

CPU 行业深度系列报告一： AI agent 驱动需求增长，CPU 价格有望上调

优于大市

核心观点

CPU 为计算机核心，性能持续提升。CPU 是程序的调用者和运行者，程序的每一条指令都要经过 CPU 的解析和执行，外界向计算机输入数据，需要 CPU 进行接收；程序运行结束后，需要 CPU 发出指示才能把计算结果输出给外界。从 CPU 发展历史来看，CPU 的字长、晶体管数量、主频、核心数持续提升。

CPU 需求侧：AI Agent 驱动 CPU 需求增长。2025 年 CPU 全球市场规模为 747.37 亿美金，YoY +10%，其中，消费级 CPU 市场规模为 445.30 亿美元，YoY +9%，服务器 CPU 市场规模为 302.07 亿美元，YoY +12%。

CPU 需求驱动因素一：AI Headnode 配比逐步从 4:1 转为 2:1，24 年英伟达发布 GB200 超级芯片，其 CPU: GPU 配置提升至 1: 2，随着英伟达新一代芯片渗透率快速提升，AI Headnode 配比有望逐步从 4:1 转为 2:1。
CPU 需求驱动因素二：AI 大模型逐步从 ChatBot 转向 Agent 工作流，CPU 负责编排子智能体之间的关系，同时，每个子智能体调用的外部资源（例如数据库、搜索引擎等）亦需要 CPU 支撑，拉动 CPU 需求开始增长。2026 年 5 月，AMD 召开财报电话会议，AMD CEO 苏姿丰上调 2030 年全球服务器 CPU 市场规模为 1200 亿，对应年增长率提升至 35%+。

CPU 供给侧：x86 架构为主，ARM 架构占比持续提升。1) **服务器 CPU：**2025 年 x86 架构、ARM 架构占比分别为 87.0%、10.0%，x86 架构为主，ARM 占比持续提升；英特尔、AMD 占据主要市场份额，AWS、英伟达自研 CPU 快速放量。2) **消费级 CPU：**电脑端 CPU 主要包括台式机 CPU 和笔记本电脑 CPU，英特尔、AMD 占据主要市场份额；平板 CPU 考虑低功耗、长续航等特点，ARM 架构 CPU 占据主要市场份额，苹果、高通、联发科为全球龙头。

全球 CPU 价格上涨，CPU 国产替代有望加速。2026 年全球 CPU 供不应求，英特尔、AMD 公司 CPU 价格上涨。国产 CPU 公司海光信息、海思鲲鹏等国产替代有望加速。

投资建议：AI Agent 拉动 CPU 需求，旨在支撑 Agent 任务编排、沙盒、工具运行等，根据 IDC 披露数据，预计 2030 年全球 CPU 市场规模增长至 1307.8 亿美元，对应 25-30 年 CAGR 为 34.1%；国内 CPU 市场亦在快速增长，国产 CPU 市占率持续提升，建议关注海光信息。

风险提示：CPU 需求不及预期风险、CPU 价格下跌风险、互联网大厂自研 CPU 替代第三方 CPU 风险。

行业研究 · 行业专题

计算机

优于大市 · 维持

证券分析师：熊莉

021-61761067

xiongli1@guosen.com.cn

S0980519030002

证券分析师：艾宪

0755-22941051

aixian@guosen.com.cn

S0980524090001

市场走势



资料来源：Wind、国信证券经济研究所整理

相关研究报告

- 《AI 云系列研究：AI 重塑云计算增长曲线，阿里云有望进入回报提升阶段》——2026-08-04
- 《人工智能行业专题：2026Q2 海外大厂 CapEx 和 ROIC 总结梳理》——2026-08-01
- 《AI 周报：多重因素共同推升租赁价格，高端 GPU 供不应求》——2026-06-28
- 《华为韬（T）定律：从“面积缩微”转向“时间缩微”的半导体新范式》——2026-05-28
- 《人工智能行业专题：2026Q1 海外大厂 CapEx 和 ROIC 总结梳理》——2026-05-06

重点公司盈利预测及投资评级

公司代码	公司名称	投资评级	昨收盘 (元)	总市值 (百万元)	EPS		PE	
					2026E	2027E	2026E	2027E
688041	海光信息	优于大市	213.50	496,246	1.83	2.90	116.7	73.6

资料来源：Wind、国信证券经济研究所预测

内容目录

一、 CPU 为计算机核心，性能持续提升	5
二、 CPU 需求侧：AI Agent 驱动 CPU 需求增长	7
三、 CPU 供给侧：x86 架构为主，ARM 架构占比持续提升	13
四、 全球 CPU 价格上涨，CPU 国产替代有望加速	17
投资建议	19
风险提示	19

图表目录

图 1: 英特尔 CPU 产品迭代历史	5
图 2: 数据中心 CPU 核心数快速提升	6
图 3: 冯·诺依曼体系结构	7
图 4: CPU 架构组成	7
图 5: 全球 CPU 市场规模 (单位: 亿美元)	7
图 6: 消费级、服务器 CPU 占比 (收入)	7
图 7: 全球 CPU 出货量 (单位: 百万颗)	8
图 8: 消费级、服务器 CPU 占比 (出货量)	8
图 9: 消费级、服务器级 CPU ASP 情况 (单位: 美元/颗)	8
图 10: 消费级 CPU 市场规模 (单位: 亿美元)	9
图 11: 消费级 CPU 出货量 (单位: 百万颗)	9
图 12: 消费级 CPU 细分领域占比 (收入)	9
图 13: 消费级 CPU 细分领域占比 (出货量)	9
图 14: 消费级 CPU 细分领域占比 (收入)	10
图 15: 消费级 CPU 细分领域占比 (出货量)	10
图 16: H200 系统, CPU: GPU=1:4	10
图 17: GB200 NVL72 系统, CPU: GPU=1:2	10
图 18: Agent 工作流	11
图 19: CPU 在不同任务中运行时间 (延迟) 情况 (阴影部分表示 CPU 使用)	11
图 20: Agentic AI 工作流示意图	12
图 21: 流水线不同任务时间	12
图 22: 代码生成工作流示意图	12
图 23: 代码生成不同任务时间	12
图 24: CPU 配比有望提升至 1:1	13
图 25: AI HeadNode 占比持续提升	13
图 26: 服务器 CPU 市场规模预测 (单位: 亿美元)	13
图 27: 服务器 CPU 出货量预测 (单位: 百万颗)	13
图 28: 指令集划分 CPU 类型	14
图 29: 全球不同指令集 CPU 厂商	14
图 30: 全球服务器 CPU 架构占比 (收入维度)	14
图 31: 全球服务器 CPU 架构占比 (出货量)	14
图 32: 全球服务器 CPU 竞争格局 (收入维度)	15
图 33: 全球服务器 CPU 竞争格局 (出货量)	15
图 34: 中国服务器 CPU 市场竞争格局	15
图 35: 全球台式机 CPU 竞争格局 (收入维度)	16
图 36: 全球台式机 CPU 竞争格局 (出货量)	16
图 37: 全球笔电 CPU 竞争格局 (收入维度)	16

图 38: 全球笔电 CPU 竞争格局 (出货量)	16
图 39: 全球平板 CPU 竞争格局 (收入维度)	17
图 40: 全球平板 CPU 竞争格局 (出货量)	17
图 41: Intel i7 Alder Lake 散货价 (美元)	17
图 42: 英特尔服务器 CPU 出货量 (百万颗)	17
图 43: AMD Ryzen7 Zen4 散货价 (美元)	18
图 44: AMD 服务器 CPU 出货量 (百万颗)	18
图 45: 海光信息产品矩阵	18
图 46: 海光营收 (亿元) 及增速	18

一、CPU 为计算机核心，性能持续提升

CPU 产品发展趋势：以全球 CPU 龙头 Intel 产品发展历史来看，CPU 的字长、晶体管数量、主频、核心数持续提升；在 AI Agent 阶段，CPU 将在核心数、主频、能耗比等维度进一步提升。

开端：1971 年 Intel 发布 Intel 4004，为商用 CPU 的开端；

百花齐放阶段（1971 年-2000 年）：个人计算机、家庭娱乐设备、工业控制领域对 CPU 提出大量需求，多家 CPU 厂商并存，其中 Intel 8008、MOS 6502、Motorola MC6800、Zilog Z80 是领导者，Intel 通过深度绑定 IBM、微软，逐步成长为 CPU 龙头，同时，IBM 要求 Intel 把 x86 指令集开放给 AMD，AMD 开始快速发展；

多核革命阶段（2001 年-2010 年）：这一时期主要参与者为 Intel、AMD，单核主频提升速度放缓（物理限制导致漏电流和发热急剧增减），2005 年 AMD Athlon 64 X2 发布，首款主流双核处理器，通过并行计算提升性能；同年，Intel 发布 Pentium Extreme Edition 840，双核心，主频 3.2GHz，支持超线程技术；

服务器 CPU 和移动端 CPU 崛起（2011 年-2022 年）：2010 年美国联邦政府制定了“云优先”的发展战略，云计算开始蓬勃发展，服务器 CPU 出货量快速增长；同时，智能移动终端开始放量，ARM 架构的崛起让 x86 架构无法延伸到移动端，高通、苹果、华为成为移动端 CPU 的领导者；

