层交换机
局域网交换机	万兆以太网交换机	模块化交换机	工作组交换机	非网管型交换机	第四层交换机
	ATM交换机		卓机型交换机		
	FDDI交换机				
	令牌环交换机				

来源：头豹咨询，国金证券研究所

4.3.3 产业链条：上游-电子元器件，下游-数据中心等领域

从产业链来看，交换机上游主要是电子元器件行业，使用的原材料主要包括：芯片、光器件、电源模块、连接器、线路板、变压器、阻容等，行业基本处于充分竞争状态，企业数量众多，产能充足，价格稳定，供给状况无重大波动，对交换机发展不构成约束和重大影响。交换机下游可应用于数据中心等领域。

图表55：交换机产业链：上游连接电子元器件，下游连接数据中心等领域



来源：智研咨询，国金证券研究所

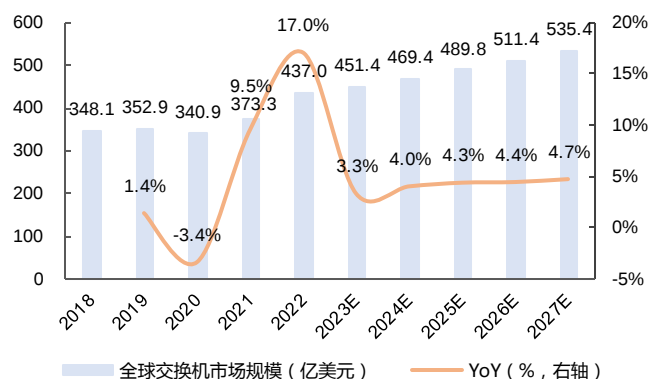
4.3.4 市场规模：全球交换机市场规模达 400 亿美元，我国约占全球市场规模 17%

据 IDC，2022 年全球交换机行业市场规模达到 437 亿美元，同比增长 17.0%，且预计未来 5 年的增速稳定在 4%左右，预计 2027 年规模将达到 535 亿美元。

据 IDC，近年来，中国交换机市场规模一直保持稳定增长趋势。2022 年中国交换机市场约为 72 亿美元，同比增长 9.4%。预计 2023 年将进一步增长至 78 亿美元。

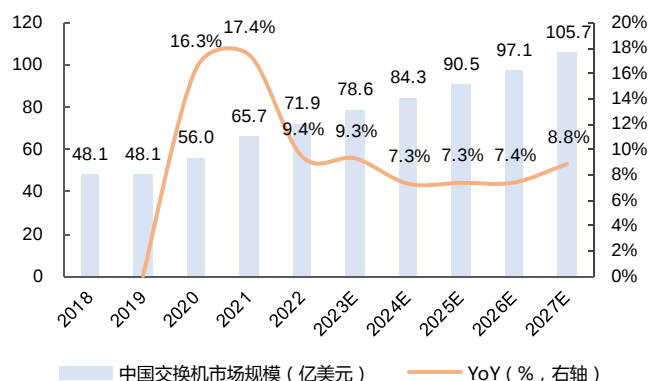


图表56: 2022 年全球交换机市场规模达 400 亿美元



来源: IDC, 国金证券研究所

图表57: 2022 年我国交换机市场规模达 70 亿美元



来源: IDC, 国金证券研究所

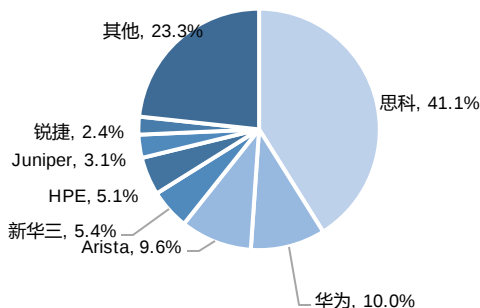
4.3.5 竞争格局: 思科是全球交换机市场绝对龙头, 华为、新华三、锐捷合计占国内市场超过八成

交换机行业集中度较高, 思科、华为、新华三等少数几家企业占据着绝大部分的市场份额, 呈现寡头竞争的市场格局。

据 IDC, 全球交换机市场方面, 2022 年思科是绝对龙头, 市场份额占有率为 41.1%; 5 家龙头企业占据全球市场的 70% 以上。

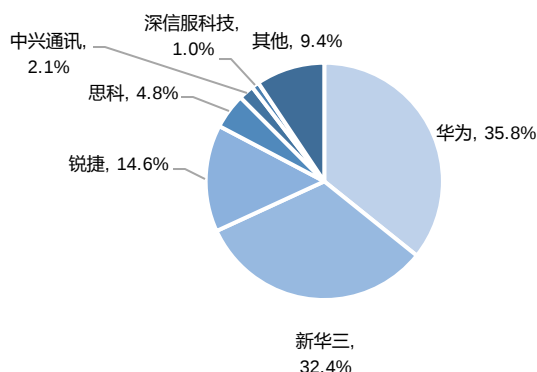
据 IDC, 中国交换机市场前三名企业占整体市场超过 80%。其中华为占比最多, 达 35.8%, 排名第一。其次为新华三, 占比约为 32.4%, 排名第二。星网锐捷排名第三, 占比 14.6%。

图表58: 2022 年思科是全球交换机市场绝对龙头



来源: IDC, 国金证券研究所

图表59: 2022 年华为、新华三、锐捷合计占国内市场超过八成



来源: IDC, 国金证券研究所

4.4 存储设备: 国外厂商垄断全球存储市场, 三星、海力士、美光合计占比超过 80%

4.4.1 基本定义: 存储器是用来存储程序和各种数据信息的记忆部件

存储器的工作原理为: 通过使用地址编址和电子静态存储技术实现数据的存储和读取。存储器通常被组织成一个二维矩阵, 其中每个单元称为一个存储位置。在需要读取或写入数据时, CPU 向存储器发送地址信号, 通过数据总线与存储器进行数据的传输。