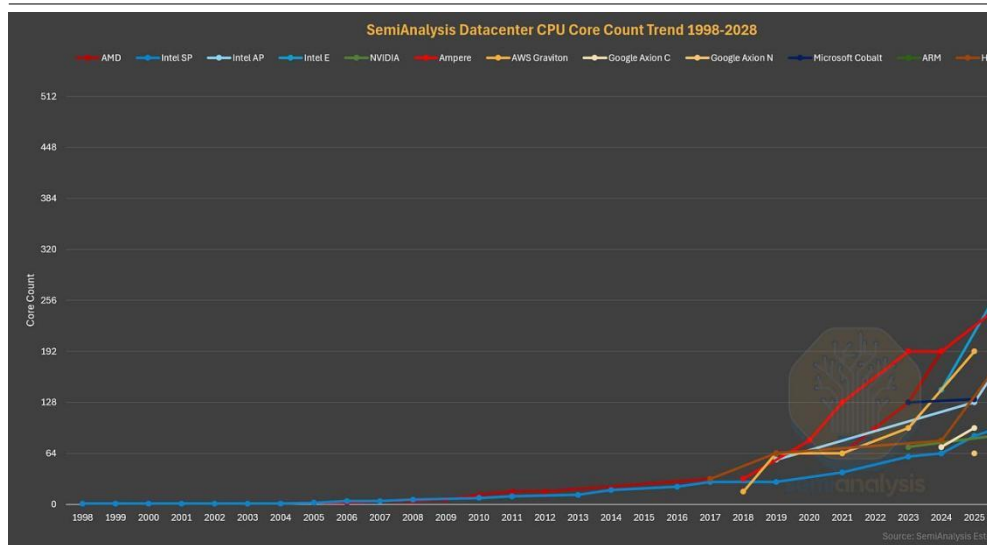
AI 时代拉动 CPU 需求（2023 年-至今）：从服务器 CPU 端，受 AI 大模型训练和推理需求驱动，互联网大厂纷纷加大资本开支，采购 AI 服务器，AI 服务器用 CPU 需求快速增长，随着 AI 推理从 ChatBot 转向 AI Agent，Agentic CPU 有望成为 CPU 需求增长的最大推手；从消费级 CPU 端，个人 PC 将承担部分本地推理任务，消费级 CPU 开始整合 NPU，提升产品整体算力，价值量有望进一步提升。

图1：英特尔 CPU 产品迭代历史

发展阶段	年代	字长	代表芯片	内核个数	晶体管个数	主频
第一代	1971 年	4 位	Intel 4004	1	2 300	108kHz
	1972 年	8 位	Intel 8008	1	3 500	200kHz
第二代	1974 年	8 位	Intel 8080	1	6 000	2MHz
第三代	1978 年	16 位	Intel 8086 Intel 8087 Intel 8088	1	29 000	5/8/10MHz
	1982 年	16 位	Intel 80286	1	14.3 万	6/8/10/12.5MHz
第四代	1985 年	32 位	Intel 80386	1	27.5 万	16MHz
	1989 年	32 位	Intel 80486	1	125 万	25MHz~100MHz
第五代	1994—2005 年	32/64 位	Intel 奔腾系列	1/2	3 210 万~5 500 万	66MHz~3.4GHz
第六代	2005 年至今	64 位	Intel 酷睿系列	2/4/8/16	几亿	

资料来源：宁爱军、曹鉴华著-《信息与智能科学导论》-人民邮电出版社（2019 年）-P28，国信证券经济研究所整理

图2：数据中心 CPU 核心数快速提升



资料来源：SemiAnalysis，国信证券经济研究所整理

CPU=运算器+控制器。按冯·诺依曼体系结构，计算机由运算器、控制器、存储器、输入设备、输出设备共 5 个部分构成，其中运算器、控制器两部分经常被一起设计，二者合称中央处理器（CPU），主要任务为由控制器指挥计算机中的其他部件一起协同工作，并由运算器执行数值运算，主要工作原理如下：1）5 个部分之间有通信线路进行连接；2）程序以连续排列的指令为单位组成（每条指令包含一项计算操作），要运行的任务（程序）保存在存储器中；3）启动后，控制器从存储器中依次读取指令，将指令包含的信息传送到运算器，由运算器解析指令的功能、执行数值运算；4）指令执行时，有可能需要从外界读取要加工的数据；控制器向输入设备发送信息，输入设备把数据传送到计算机内部；5）指令执行结束后，控制器可以把运算器的计算结果存入存储器；控制器也可以向输出设备发送信息，由输出设备把计算结果传送到计算机外部；6）控制器再从存储器中读取下一条指令并送入运算器；7）上述“取指令-执行指令-保存结果”的过程多次重复执行，直到程序的最后一条指令执行完成。

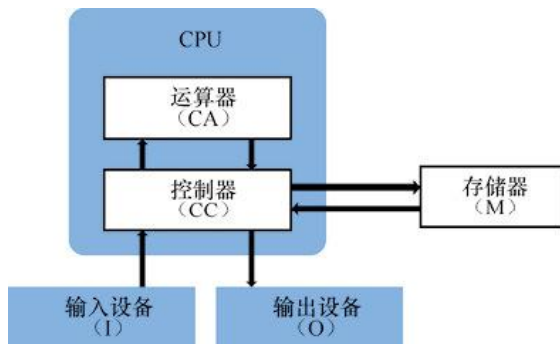
CPU 为计算机的核心，指挥计算机中的其它部件工作。CPU 是程序的调用者和运行者，程序的每一条指令都要经过 CPU 的解析和执行，外界向计算机输入数据，需要 CPU 进行接收；程序运行结束后，需要 CPU 发出指示才能把计算结果输出给外界。CPU 主要由控制单元（Control Unit）、算术逻辑单元（ALU）、寄存器集（Register Set）组成，具体功能如下：

控制单元（Control Unit）：负责管理处理器的整体操作，主要包括从内存中读取下一条指令，通过指令译码确定要执行的操作，以及将要执行的指令分发到处理器中的合适单元；

算术逻辑单元（ALU）：执行算术运算和位操作运算的组合电路；

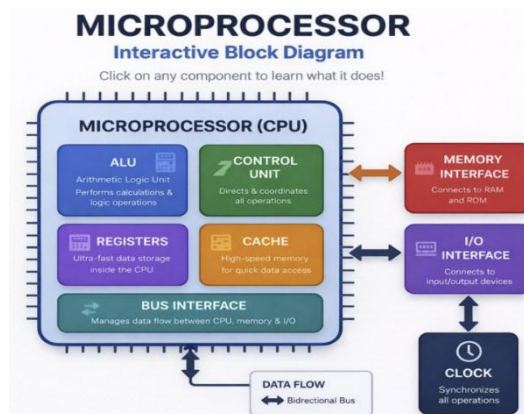
寄存器集（Register Set）：提供临时存储以及指令输入和输出的源位置与目标位置。

图3：冯·诺依曼体系结构



资料来源：张戈著-《CPU 通识课》-人民邮电出版社（2022 年）-P18，国信证券经济研究所整理

图4：CPU 架构组成



资料来源：Testbook，国信证券经济研究所整理

二、CPU 需求侧：AI Agent 驱动 CPU 需求增长

CPU 分类：CPU 按下游应用领域可分为消费级 CPU 和服务器 CPU，其中消费级 CPU 主要应用于台式机、笔记本电脑、平板（此处不考虑手机端 SOC），服务器 CPU 主要应用于通用服务器和 AI 服务器等。

市场规模：2025 年 CPU 全球市场规模为 747.37 亿美金，同比+10%，对应 21-25 年 CAGR 为 2.0%，其中消费级 CPU 市场规模为 445.30 亿美元，同比+9%，对应 21-25 年 CAGR 为 1.2%，占比为 59.6%；服务器 CPU 市场规模为 302.07 亿美元，同比+12%，对应 21-25 年 CAGR 为 3.4%，占比为 40.4%。2026Q1 全球 CPU 市场规模为 203.04 亿美金，同比+17%，占比为 52.5%；服务器 CPU 市场规模为 96.70 亿美元，同比+45%，占比为 47.5%。

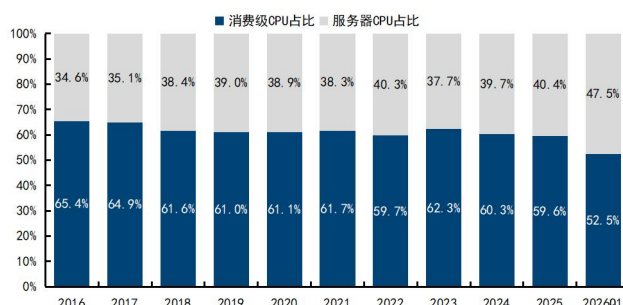
出货量：2025 年 CPU 全球出货量为 4.72 亿颗，同比+4%，对应 21-25 年 CAGR 为 -5.3%，其中消费级 CPU 出货量为 4.39 亿颗，同比+3%，对应 21-25 年 CAGR 为 -5.4%，占比为 93.0%；服务器 CPU 出货量为 0.33 亿颗，同比+25%，对应 21-25 年 CAGR 为 -3.2%，占比为 7.0%。2026Q1 全球 CPU 出货量为 1.11 亿颗，同比+2%，其中消费级 CPU 出货量为 1.02 亿颗，同比+1%，占比为 91.5%；服务器 CPU 出货量为 0.09 亿颗，同比+17%，占比为 8.5%。