4.4.2 产品分类: NAND 和 DRAM 市场是主要细分市场

存储器主要可分为外部存储器和内部存储器。其中, 外部存储器包含半导体存储器、磁存储器和光存储器, 内部存储器包含易失性存储器和新型非易失性存储器。按照读写方式, 半导体存储器可分为只读存储器 ROM (EEPROM、PROM、EPROM) 和闪存存储介质 (NOR Flash、NAND Flash)。磁存储器件主要包含磁带、磁盘。光存储器件主要包含蓝光光盘 (BD)、归档光盘 (AD)、全息存储、玻璃存储等。



图表60：存储器主要可分为外部存储器和内部存储器



来源：头豹咨询，国金证券研究所

据 WSTS 数据，全球储存市场中，DRAM 占据 61% 的市场份额，NAND 占据 36% 的市场份额，是最主要的两大存储细分市场。

图表61：NAND 和 DRAM 对比

对比角度	NAND	DRAM
存储原理	浮栅型	电容充放电型
读写速度	较慢	极快
存储容量	高 (GB / TB)	中 (MB / GB)
易失性	非易失性	易失性
分类	2D NAND、3D NAND	DDR、LPDDR (低功耗) 和 GDDR
应用领域	2D NAND 可用于医疗、机器人、可穿戴、安防摄像头、机顶盒等各种 IoT 设备；3D NAND 多见于数据中心、智能手机、PC 等	DDR 主要应用于服务器和 PC 端；LPDDR 主要应用于手机端；GDDR 主要应用于图像处理领域

来源：TechInsights、东芯半导体招股说明书，国金证券研究所

4.4.3 市场规模：全球 NAND 市场规模达 600 亿美元，全球 DRAM 市场规模近千亿美元

据中商产业研究院，2019 年受智能手机终端需求疲软影响，加之国产企业实现量产，整体价格下降，全球 NAND 市场规模出现较大下降，随着 5G 推动智能手机需求回暖加之 PC 等需求稳步增长，市场回暖，2021 年全球 NAND 市场规模达 600 亿美元。预计随着

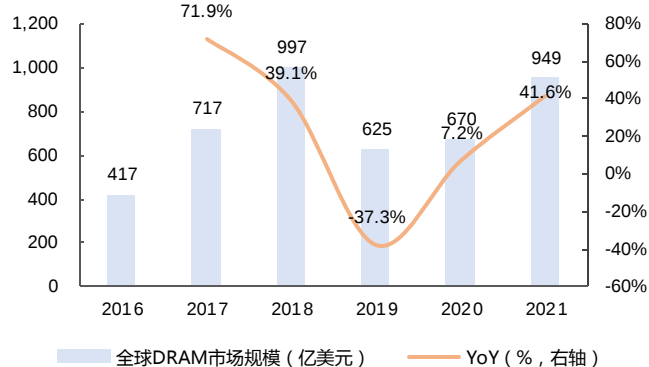
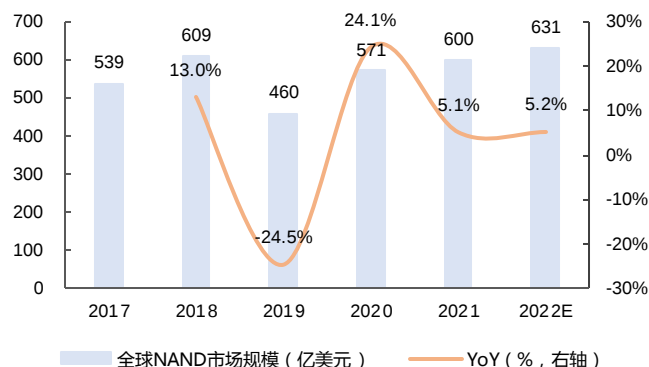


物联网、自动驾驶等新兴产业智能化推进，存储芯片需求将出现快速增长。

据 Trend Force，2021 年，全球 DRAM 市场规模达到 949 亿美元，同比增长 41.6%。

图表62：2022 年全球 NAND 市场规模近 700 亿美元

图表63：2021 年全球 DRAM 市场规模近千亿美元



来源：中商产业研究院，国金证券研究所

来源：Statista, TrendForce, 国金证券研究所

4.4.4 竞争格局：国外厂商垄断全球存储市场，三星据 NAND、DRAM 市场份额首位

全球存储市场绝大部分份额由国外厂商占有，呈现寡头垄断格局，行业集中度较高。

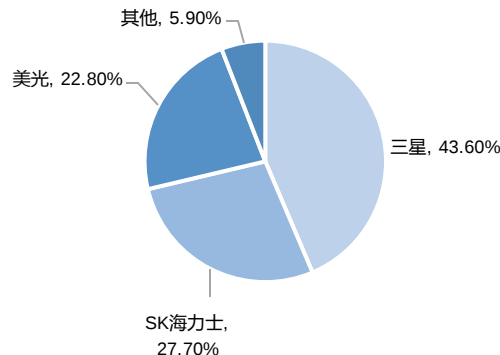
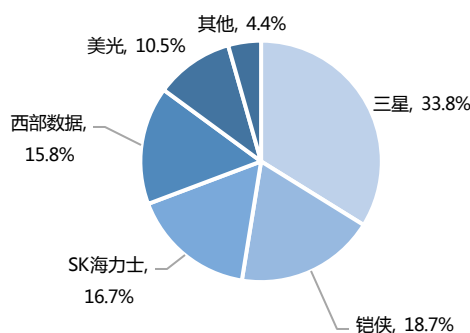
据 CFM 闪存市场数据，2022 年全球 NAND flash 市场前三大厂商分别为三星、铠侠和 SK 海力士，市场份额分别为 33.8%、18.7%和 16.7%。