图5：全球 CPU 市场规模（单位：亿美元）



资料来源：IDC，国信证券经济研究所整理

图6：消费级、服务器 CPU 占比（收入）



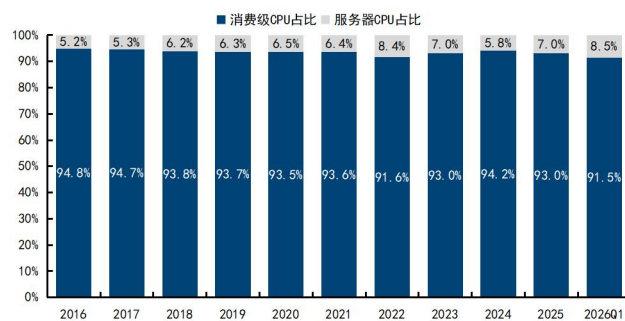
资料来源：IDC，国信证券经济研究所整理

图7: 全球 CPU 出货量（单位：百万颗）



资料来源：IDC，国信证券经济研究所整理

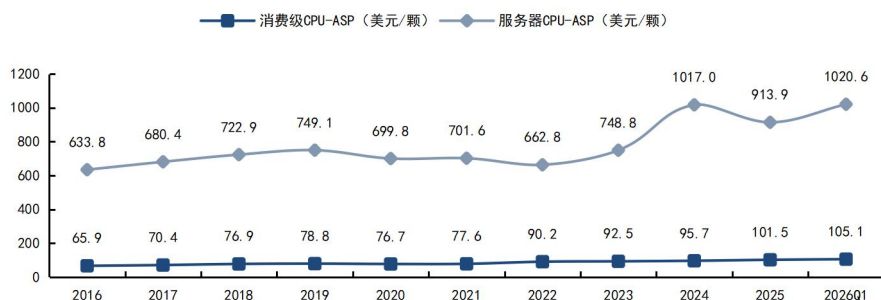
图8: 消费级、服务器 CPU 占比（出货量）



资料来源：IDC，国信证券经济研究所整理

ASP 趋势向上, 服务器级 CPU 价格显著高于消费级 CPU。根据 IDC 披露数据, 2026Q1 年消费级 CPU、服务器级 CPU 分别为 105.1、1020.6 美元/颗, 服务器 CPU 的 ASP 显著高于消费级 CPU, 主要由于服务器 CPU 核心数更多, 单核频率更高; 从历史维度来看, 消费级 CPU ASP 持续提升, 未来受益于 AI PC 渗透率加速, ASP 有望进一步提升, 服务器 CPU ASP 在 2024 年提升明显, 主要由于高核心数 CPU 销售占比提升, 拉动整体 ASP 提升, 2026Q1 服务器 CPU 受益于新一代架构 CPU 放量, AI 需求拉动超高核心数 CPU 需求, ASP 创下新高。

图9: 消费级、服务器级 CPU ASP 情况（单位：美元/颗）



资料来源：IDC，国信证券经济研究所整理

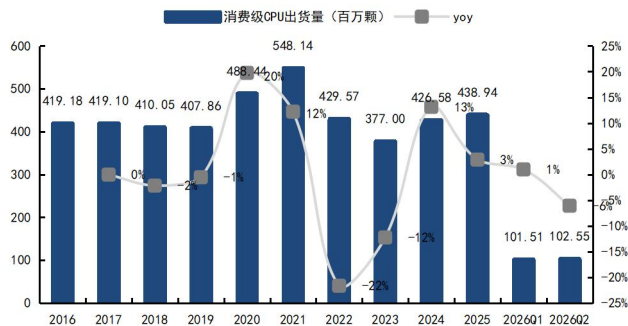
消费级 CPU：主要应用于台式机、笔记本电脑、平板等领域（此外不考虑手机端 SOC）；2025 年消费级市场规模为 445.30 亿美元，同比+9%，对应 21-25 年 CAGR 为 1.2%，其中，台式机、笔记本电脑、平板用 CPU 市场规模分别为 145.73、267.45、32.12 亿美元，分别同比+18%、+5%、+5%，占比分别为 32.7%、60.1%、7.2%。从出货量来看，消费级 CPU 出货量为 4.39 亿颗，同比+3%，对应 21-25 年 CAGR 为 -5.4%，其中，台式机、笔记本电脑、平板用 CPU 出货量分别为 0.81、2.10、1.48 亿颗，占比分别为 18.3%、47.9%、33.8%。2026Q2 消费级市场规模为 121.27 亿美元，同比+15%，其中，台式机、笔记本电脑、平板用 CPU 市场规模分别为 33.79、80.08、7.40 亿美元，分别同比-1%、+27%、-6%，占比分别为 27.9%、66.0%、6.1%。从出货量来看，2026Q2 消费级 CPU 出货量为 1.03 亿颗，同比-6%，其中，台式机、笔记本电脑、平板用 CPU 出货量分别为 0.16、0.53、0.33 亿颗，占比分别为 16.0%、51.9%、32.1%。

图10: 消费级 CPU 市场规模（单位：亿美元）



资料来源：IDC，国信证券经济研究所整理

图11: 消费级 CPU 出货量（单位：百万颗）



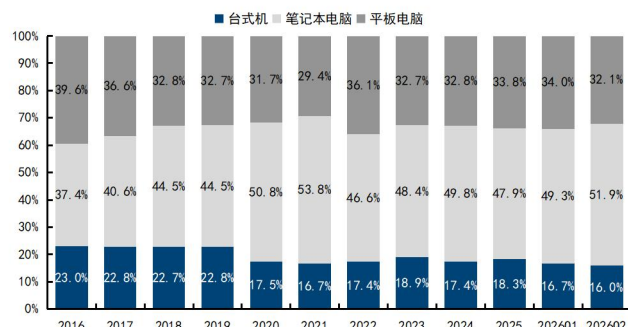
资料来源：IDC，国信证券经济研究所整理

图12: 消费级 CPU 细分领域占比（收入）



资料来源：IDC，国信证券经济研究所整理

图13: 消费级 CPU 细分领域占比（出货量）



资料来源：IDC，国信证券经济研究所整理

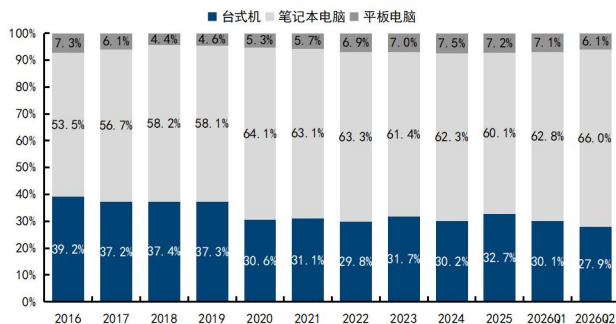
服务器 CPU：主要用于通用服务器、AI 服务器，服务器 CPU 市场规模为 302.07 亿美元，同比+12%，对应 21-25 年 CAGR 为 3.4%；服务器 CPU 出货量为 0.33 亿颗，同比+25%，对应 21-25 年 CAGR 为-3.2%。2026Q1 服务器 CPU 市场规模为 96.70 亿美元，同比+45%；服务器 CPU 出货量为 0.09 亿颗，同比+17%。

a) 2010 年，美国联邦政府制定了“云优先”的发展战略，为政府提供了广泛采用基于云解决方案的权利，并积极鼓励其下各部门积极采用私营机构云服务以提升政府决策能力，同时白宫在预算申请文件中将云计算列为促进美国政府技术设施的重要技术，云计算开始进入快速发展期，海外云厂商（例如谷歌、微软、亚马逊等）加大资本开支，用于通用服务器采购，服务器 CPU 开始快速放量；

b) 2023 年，在 AI 大模型训练和推理需求驱动下，海外云厂商加大资本开支，采购 AI 服务器，AI 服务器机柜内部的 CPU（Head Node）开始放量；

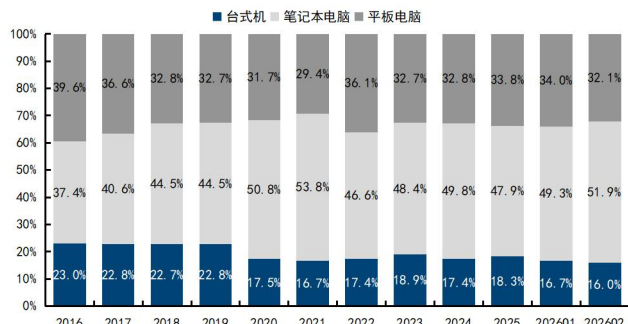
c) 2026 年，AI 大模型推理开始从 ChatBot 转向 Agent，Agent 的任务编排、工具调用、沙盒等需要 CPU 支持，Agentic CPU 成为需求增长的主要动力。