根据 IC Insights，2021 年全球 DRAM 市场几乎由三星、SK 海力士和美光所垄断，分别占比 44%、28%和 23%。

国内存储产业的重点公司有：长江存储，在 NAND Flash 市场份额约占全球 1%；长鑫存储，中国 DRAM 龙头企业，计划于今年在科创板上市，估值不低于 1000 亿元；北京君正，其收购的北京矽成在汽车 DRAM 领域市占率 15%居全球第二；兆易创新，产品线 Nor+NAND+DRAM 全覆盖，其中 Nor 全球第三、大陆第一；江波龙，2022 年在 eMMC 及 UFS 市场份额位列全球第六；佰维存储，在 eMMC 及 UFS 全球市占率 2.4%，居全球第八、国内厂商第二；澜起科技，在内存接口芯片市场位列全球前二。

图表64：2022 年三星占全球 NAND 市场三分之一份额

图表65：2021 年三星、SK 海力士、美光垄断全球 DRAM 市场



来源：CFM 闪存市场，国金证券研究所

来源：IC Insights, 国金证券研究所

5. 下游-数据中心/云服务——东数西算铸基，算力租赁牵引

5.1 数据中心：数据存储和交换的中心，Equinix 是全球 IDC 龙头

5.1.1 基本定义：拥有完善的设备、专业化的管理、完善的应用级服务的服务平台

IDC 全称为 Internet Data Center，互联网数据中心。只提供场地和机柜的数据中心，一般称为 DC(Data Center)，而同时提供带宽服务的，一般称 IDC(互联网数据中心，Internet Data Center)，两者有时不作严格区分。IDC 是指一种拥有完善的设备(包括高

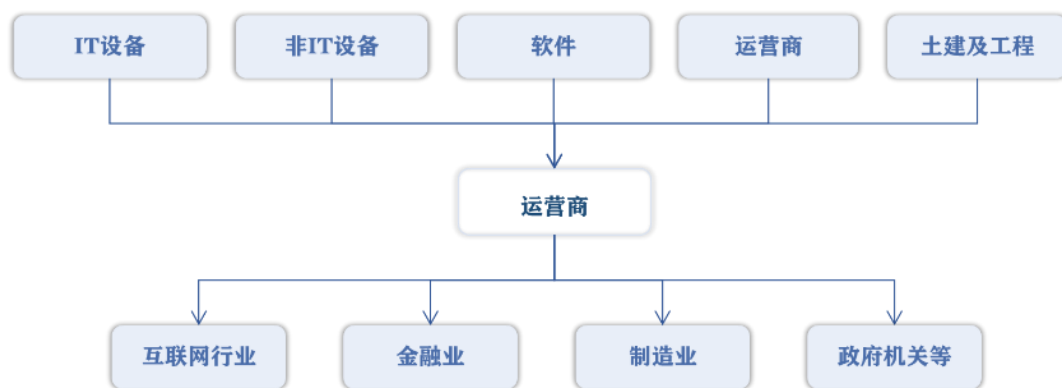


速互联网接入带宽、高性能局域网络、安全可靠的机房环境等)、专业化的管理、完善的应用级服务的服务平台。在 IDC 平台基础上, IDC 服务商为企业和 ISP、ICP、ASP 等客户提供互联网基础平台服务以及各种增值服务。

5.1.2 产业链条: 上游-基础设施, 中游-数据中心运营服务商, 下游-应用行业

数据中心产业链中, 上游为基础设施主要包括 IT 设备、非 IT 设备、软件及建设工程; 中游为数据中心运营服务商; 下游为应用行业, 主要包括互联网行业、金融业、软件业及制造业等。

图表66: IDC 产业链: 上游连接基础设施, 下游连接应用行业

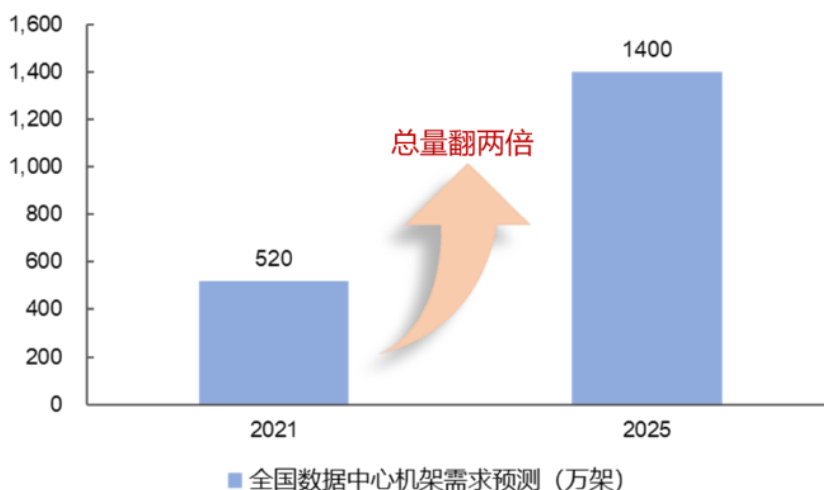


来源: 中商产业研究院, 国金证券研究所

5.1.3 市场规模: 全球 IDC 市场达 800 亿美元, 我国 IDC 市场规模在 2,500 亿元左右

2022 年 2 月, “东数西算”工程正式拉开了构建全国一体化大数据中心体系的大幕。随着 8 大国家算力枢纽节点建设方案均进入深化实施阶段, 全国数据中心机架规模逐步增长。根据相关统计数据, 截止 2021 年底, 全国在用数据中心机架规模达 520 万架, 其中大型以上数据中心机架规模约为 420 万架(按 2.5Kw 标准机架统计), 预计至 2025 年“十四五”规划期末, 拟实现数据中心机架规模增长至 1400 万架, 规模总量翻两倍, 总增量投资约 7000 亿元。

图表67: 全国数据中心机架需求预计将在 2025 年达到 1400 万架



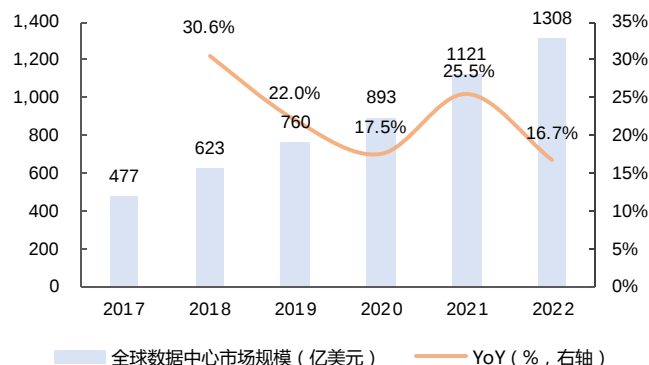
来源: 中国通服数字基建产业研究院《中国数据中心产业发展白皮书(2023年)》, 国金证券研究所

从全球市场来看, 全球数据中心产业规模在 2022 年达到 1308 亿美元, 2019-2022 年 CAGR 约 20%。

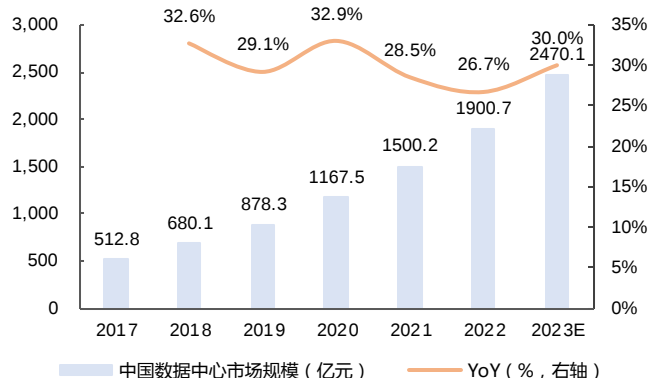
受新基建、数字化转型及数字中国远景目标等国家政策促进及企业降本增效需求的驱动, 我国数据中心市场规模持续高速增长。2021 年, 我国数据中心市场规模达到 1500 亿元, 近五年 CAGR 达到 29.96%。随着我国各地区、各行业数字化转型的深入推进, 我国数据中心市场规模将保持持续增长态势, 预计 2023 年市场规模将达到 2470 亿元。



图表68: 2022 年全球数据中心市场规模达 1300 亿美元



图表69: 2023 年我国数据中心市场规模在 2,500 亿元左右



来源: 中国通服数字基建产业研究院, 国金证券研究所

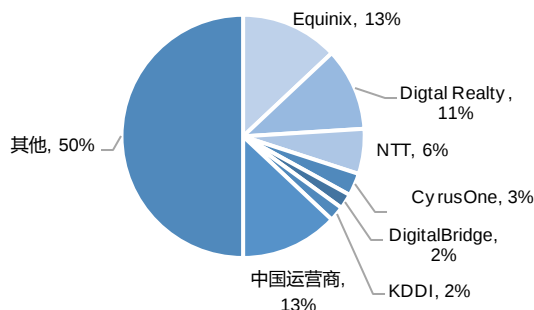
来源: 中国信通院, 国金证券研究所

5.1.4 竞争格局: Equinix 是全球 IDC 龙头, 国内市场三大运营商合计占比超 50%

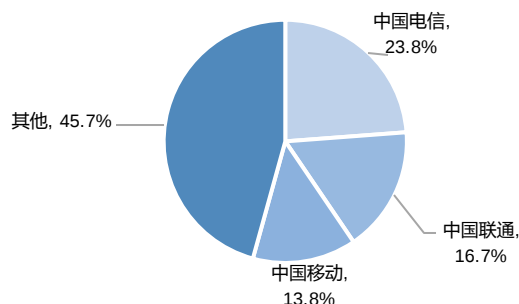
全球 IDC 企业的竞争格局中, 美国第三方数据中心 Equinix 公司是全球数据中心的龙头企业, 2022 年 Q2, 该公司全球市场份额占比达 13%。其次是 Digital Realty Trust 公司, 同样位于美国, 市场份额占比约 11%, NTT 排名第三, 在全球 IDC 市场占有 6% 的份额。中国运营商合计在国际市场占有 13% 的市场份额。

据 IDC, 2020 年我国数据中心市场格局以运营商数据中心为主, 凭借其网络带宽和机房资源优势, 三大电信运营商市场份额占比超 50%; 万国数据、世纪互联、光环新网等第三方数据中心近年来逐渐兴起, 满足核心城市的数据中心需求, 弥补供需缺口。

图表70: 2022 年 Equinix 是全球 IDC 龙头



图表71: 2020 年三大运营商占据国内市场超 50% 份额



来源: Synergy, 国金证券研究所

来源: IDC, 国金证券研究所

5.2 云服务: 分布式计算, AWS 独占全球市场份额 40%

5.2.1 基本定义: 通过网络以按需、易扩展的方式获得所需服务

云服务, 即云计算服务, 是一种按使用量付费的模式, 这种模式提供可用的、便捷的、按需的网络访问, 进入可配置的计算资源共享池 (资源包括网络, 服务器, 存储, 应用, 软件, 服务), 这些资源能够被快速提供, 只需投入很少的管理工作, 或服务供应商进行很少的交互。