图14: 消费级 CPU 细分领域占比（收入）



资料来源：IDC，国信证券经济研究所整理

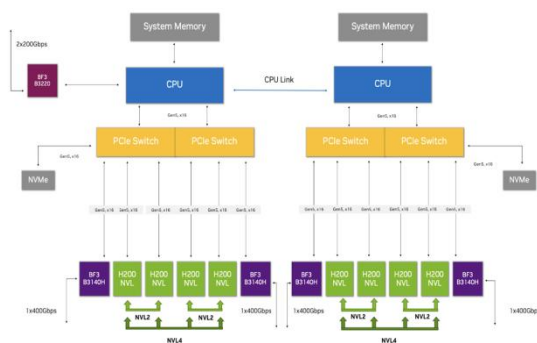
图15: 消费级 CPU 细分领域占比（出货量）



资料来源：IDC，国信证券经济研究所整理

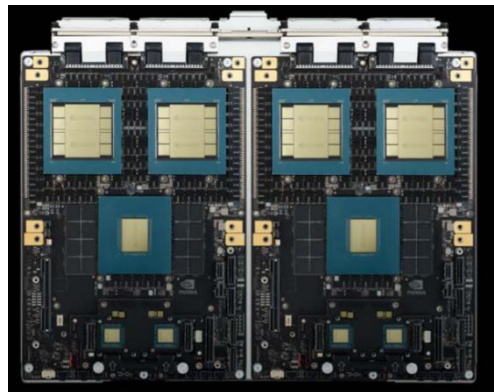
CPU 需求驱动因素一：AI Headnode 配比逐步从 4:1 转为 2:1，拉动 CPU 需求增长。 AI Headnode CPU 主要负责管理 GPU，并持续为其提供数据，具备大容量缓存、高带宽内存、高单核性能等特点。过去单个计算节点（单台服务器）通常为 2 颗 CPU 配 8 颗 GPU，CPU: GPU 配比为 1:4；2024 年 3 月，英伟达发布 GB200 超级芯片，其 CPU: GPU 配置提升至 1: 2，且后续发布的 GB300、Rubin 等芯片，仍采取 1:2 的配比，随着英伟达新一代芯片渗透率快速提升，AI Headnode 配比有望逐步从 4:1 转为 2:1。

图16: H200 系统，CPU: GPU=1:4



资料来源：英伟达，国信证券经济研究所整理

图17: GB200 NVL72 系统，CPU: GPU=1:2



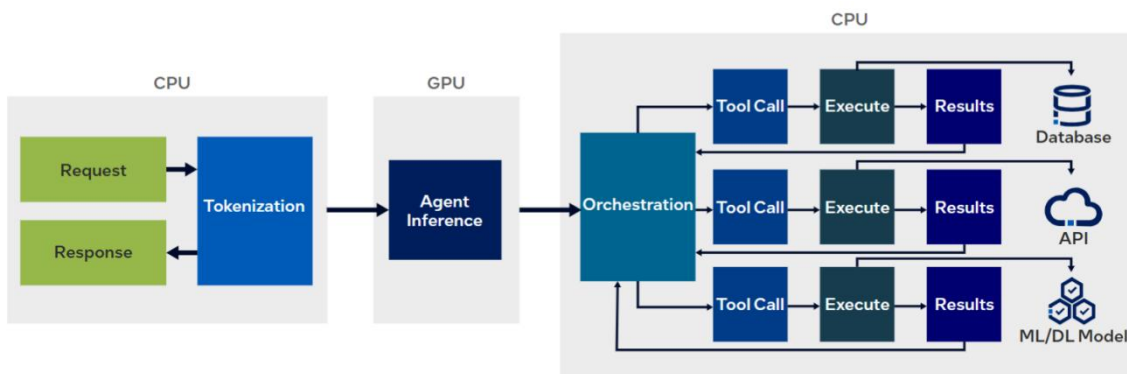
资料来源：英伟达，国信证券经济研究所整理

CPU 需求驱动因素二：Agentic CPU 需求快速提升。 AI 大模型推理模式，逐步从单次推理（以 GPU 为主）转向 Agent 工作流。1) **单次推理阶段**：用户发送请求，CPU 对其进行 Token 化处理，将结果传递给 GPU 进行推理，GPU 推理结束后，CPU 对返回结果进行去 Token 化处理，将最终结果反馈给用户；在单次推理中，CPU 仅限于将请求发送给 GPU，并将结果返回给用户，计算密集型工作（例如矩阵乘法、注意力机制计算、Token 生成）全部在 GPU 上完成，所以 CPU 负载较轻，不会成为推理瓶颈。2) **Agent 工作流阶段**：在编排工作流中，CPU 从辅助角色转变为指挥层，用户发送请求，CPU 对其进行 Token 化处理，进而生成一个执行计划，随着该计划由多个并行的子智能体（sub-agents）执行，每个子智能体调用不同的工具，并等待各自的完成时间。CPU 负责编排子智能体之间的关系（例如独立关系、结果依附关系、完成顺序等），同时，每个子智能体调用的外部资源（例如数据库、搜索引擎、向量库、API 等）亦需要 CPU 支撑，当所有子智能体完成任务，CPU 会整合其结果发送到反思推理循环（Reflection Inference Loop）中，

评估原始请求是否得到满足，如果没有，则会开启新一轮规划、推理和工具执行循环，直到结果满足用户需求，最后 CPU 会进行去 Token 化并返回用户。

基于以上 Agent workflow 特征，CPU 将向以下三个维度迭代升级：1) 高主频：能以极低的延迟监控每一个 Token，判断子智能体何时完成任务，并可根据需要将这些结果整合进其他子智能体；2) 高核心数：多个子智能体可以并行运行，同时处理多项任务；3) 低延迟访问内存（上下文和中间状态）：缓存设计、存储层级设计愈发关键。

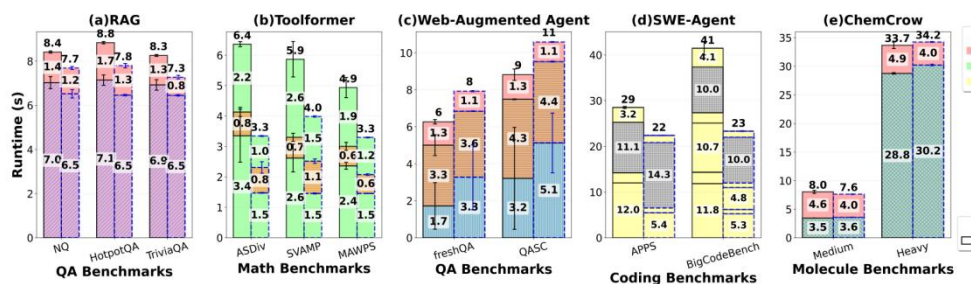
图18: Agent workflow



资料来源：英特尔，国信证券经济研究所整理

CPU 成为 Agent workflow 延迟瓶颈，拉动 CPU 需求增长。根据《Towards Understanding, Analyzing, and Optimizing Agentic AI Execution: A CPU-Centric Perspective》披露数据，作者对五种代表性 Agent 工作负载进行测试，在 RAG、Web-Augmented Agent、SWE-Agent、ChemCrow 四个 Agent 工作负载中，CPU 承担的工具任务占据主要的延迟时长，例如在 RAG 测试中，CPU 承载的 ENNS 检索任务占到总延迟的 80%-90%，在 Web-Augmented，CPU 承载的 URL 获取和 LexRank 摘要任务占到总延迟的 80%-90%，CPU 成为 Agent workflow 延迟瓶颈，数量更多、性能更强大的 CPU 是降低 Agent workflow 延迟的关键。

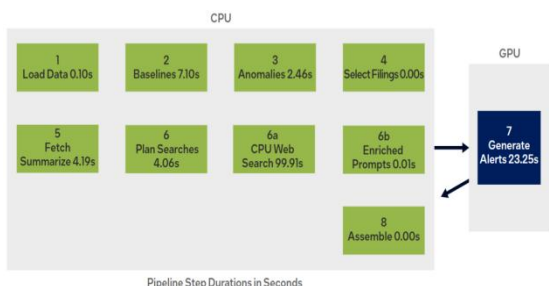
图19: CPU 在不同任务中运行时间（延迟）情况（阴影部分表示 CPU 使用）



资料来源：Ritik Raj 等著《Towards Understanding, Analyzing, and Optimizing Agentic AI Execution: A CPU-Centric Perspective》-Arxiv (2025 年) -P4，国信证券经济研究所整理

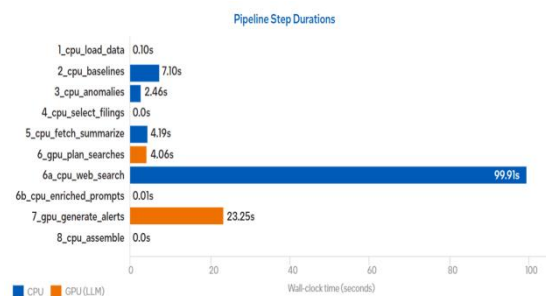
通用 Agent AI workflows 中的 CPU 延迟：根据 Intel 披露数据，基于一个涉及 SEC 监管文件的真实金融服务 workflow，构建一个异常检测的 workflow。首先，CPU 负责执行基础工作，包括加载数据、建立每家公司的典型申报模式、标记偏离模式公司、下载实际申报文档，并将多页文档压缩成简短摘要，随后 Agent 代用一个工具来计算分值，并将其与预设阈值进行比较。此外，为了模拟通过额外材料扩充数据集的 workflow，加入网络搜索增强步骤，结果显示，CPU 消耗了整个流水线的大部分时间。