5.2.2 产品分类: 基础设施即服务(IaaS)/ 平台即服务(PaaS)/ 软件即服务(SaaS)

云计算的服务类型通常可分为三类, 即基础设施即服务(IaaS)、平台即服务(PaaS)和软件即服务(SaaS)。

- 基础设施即服务 (IaaS), 向云计算提供商的个人或组织提供虚拟化计算资源, 如虚拟机、存储、网络和操作系统。
- 平台即服务(PaaS), 为开发人员提供通过全球互联网构建应用程序和服务的平台, PaaS 为开发、测试和管理软件应用程序提供按需开发环境。

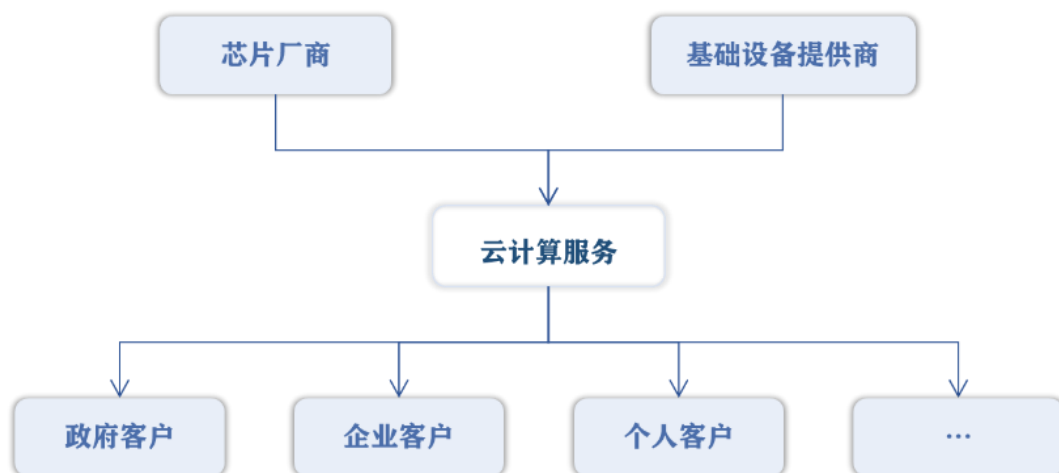


- 软件即服务(SaaS)，通过互联网提供按需软件付费应用程序，云计算提供商托管和管理软件应用程序，并允许其用户连接到应用程序并通过全球互联网访问应用程序。

5.2.3 产业链条：上游-芯片/基础设备，中游-IDC 制造商/云服务提供商，下游-客户端

云计算行业产业链的上游是芯片厂商和基础设备提供商，包括芯片、服务器、路由器、交换机等；中游参与者则是 IDC 制造商和云服务提供商，包括 IaaS、PaaS、SaaS；下游则是对应的最终客户端，包括互联网、金融等行业的政府客户、企业客户以及个人用户等。

图表72：云计算产业链：上游连接芯片/基础设备，下游连接客户端



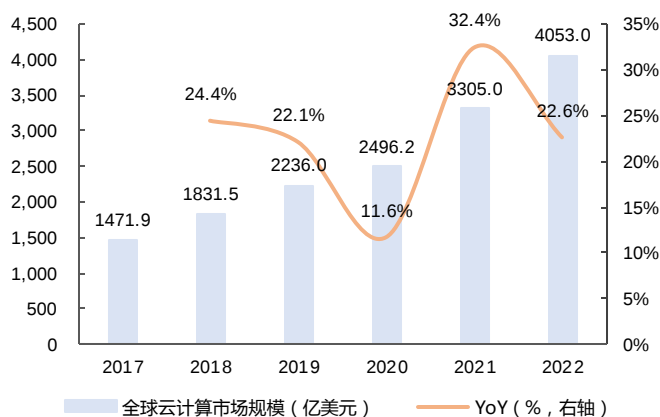
来源：华经产业研究院，国金证券研究所

5.2.4 市场规模：全球云计算规模达 4,000 亿美元，中国市场规模 4,500 亿元

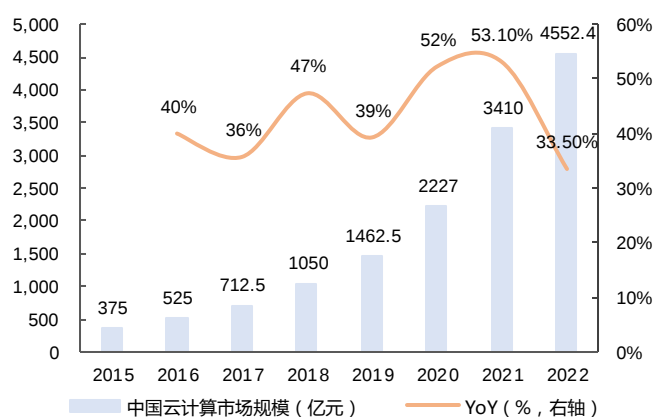
据中国信通院，2021 年全球云计算市场规模达到 3,305 亿美元，同比增长 32.44%。初步统计 2022 年全球云计算市场规模 4053 亿美元，增幅 22.6%。

中国云计算产业正处蓬勃发展阶段。据智研咨询，2021 年中国云计算市场规模达 3,410 亿元，增速 53.1%。初步统计 2022 年中国云计算市场规模达到 4552.4 亿元，增长 33.5%。