图20: Agentic AI workflow 示意图



资料来源：英特尔，国信证券经济研究所整理

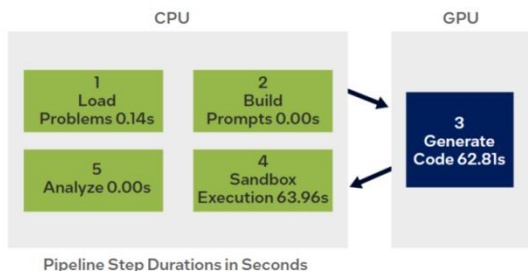
图21: 流水线不同任务时间



资料来源：英特尔，国信证券经济研究所整理

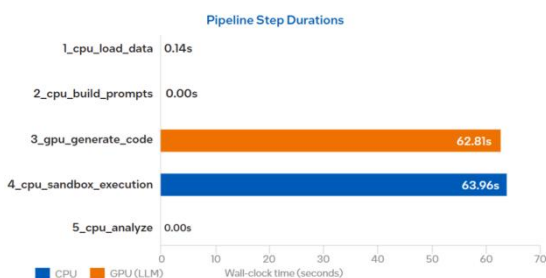
编程场景下 CPU 延迟：根据 Intel 披露数据，模拟开发人员要求 AI 助手编写和测试代码情况，首先，GPU 上的 AI 大模型生成候选代码解决方案，其次，CPU 执行并验证代码，实际过程中会使用容器/虚拟机，增加额外的 CPU 开销，结果显示，CPU 仍成为编程 Agent 性能瓶颈，其中沙箱执行阶段时间占比超过任务总时长的 50%。

图22: 代码生成 workflow 示意图



资料来源：英特尔，国信证券经济研究所整理

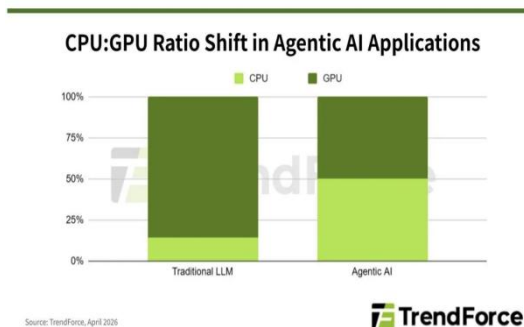
图23: 代码生成不同任务时间



资料来源：英特尔，国信证券经济研究所整理

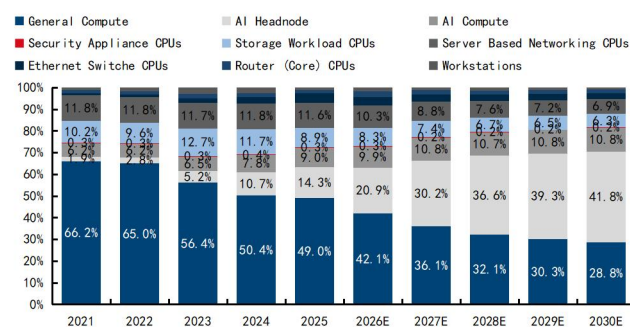
CPU 配比有望提升至 1:1，HeadNode CPU 占比持续提升。根据 Trendforce 披露数据，传统 AI 数据中心每 GW 需要约 3000 万个 CPU 核心，而 AI Agent 时代，这一需求将提升至每 GW 1.2 亿个 CPU 核心，未来 CPU: GPU 配比比率有望提升至 1:1 到 1:2 之间。根据 IDC 披露数据，在服务器 CPU 中，AI HeadNode CPU 占比持续提升，从出货量维度来看，2024 年 AI HeadNode 出货量占比约 10.7%，预计 2030 年提升至 41.8%。

图24: CPU 配比有望提升至 1:1



资料来源: Trendforce, 国信证券经济研究所整理

图25: AI HeadNode 占比持续提升



资料来源: IDC, 国信证券经济研究所整理

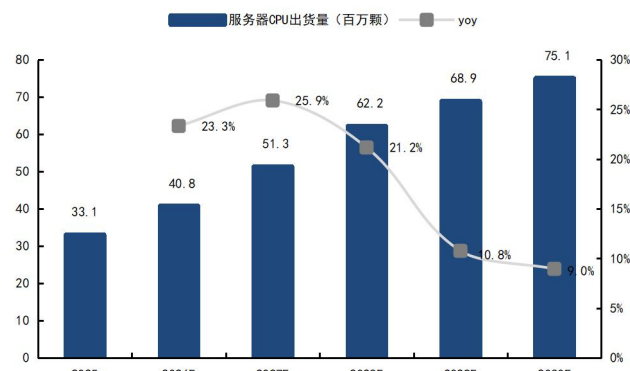
AMD 大幅上调服务器 CPU 市场空间，预计年化增速超 35%。根据 IDC 预测数据，2025 年全球服务器 CPU 市场规模为 302.1 亿美元，预计 2030 年增长至 1307.8 亿美元，对应 25-30 年 CAGR 为 34.1%；2025 年全球服务器 CPU 出货量为 0.33 亿颗，预计 2030 年增长至 0.75 亿颗，对应 25-30 年 CAGR 为 17.8%。2026 年 5 月，AMD 召开财报电话会议，AMD CEO 苏姿丰上调 2030 年全球服务器 CPU 市场规模为 1200 亿，对应年增长率提升至 35%+，主要由 AI Agent 对 CPU 需求拉动。

图26: 服务器 CPU 市场规模预测（单位：亿美元）



资料来源: Trendforce, 国信证券经济研究所整理

图27: 服务器 CPU 出货量预测（单位：百万颗）



资料来源: IDC, 国信证券经济研究所整理

三、CPU 供给侧：x86 架构为主，ARM 架构占比持续提升

指令集划分 CPU 类型，目前以 x86 和 ARM 架构为主。1) 指令集：指令集架构是指 CPU 用来完成计算和控制的一套指令的集合，主要规定了指令格式、寻址访存（包括寻址范围、寻址模式、寻址粒度、访存方式、地址对齐等）、数据类型和寄存器，是连接上层软件和底层硬件的桥梁。2) 指令集分类：指令集架构一般可以分为复杂指令集和精简指令集两大类，复杂指令集以 x86 架构为代表，主要公司有英特尔、AMD、海光信息等，精简指令集有 ARM、RISC-V、MIPS、PowerPC 等架构，目前 ARM 架构为其代表，主要公司有高通、三星、苹果、华为等。

图28: 指令集划分 CPU 类型

复杂指令集	精简指令集
复杂的指令集架构, 原生的处理器指令集架构	精简的指令集架构, 产生于20世纪80年代初
指令执行周期需要几个时钟周期	一个时钟周期执行一条指令
基于硬件的设计, 使用硬件电路来完成更多指令	基于软件的设计, 编译器承担了更多的任务, 用更少的指令编写更高效的软件
指令对RAM需求量少	指令对RAM需求量大
指令长度不固定	指令长度等长
指令数量多	指令数量少
程序源代码长度较短	程序源代码长度较长
控制器实现方式绝大多数为微程序	控制器实现方式绝大多数为硬件电路

资料来源: 王晓云等著-《算力时代: 一场新的产业革命》-中信出版社(2022年)-P108-118, 国信证券经济研究所整理

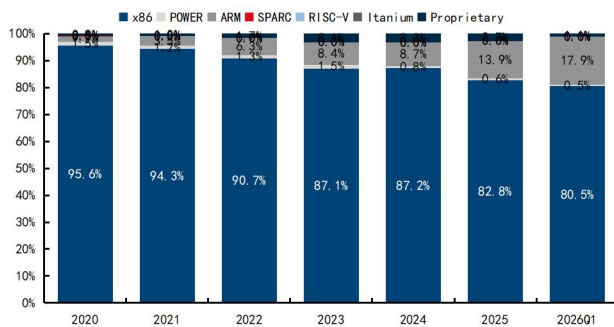
图29: 全球不同指令集 CPU 厂商

指令集	CISC (复杂指令集)	RISC (精简指令集)
产品架构	X86	ARM MIPS RISC-V Power PC Alpha
国际厂商 (列举)	英特尔 AMD	ARM 高通 三星 苹果 MIPS Microsemi IBM /
中国厂商 (列举)	海光 兆芯	飞腾 华为 展讯 龙芯 北京君正 华米 苏州国芯 申威

资料来源: 艾瑞咨询, 国信证券经济研究所整理

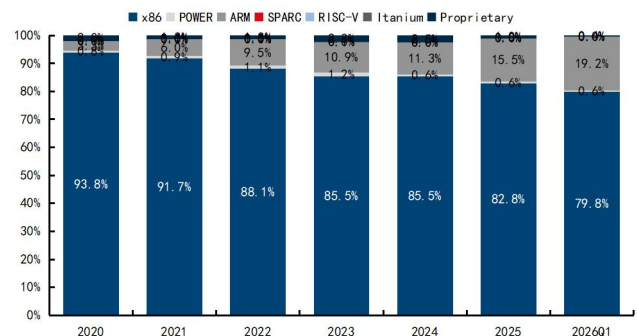
服务器 CPU: x86 架构为主, ARM 占比持续提升。根据 IDC 披露数据, 从收入维度来看, 2025 年 x86 架构服务器 CPU 营收 250.1 亿美金, 全球占比 82.8%, 占据主要市场份额, ARM 架构服务器 CPU 营收 42.1 亿美元, 全球占比 13.9%, 相比于 2020 年提升 12.0 个 pct, 占比快速提升; 从出货量维度来看, 2025 年 x86 架构服务器 CPU 全球出货 2738 万颗, 全球占比 82.8%, ARM 架构服务器 CPU 全球出货 513 万颗, 全球占比 15.5%。2026Q1, x86 架构服务器 CPU 营收 77.8 亿美元, 全球占比 80.5%; ARM 架构服务器 CPU 营收 17.3 亿美元, 全球占比 17.9%。从出货量来看, 2026Q1 x86 架构服务器 CPU 全球出货 756 万颗, 全球占比 79.8%; ARM 架构服务器 CPU 全球出货 182 万颗, 全球占比 19.2%。

图30: 全球服务器 CPU 架构占比 (收入维度)



资料来源: IDC, 国信证券经济研究所整理

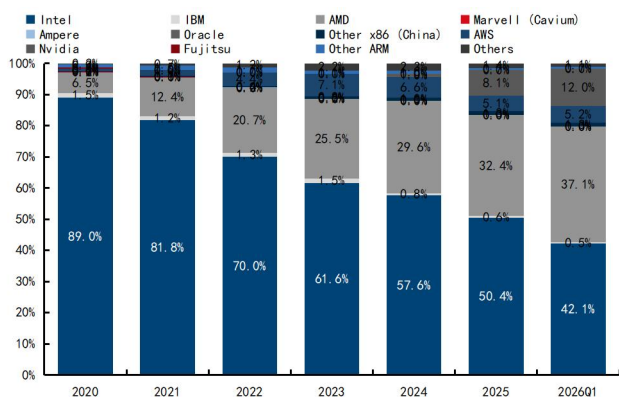
图31: 全球服务器 CPU 架构占比 (出货量)



资料来源: IDC, 国信证券经济研究所整理

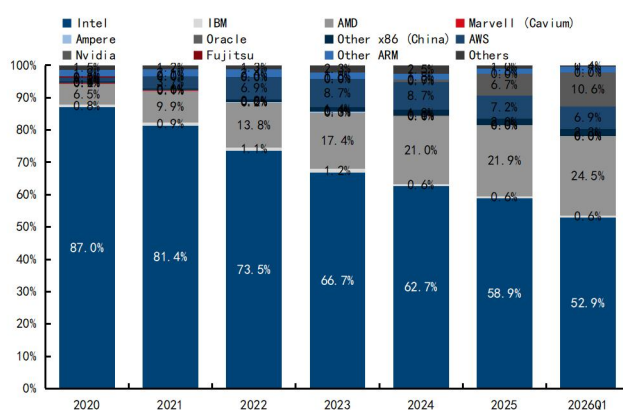
英特尔、AMD 占据主要服务器 CPU 市场份额, AWS、英伟达自研 CPU 快速放量。根据 IDC 披露数据, 从收入维度来看, 2025 年英特尔、AMD 市占率分别为 50.4%、32.4%, 合计占比 82.8%, 占据主要市场份额, 此外, AWS、英伟达自研 CPU 市占率分别为 5.1%、8.1%; 从出货量维度来看, 2025 年英特尔、AMD 市占率分别为 58.9%、21.9%, AWS、英伟达自研 CPU 市占率分别为 6.7%、7.2%, 大厂自研 CPU 占比快速提升。2026Q1, 从收入维度来看, 英特尔、AMD 市占率分别为 42.1%、37.1%, 合计占比 79.2%, 此外, AWS、英伟达自研 CPU 市占率分别为 5.2%、12.0%; 从出货量维度来看, 2026Q1 英特尔、AMD 市占率分别为 52.9%、24.5%, 合计占比 77.4%, AWS、英伟达自研 CPU 市占率分别为 6.9%、10.6%。

图32: 全球服务器 CPU 竞争格局（收入维度）



资料来源：IDC，国信证券经济研究所整理

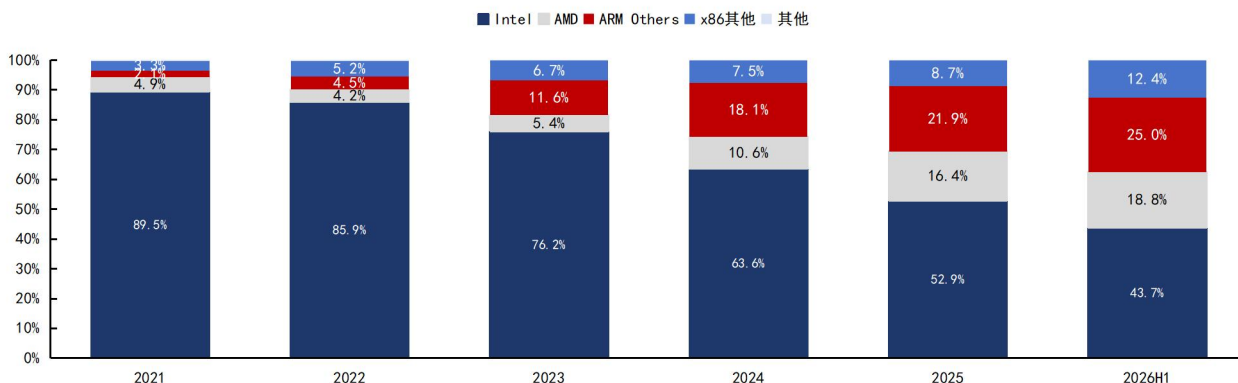
图33: 全球服务器 CPU 竞争格局（出货量）



资料来源：IDC，国信证券经济研究所整理

中国服务器 CPU 市场：英特尔为主，国产 CPU 占比持续提升。根据 IDC 披露数据，2026 年上半年中国不同类型 CPU 通用服务器中，Intel、AMD CPU 服务器占比分别为 43.7%、18.8%，Intel 占据主要市场份额。同时，2026 年上半年 ARM Other（主要为鲲鹏）市占率为 25.0%，市占率快速提升；此外，其他国产 CPU（例如 x86 其他）市占率也从 2022 年 3.3%提升至 2025 年的 12.4%。

图34: 中国服务器 CPU 市场竞争格局

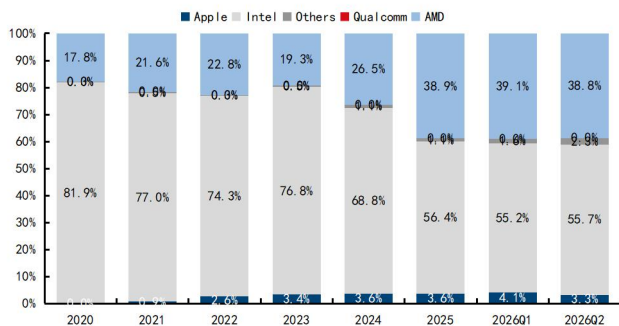


资料来源：IDC，国信证券经济研究所整理

全球消费级 CPU 市场：电脑端英特尔、AMD 为主，平板端苹果、高通、联发科占据主要市场份额。电脑端 CPU 主要包括台式机 CPU 和笔记本电脑 CPU，平板 CPU 考虑低功耗、长续航等特点，ARM 架构 CPU 占据主要市场份额，苹果、高通、联发科为全球龙头。

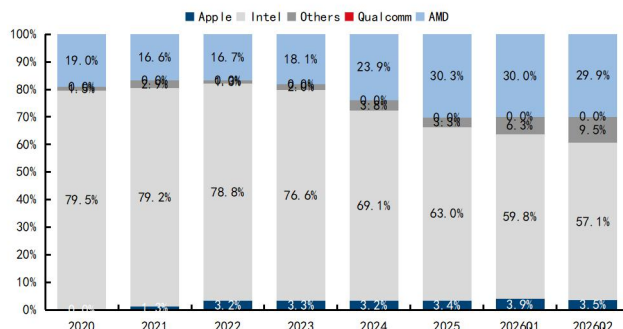
2025 年全球台式机笔记本 CPU 市场：根据 IDC 披露数据，从收入维度来看，英特尔、AMD 市占率分别为 56.4%、38.9%，合计占比 95.3%；从出货量维度来看，英特尔、AMD 市占率分别为 63.0%、30.3%，合计占比 93.3%，占据主要市场份额，AMD 的市占率持续提升。2026Q2，从收入维度来看，英特尔、AMD 市占率分别为 55.7%、38.8%，合计占比 94.5%；从出货量维度来看，2026Q2 英特尔、AMD 市占率分别为 57.1%、29.9%，合计占比 87.0%。