图表73：2022 年全球云计算市场规模达 4,000 亿美元



图表74：2022 年我国云计算市场规模达 4,500 亿元



来源：中国信通院，国金证券研究所

来源：智研咨询，国金证券研究所

5.2.5 竞争格局：AWS 约占全球云计算市场份额 40%，国内四大云计算厂商占比 79%

据 Gartner，全球云计算市场竞争格局为一超多强，2022 年 AWS 在全球云基础设施服务市场份额为 40%，保持行业第一位置，Azure、阿里云、谷歌云分别占据第二至四位，分别占比 21.5%、7.7%、7.5%。

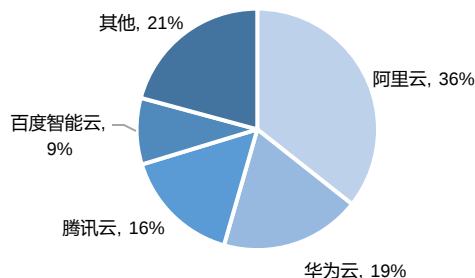
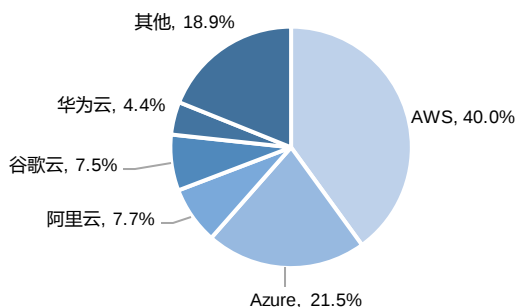
据 Canalys，2022 年国内前四大云计算厂商分别为阿里云、华为云、腾讯云和百度智能云，占云服务客户支出总额的 79%。阿里云在 2022 年仍然排在第一位，市场份额 36%；



华为云以 19% 的市场份额排名第二，每年增长 13%，领先于整体市场增长；腾讯云排名第三，市场份额 16%，受到内部业务重组的影响，收入规模增长放缓；百度智能云市场份额占比 9%，同比增长 11%。

图表75：2022 年 AWS 约占全球云计算市场份额 40%

图表76：2022 年国内四大云计算厂商共占市场份额 79%



来源：Gartner，国金证券研究所

来源：Canalys，国金证券研究所

6. 投资建议

推荐关注 AI 算力产业相关计算机标的，如浪潮信息，中科曙光，宝信软件，神州数码，高新发展等。

图表77：推荐 AI 算力产业相关的计算机赛道 22 家主要标的

股票代码	公司名称	推荐理由
寒武纪	688256.SH	寒武纪系国内 AI 芯片龙头。公司产品线涵盖云端、边缘端和 IP 授权及软件四大板块，各产线产品持续高端化迭代。云端智能芯片为业务核心，同时延伸发展训练整机与智能计算集群，较前代产品算力提升显著。公司 2023 年业绩快报对应 EPS 约为-2.01 元，对应-86XPE。预计 24 年摊薄 EPS 为-1.34 元，对应-130XPE。
海光信息	688041.SH	海光信息系国内稀缺的 CPU+DCU 涉及厂商。公司 CPU 产品性能不仅国内领先并已逐步追上海外主流产品，DCU 产品迭代顺利，对半精度/单精度能力进行针对性加强。公司 2023 年业绩快报对应 EPS 约为 0.54 元，对应 142XPE。预计 24 年摊薄 EPS 为 0.73 元，对应 106XPE。
龙芯中科	688047.SH	龙芯中科系国产 CPU 重要厂商。主营业务包括处理器及配套芯片的研制、销售及服务，公司研制的芯片包括龙芯 1 号、龙芯 2 号、龙芯 3 号三大系列处理器芯片及桥片等配套芯片。公司 2023 年业绩快报对应 EPS 约为-0.82 元，对应-106XPE。预计 24 年摊薄 EPS 为 0.22 元，对应 392XPE。
景嘉微	300474.SZ	景嘉微系国产 GPU 图显龙头。公司立足并成长于军用图形显控和小型化雷达领域，加强在民用高性能 GPU 芯片等消费领域的产品开发及推广，从“0”到“1”打造了三代图形显控产品(JM5 系列、7 系列、9 系列)；2023 年公司开始布局算力产品。预计 23、24 年摊薄 EPS 分别为 0.15、0.54 元，对应 483X、138XPE。
深桑达	000032.SZ	深桑达系中国电子数据创新业务的核心参与企业。公司已形成数据治理工程总体设计、数据金库、数据要素加工交易平台、安全可信数据空间、数据要素蝶变器、数据资产登记服务平台等数据安全与数据要素化工程系列产品。预计 23、24 年摊薄 EPS 分别为 0.22、0.38 元，对应 85X、47XPE。
优刻得	688158.SH	优刻得系国内中立云计算服务领域的领先企业。公司主要通过公有云、私有云、混合云三种模式，为客户提供计算、网络、存储等 IaaS 和 PaaS 产品，此外，还涉及大数据和人工智能产品的开发和服务。公司 2023 年业绩快报对应 EPS 约为-0.72 元，对应-18XPE。预计 24 年摊薄 EPS 为-0.37 元，对应-35XPE。
浪潮信息	000977.SZ	浪潮信息系算力系统供应商。公司支持多元异构算力，可适配多种架构的 AI 加速芯片。公司致力于积极推进更加多元化的算力系统研发创新，为客户提供持续稳定的产品服务。预计 23、24 年摊薄 EPS 分别为 1.21、1.61 元，对应 35X、27XPE。
神州数码	000034.SZ	神州数码系国内 IT 分销龙头。公司主营业务包括 IT 分销及增值服务业务、云计算和数字化转型业务和信创业务，公司 IT 分销及增值服务业务可根据最终销售客户和产品进一步划分为消费电子业务和企业增值业务。预计 24、25 年摊薄 EPS 分别为 1.75、2.20 元，对应 17X、13XPE。