图35: 全球台式机 CPU 竞争格局（收入维度）



资料来源: IDC, 国信证券经济研究所整理

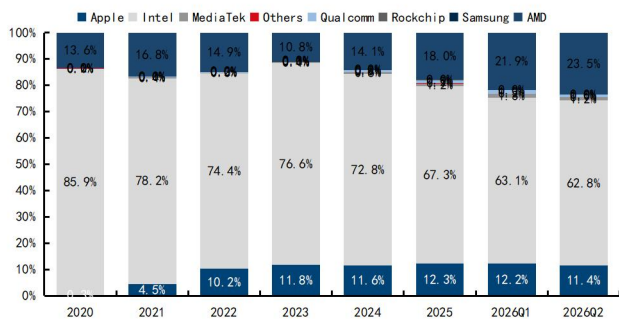
图36: 全球台式机 CPU 竞争格局（出货量）



资料来源: IDC, 国信证券经济研究所整理

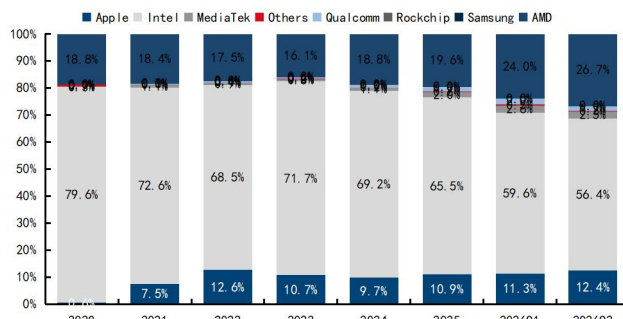
2025 年全球笔记本 CPU 市场: 根据 IDC 披露数据, 从收入维度来看, 英特尔、AMD、苹果市占率分别为 67.3%、18.0%、12.3%, 合计占比 97.5%; 从出货量维度来看, 英特尔、AMD、苹果市占率分别为 65.5%、19.6%、10.9%, 合计占比 96.0%, 英特尔为全球龙头, 苹果自研 CPU 市占率持续提升。从出货量维度来看, 2026Q1 英特尔、AMD、苹果市占率分别为 59.6%、24.0%、11.3%, 合计占比 94.9%。2026Q2, 从收入维度来看, 英特尔、AMD、苹果市占率分别为 62.8%、23.5%、11.4%, 合计占比 97.7%; 从出货量维度来看, 2026Q2 英特尔、AMD、苹果市占率分别为 56.4%、26.7%、12.4%, 合计占比 95.5%。

图37: 全球笔电 CPU 竞争格局（收入维度）



资料来源: IDC, 国信证券经济研究所整理

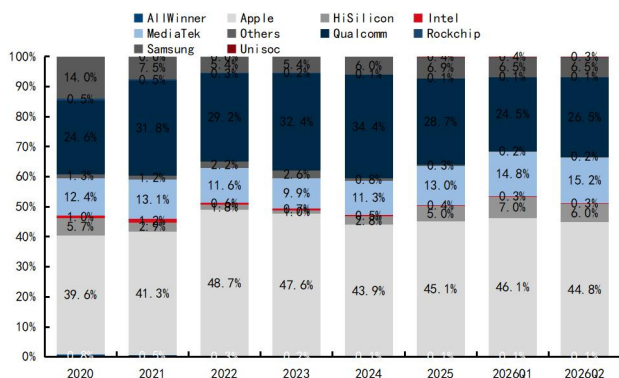
图38: 全球笔电 CPU 竞争格局（出货量）



资料来源: IDC, 国信证券经济研究所整理

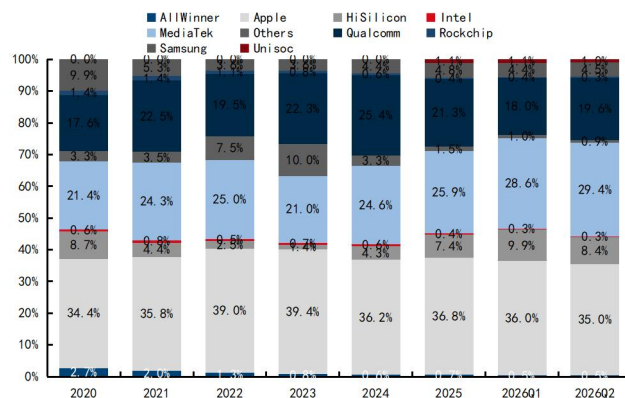
2025 年全球平板 CPU 市场: 根据 IDC 披露数据, 从收入维度来看, 苹果、高通、联发科市占率分别 45.1%、28.7%、13.0%, 合计占比 86.8%; 从出货量维度来看, 苹果、高通、联发科市占率分别为 36.8%、21.3%、25.9%, 合计占比 83.9%, 各家市占率基本维稳。2026Q1, 从收入维度来看, 苹果、高通、联发科市占率分别为 46.1%、24.5%、14.8%, 合计占比 85.4%; 从出货量维度来看, 2026Q1 苹果、高通、联发科市占率分别为 36.0%、18.0%、28.6%, 合计占比 82.6%。2026Q2, 从收入维度来看, 苹果、高通、联发科市占率分别为 44.8%、26.5%、15.2%, 合计占比 86.5%; 从出货量维度来看, 2026Q2 苹果、高通、联发科市占率分别为 35.0%、19.6%、29.4%, 合计占比 84.0%。

图39: 全球平板 CPU 竞争格局（收入维度）



资料来源: IDC, 国信证券经济研究所整理

图40: 全球平板 CPU 竞争格局（出货量）

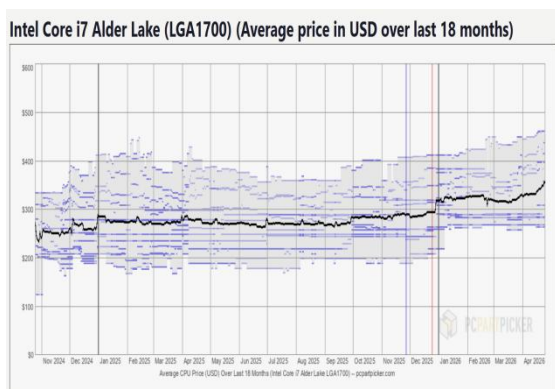


资料来源: IDC, 国信证券经济研究所整理

四、全球 CPU 价格上涨，CPU 国产替代有望加速

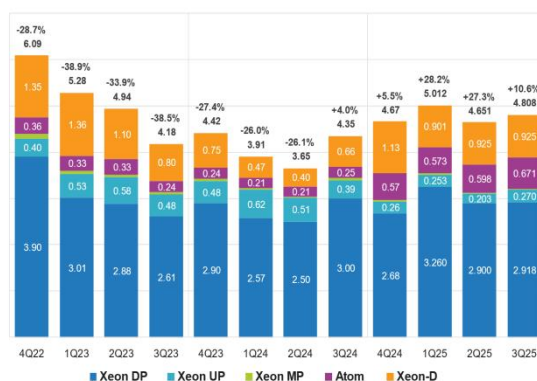
英特尔：公司为全球 CPU 龙头公司，成立于 1968 年，1971 年发布全球首款商用 CPU Intel 4004，上世纪通过绑定 IBM、微软，成长为全球 CPU 龙头。英特尔业务主要包括 CPU 业务、晶圆代工业务等，其中，英特尔 CPU 全面覆盖消费级 CPU、服务器级 CPU，晶圆代工业务 18A 工艺全面量产。根据 IDC 披露数据，25Q3 英特尔 Xeon DP、Xeon UP、Xeon MP、Atom、Xeon-P 出货量分别为 291.8、27、67.92.5 万颗，占比分别为 61.0%、5.6%、14.0%、19.3%。根据 Trendforce 披露信息，英特尔在 26 年 3 月提高了消费级 CPU 价格，并在 4 月 1 日上调了服务器 CPU 价格，预计下半年还有一轮涨价，预计涨幅为 8%-10%；根据 PCPARTPICKER 跟踪的散货价格，英特尔消费级 CPU-i7 Alder lake 自 25 年底以来，价格稳步上涨。根据英特尔财报数据，26Q1 英特尔实现收入 136 亿美金，同比+7%，实现净利润(Non-GAAP) 15 亿美金，同比+156%，毛利率为 39.4%，同比+2.5 个 pct，开始进入业绩释放期。

图41: Intel i7 Alder Lake 散货价（美元）



资料来源: PCPARTPICKER, 国信证券经济研究所整理

图42: 英特尔服务器 CPU 出货量（百万颗）

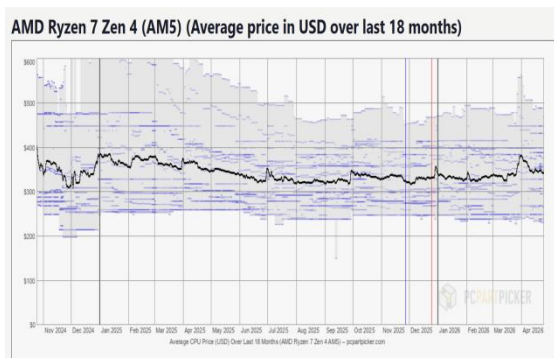


资料来源: IDC, 国信证券经济研究所整理

AMD：公司为全球 CPU 领先公司，成立于 1969 年，1982 年 AMD 和英特尔签署技术交换协议，开始生产 x86 架构 CPU，公司业务主要为 CPU 和 AI 芯片业务，其中 CPU

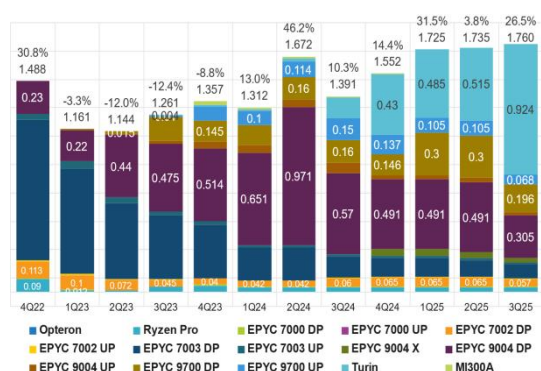
业务全面覆盖消费级 CPU 和服务级 CPU，AI 芯片业务快速发展，2023 年发布 MI300X，进入全球 AI 芯片第一梯队。根据 IDC 披露数据，25Q3AMD Turin、EPYC9700DP、EPYC9004UP 出货量分别为 92.4、19.6、30.5 万颗，占比分别为 52.5%、11.1%、17.3%。根据 Trendforce 披露信息，AMD 计划在 26 年提高其服务器 CPU 产品线价格，预计第二季度将进行一轮提价，第三季度进行另一轮提价，累计涨幅约为 16%-17%；根据 PCPARTICKER 跟踪的散货价格，AMD 消费级 CPU-Ryzen Zen4 自 25 年底以来，价格稳步上涨。

图43: AMD Ryzen7 Zen4 散货价（美元）



资料来源：PCPARTICKER，国信证券经济研究所整理

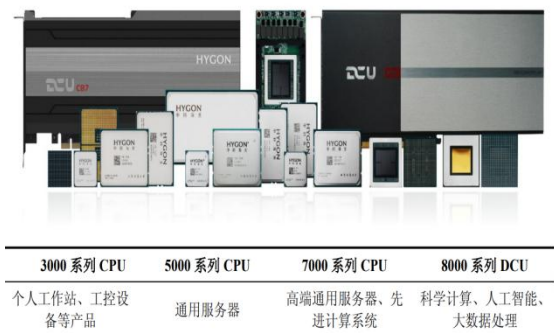
图44: AMD 服务器 CPU 出货量（百万颗）



资料来源：IDC，国信证券经济研究所整理

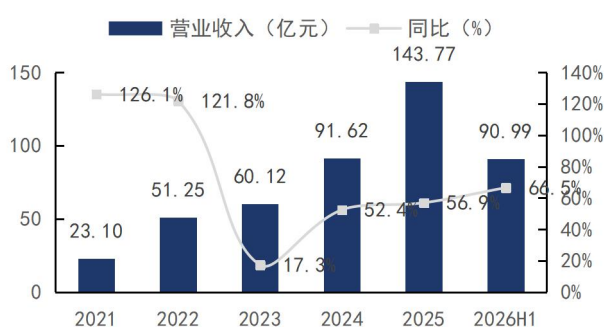
海光信息：国产 CPU 领军，互联网领域 CPU 占比有望提升。公司成立于 2014 年，是国内唯一拥有 AMD x86 指令集永久授权的高端处理器企业，业务主要包括 CPU 业务和 DCU 业务。公司 CPU 产品基于自主的 C86 架构（兼容 x86 生态），分为 3000、5000、7000 系列，全面覆盖消费级、数据中心级 CPU 产品，目前在信创领域进行 CPU 国产替代，2026 年随着 Agent 爆发，全球 CPU 供不应求，公司 CPU 产品在互联网客户占比有望提升。此外，公司 DCU 业务快速发展，采用 GPGPU 架构，拥有完善的统一底层硬件驱动平台，能够适配不同 API 接口和编译器，并支持常见的函数库，全面适配国内外主流大模型，广泛应用于国内互联网等商业场景。

图45: 海光信息产品矩阵



资料来源：公司财报，国信证券经济研究所整理

图46: 海光营收（亿元）及增速



资料来源：公司财报，国信证券经济研究所整理

海思半导体：国产 CPU 龙头，国产 ARM 架构 CPU 领军者。公司成立于 2004 年，主要产品包括移动终端麒麟 SOC 和数据中心鲲鹏 CPU。鲲鹏 920 是国内领先的 ARM-Based 处理器，通过优化分支预测算法、提升运算单元数量、改进内存子系统架构等一系列微架构设计，大幅提高处理器性能，能效比优势突出。

投资建议

投资建议：AI Agent 拉动 CPU 需求，旨在支撑 Agent 任务编排、沙盒、工具运行等，根据 IDC 披露数据，预计 2030 年全球 CPU 市场规模增长至 1307.8 亿美元，对应 25-30 年 CAGR 为 34.1%；国内 CPU 市场亦在快速增长，国产 CPU 市占率持续提升，建议关注海光信息。

风险提示

CPU 需求不及预期风险

本轮 CPU 需求主要由 AI Agent 拉动，如果 Agent 渗透率增长不及预期，互联网大厂可能暂缓 CPU 采购，存在 CPU 需求不及预期风险。

CPU 价格下跌风险

本轮 CPU 涨价主要由于供不应求，若产能端获得改善，CPU 价格存在下跌风险。

互联网大厂自研 CPU 替代第三方 CPU 风险

海外英伟达、AWS 等公司自研 CPU 芯片，国内阿里发布倚天 CPU，存在大厂依赖自研 CPU，减少对第三方 CPU 采购风险。

免责声明

分析师声明

作者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，通过合理判断并得出结论，力求独立、客观、公正，结论不受任何第三方的授意或影响；作者在过去、现在或未来未就其研究报告所提供的具体建议或所表述的意见直接或间接收取任何报酬，特此声明。

国信证券投资评级

投资评级标准	类别	级别	说明
报告中投资建议所涉及的评级（如有）分为股票评级和行业评级（另有说明的除外）。评级标准为报告发布日后 6 到 12 个月内的相对市场表现，也即报告发布日后的 6 到 12 个月内公司股价（或行业指数）相对同期相关证券市场代表性指数的涨跌幅作为基准。A 股市场以沪深 300 指数（000300.SH）作为基准；新三板市场以三板成指（899001.CSI）为基准；香港市场以恒生指数（HSI.HI）作为基准；美国市场以标普 500 指数（SPX.GI）或纳斯达克指数（IXIC.GI）为基准。	股票 投资评级	优于大市	股价表现优于市场代表性指数 10%以上
		中性	股价表现介于市场代表性指数±10%之间
		弱于大市	股价表现弱于市场代表性指数 10%以上
		无评级	股价与市场代表性指数相比无明显观点
	行业 投资评级	优于大市	行业指数表现优于市场代表性指数 10%以上
		中性	行业指数表现介于市场代表性指数±10%之间
		弱于大市	行业指数表现弱于市场代表性指数 10%以上

重要声明

本报告由国信证券股份有限公司（已具备中国证监会许可的证券投资咨询业务资格）制作；报告版权归国信证券股份有限公司

关本报告的摘要或节选都不代表本报告正式完整的观点，一切须以我公司向客户发布的本报告完整版本为准。

本报告基于已公开的资料或信息撰写，但我公司不保证该资料及信息的完整性、准确性。本报告所载的信息、资料、建议及推测仅反映我公司于本报告公开发布当日的判断，在不同时期，我公司可能撰写并发布与本报告所载资料、建议及推测不一致的报告。我公司不保证本报告所含信息及资料处于最新状态；我公司可能随时补充、更新和修订有关信息及资料，投资者应当自行关注相关更新和修订内容。我公司或关联机构可能会持有本报告中所提到的公司所发行的证券并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问或金融产品等相关服务。本公司的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告意见或建议不一致的投资决策。

本报告仅供参考之用，不构成出售或购买证券或其他投资标的的要约或邀请。在任何情况下，本报告中的信息和意见均不构成对任何个人的投资建议。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。投资者应结合自己的投资目标和财务状况自行判断是否采用本报告所载内容和信息并自行承担风险，我公司及雇员对投资者使用本报告及其内容而造成的一切后果不承担任何法律责任。

证券投资咨询业务的说明

本公司具备中国证监会核准的证券投资咨询业务资格。证券投资咨询，是指从事证券投资咨询业务的机构及其投资咨询人员以下列形式为证券投资人或者客户提供证券投资分析、预测或者建议等直接或者间接有偿咨询服务的活动：接受投资人或者客户委托，提供证券投资咨询服务；举办有关证券投资咨询的讲座、报告会、分析会等；在报刊上发表证券投资咨询的文章、评论、报告，以及通过电台、电视台等公众传播媒体提供证券投资咨询服务；通过电话、传真、电脑网络等电信设备系统，提供证券投资咨询服务；中国证监会认定的其他形式。

发布证券研究报告是证券投资咨询业务的一种基本形式，指证券公司、证券投资咨询机构对证券及证券相关产品的价值、市场走势或者相关影响因素进行分析，形成证券估值、投资评级等投资分析意见，制作证券研究报告，并向客户发布的行为。

国信证券经济研究所

深圳

深圳市福田区福华一路 125 号国信金融大厦 36 层

邮编：518046 总机：0755-82130833

上海

上海浦东民生路 1199 弄证大五道口广场 1 号楼 12 层

邮编：200135

北京

北京西城区金融大街兴盛街 6 号国信证券 9 层

邮编：100032