中科曙光	603019.SH	中科曙光系国产算力领军者。公司积极建设“全国一体化算力服务平台”，致力于链接遍布各地各类算力中心，为千行百业提供“先进、绿色、无损算力”和“集算力、数据、应用、运营、运维为一体的服务”。预计 23、24 年摊薄 EPS 分别为 1.26、1.66 元，对应 38X、29XPE。
同方股份	600100.SH	同方股份系国内高科技企业领先者。公司主要从事核技术应用、数字信息、智慧能源等业务，其中，数字信息业务主要包括了计算机产品、知识内容与服务领域、大数据产业等领域。
中国长城	000066.SZ	中国长城系中国电子自主计算产业的主力军。公司聚焦计算产业与系统装备两大主业，形成了 8 项业务 10 家骨干企业聚力发展的新格局。公司致力于构建统一的“鹏腾”生态，开创通用算力新格局。预计 23、24 年摊薄 EPS 分别为 0.03、0.11 元，对应 294X、88XPE。
拓维信息	002261.SZ	拓维信息系华为云首批盘古大模型生态合作伙伴、“昇腾智造/智行”双领域解决方案合作伙伴。公司自主品牌“兆瀚”创新发布了兆瀚 SA200I A2 智能边缘小站、兆瀚 RA2311-C 推理服务器、兆瀚 RA5900-B 训练服务器等 3 款昇腾 AI 产品。预计 23、24 年摊薄 EPS 分别为 0.12、0.24 元，对应 127X、62X PE。
宝信软件	600845.SH	宝信软件系国内 IDC 龙头。公司多年专注于自主研发工业互联网平台宝联登 xln ³ Plat 和 AI 中台、xAI 套件等，并积极推进钢铁 AI 大模型研发，为智能制造发展奠定基石。预计 24、25 年摊薄 EPS 分别为 1.06、1.39 元，对应 46X、28XPE。
万国数据	9698.HK	万国数据系国内第三方 IDC 运营商市场的领头羊。公司核心业务为托管服务，并有托管管理服务和云管理两项增值服务。预计 23、24 年摊薄 EPS 分别为 -2.85、-0.90 元，对应 -3X、-6XPE。
奥飞数据	300738.SZ	奥飞数据系国内领先的 IDC 服务商。公司对内通过不断的资本投入加快自建数据中心，对外收购完善全国布局，卡位核心城市加速全国机柜部署。预计 23、24 年摊薄 EPS 分别为 0.20、0.30 元，对应 66X、44X PE。
数据港	603881.SH	数据港系新兴高耗能产业数据中心企业。公司在浸没式液冷、光伏发电、余热回收等方面的实践应用进行了一系列积极探索。相关技术有助于降低数据中心全生命周期能耗和碳排放，解决 IDC 选址的困境。预计 23、24 年摊薄 EPS 分别为 0.32、0.42 元，对应 61X、46XPE。
紫光股份	000938.SZ	紫光股份系全球新一代云计算基础设施建设和行业智能应用服务的领先者。公司致力于构建“云—网—算—存—端”全栈业务布局，进一步提出行业赋能战略“AlforALL”。预计 23、24 年摊薄 EPS 分别为 0.80、1.00 元，对应 27X、22XPE。
高新发展	000628.SZ	高新发展是功率半导体行业的新兴公司，2022 年通过并购整合森未科技和芯未半导体，正式进入功率半导体行业。公司主营业务为建筑业和功率半导体业务。功率半导体业务为公司战略转型确立的具备硬核技术的新主业，建筑业是公司目前第一大收入及利润来源，是公司打造强大功率半导体主业的坚实基础。公司 2023 年业绩快报对应 EPS 约为 1.05 元，对应 73XPE。
云赛智联	600602.SH	云赛智联定位于中立第三方高端数据中心运营商，数据中心建设是云赛智联“做优 IDC”战略路径的重要基石。公司主营业务是以云计算和大数据、行业解决方案及智能化产品为核心业务的专业化信息技术服务。预计 23、24 年摊薄 EPS 分别为 0.14、0.19 元，对应 82X、67XPE。
锐捷网络	301165.SZ	锐捷网络是领先的 ICT 基础设施及行业解决方案提供商，主营业务为网络设备、网络安全产品及云桌面解决方案的研发、设计和销售。公司一直致力于将技术与应用充分融合，贴近用户进行产品方案创新，助力各行业用户实现数字化转型和价值升级。预计 23、24 年摊薄 EPS 分别为 0.71、1.33 元，对应 54X、29X PE。
工业富联	601138.SH	工业富联是全球领先的通信网络设备、云服务设备、精密工具及工业机器人专业设计制造服务商，为客户提供以工业互联网平台为核心的新形态电子设备产品智能制造服务。预计 23、24 年摊薄 EPS 分别为 1.06、1.28 元，对应 14X、18XPE。
华勤技术	603296.SH	华勤技术是专业从事智能硬件产品的研发设计、生产制造和运营服务的平台型公司，隶属智能硬件 ODM 行业。预计 23、24 年摊薄 EPS 分别为 3.68、4.10 元，对应 23X、20XPE。

来源：iFind，国金证券研究所

注：上述公司的盈利预测及估值来自 iFind 一致预期。

7. 风险提示

■ 国产替代进程不及预期的风险。

国产替代已成为 AI 算力行业发展的重要方向。国产替代进程不及预期可能会带来一系列挑战。首先，技术积累和创新能力不足可能导致国产替代产品在性能、稳定性和可靠性上无法达到国际先进水平，从而限制了其在市场上的竞争力。其次，供应链的不稳定性和制造成本的上升可能影响到国产替代产品的生产和交付进度，进而影响到行业的供需关系和市场份额。最后，政策环境的变化和市场竞争的激烈程度也可能影响到国产替代进程的推进速度，进而影响到行业的结构调整和发展方向。

■ AI 应用落地不及预期的风险。

若相关应用公司不能找到人工智能算法较好的商业应用落地场景，或相关场景客户没有较强的付费意愿，可能算法应用落地会不及预期。



■ 技术变化较快，技术路线调整的风险。

随着技术的不断进步，企业原有的设备和系统可能无法满足新的需求，导致技术落后和竞争力下降。此外，企业还可能面临技术实施的风险。技术实施过程中可能出现各种问题，如系统故障、数据泄露等，给企业带来损失。

■ 数据中心能耗政策推进不及预期的风险。

数据中心是支撑 AI 算力行业的核心基础设施，然而其巨大的能耗却面临着环境和资源的压力。政府和监管机构的能耗政策对数据中心的运营和扩展产生着直接影响。如果数据中心能耗政策推进不及预期，可能导致行业面临更加严格的环境法规和能源限制，增加了行业的运营成本，并可能限制数据中心的规模和发展。

■ 行业竞争加剧的风险。

随着 AI 算力行业的迅速发展，竞争也在逐渐加剧。行业竞争加剧可能导致价格战和利润下降、行业内部出现分裂、混乱与垄断、同质化竞争过度等现象，从而影响到行业的竞争公平性和创新活力。



行业投资评级的说明:

买入: 预期未来 3-6 个月内该行业上涨幅度超过大盘在 15%以上;

增持: 预期未来 3-6 个月内该行业上涨幅度超过大盘在 5% - 15%;

中性: 预期未来 3-6 个月内该行业变动幅度相对大盘在 -5% - 5%;

减持: 预期未来 3-6 个月内该行业下跌幅度超过大盘在 5%以上。



特别声明:

国金证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

形式的复制、转发、转载、引用、修改、仿制、刊发，或以任何侵犯本公司版权的其他方式使用。经过书面授权的引用、刊发，需注明出处为“国金证券股份有限公司”，且不得对本报告进行任何有悖原意的删节和修改。

本报告的产生基于国金证券及其研究人员认为可信的公开资料或实地调研资料，但国金证券及其研究人员对这些信息的准确性和完整性不作任何保证。本报告反映撰写研究人员的不同设想、见解及分析方法，故本报告所载观点可能与其他类似研究报告的观点及市场实际情况不一致，国金证券不对使用本报告所包含的材料产生的任何直接或间接损失或与此有关的其他任何损失承担任何责任。且本报告中的资料、意见、预测均反映报告初次公开发布时的判断，在不作事先通知的情况下，可能会随时调整，亦可因使用不同假设和标准、采用不同观点和分析方法而与国金证券其它业务部门、单位或附属机构在制作类似的其他材料时所给出的意见不同或者相反。

本报告仅为参考之用，在任何地区均不应被视为买卖任何证券、金融工具的要约或要约邀请。本报告提及的任何证券或金融工具均可能含有重大的风险，可能不易变卖以及不适合所有投资者。本报告所提及的证券或金融工具的价格、价值及收益可能会受汇率影响而波动。过往的业绩并不能代表未来的表现。

客户应当考虑到国金证券存在可能影响本报告客观性的利益冲突，而不应视本报告为作出投资决策的唯一因素。证券研究报告是用于服务具备专业知识的投资者和投资顾问的专业产品，使用时必须经专业人士进行解读。国金证券建议获取报告人员应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。报告本身、报告中的信息或所表达意见也不构成投资、法律、会计或税务的最终操作建议，国金证券不就报告中的内容对最终操作建议做出任何担保，在任何时候均不构成对任何人的个人推荐。

在法律允许的情况下，国金证券的关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并可能为这些公司正在提供或争取提供多种金融服务。

本报告并非意图发送、发布给在当地法律或监管规则下不允许向其发送、发布该研究报告的人员。国金证券并不因收件人收到本报告而视其为国金证券的客户。本报告对于收件人而言属高度机密，只有符合条件的收件人才能使用。根据《证券期货投资者适当性管理办法》，本报告仅供国金证券股份有限公司客户中风险评级高于 C3 级(含 C3 级)的投资者使用；本报告所包含的观点及建议并未考虑个别客户的特殊状况、目标或需要，不应被视为对特定客户关于特定证券或金融工具的建议或策略。对于本报告中提及的任何证券或金融工具，本报告的收件人须保持自身的独立判断。使用国金证券研究报告进行投资，遭受任何损失，国金证券不承担相关法律责任。

若国金证券以外的任何机构或个人发送本报告，则由该机构或个人为此发送行为承担全部责任。本报告不构成国金证券向发送本报告机构或个人的收件人提供投资建议，国金证券不为此承担任何责任。

此报告仅限于中国境内使用。国金证券版权所有，保留一切权利。

上海
电话: 021-80234211
邮箱: researchsh@gjzq.com.cn
邮编: 201204
地址: 上海浦东新区芳甸路 1088 号
紫竹国际大厦 5 楼

北京
电话: 010-85950438
邮箱: researchbj@gjzq.com.cn
邮编: 100005
地址: 北京市东城区建内大街 26 号
新闻大厦 8 层南侧

深圳
电话: 0755-86695353
邮箱: researchsz@gjzq.com.cn
邮编: 518000
地址: 深圳市福田区金田路 2028 号皇岗商务中心
18 楼 1806



【小程序】
国金证券研究服务



【公众号】
国金证券研究