

行业周度点评

计算机

北美云厂商 Capex 增速新高，算力投入仍有保障

2024 年 08 月 18 日

评级 同步大市

评级变动: 维持

行业涨跌幅比较



%	1M	3M	12M
计算机	-3.41	-16.71	-36.15
沪深 300	-4.45	-9.04	-12.67

何晨 分析师

执业证书编号:S0530513080001
hechen@hncasing.com

黄奕景 研究助理

huangyijing@hncasing.com

相关报告

- 1 计算机行业点评：北美云厂商 24Q2 资本支出同环比大幅增加，算力投入确定性依旧 2024-08-06
- 2 计算机行业 7 月月报：AI 景气度不减，自主可控为大势所趋 2024-08-01
- 3 计算机行业点评：英伟达发布 NIM 软件包，简化 AI 部署 2024-08-01

重点股票	2023A		2024E		2025E		评级
	EPS (元)	PE (倍)	EPS (元)	PE (倍)	EPS (元)	PE (倍)	
海光信息	0.54	141.39	0.72	106.44	1.04	74.23	增持

资料来源: iFinD, 财信证券

投资要点:

- **【行业观点】北美四大云厂商发布 24Q2 财报，合计资本支出同比增长 59%，环比增长 19%，24 年全年资本支出指引维持乐观。**Meta、Microsoft、Google 和 Amazon 24Q2 的资本支出合计 531 亿元，同比增长 59%，环比增长 19%，增速达到近 4 年新高。AI 算力景气度持续得到验证，虽然近期有投资者担心算力投入和现金流成长的匹配度问题，但我们认为 1) 在此阶段算力投资的军备竞赛仍是第一要务，根据 Google 的 CEO 所言，公司认为在现阶段对于 AI 基础设施投资不足的风险远远大于投资过剩的风险。2) 当下 AI 业务暂未产生较多正现金流的情况下，各云厂商的现金流也足以覆盖未来 1-2 年的算力投资高速增长，24Q2 四大云厂商合计经营性现金流 1085 亿元，回购股份、股息支付和资本支出三者合计占用经营性现金流的 82%，现金流仍然比较充沛。根据我们的测算，从保障一定现金流冗余与股份回购+股息支付的稳定性的角度出发，北美四大云厂商在 2024 年还有足够的能力和意愿维持算力投资的高增速，我们在当期仍然看好 AI 高强度投资给 AI 算力产业链带来的业绩增长；后续在 2025-2026 年的某个阶段，各云厂商的经营策略、包括市场的考察重心将转向 AI 带来的收入增速和投资强度的匹配。
- **【行业观点】英伟达发布全新软件包 NIM。**NVIDIA NIM 是一套经过优化的云原生微服务，旨在缩短 AI 上市时间。NIM 作为英伟达在 AI 应用层的全新软件工具包，进一步巩固了英伟达在 AI 领域的竞争优势和护城河。国内的企业级 AI 开发者也有望通过英伟达中国的解决方案合作伙伴使用 NIM 服务进行 AI 模型部署。
- **【行情回顾】报告期内申万计算机指数下跌 2.63%，行业排名第 27 位。**本报告期内（2024 年 7 月 29 日-2024 年 8 月 16 日）市场，上证指数下跌 0.40%，沪深 300 指数下跌 1.87%，创业板指下跌 4.10%，申万计算机指数下跌 2.83%，落后上证指数 2.43 个百分点，落后沪深 300 指数 0.96 个百分点，在申万一级行业中排名第 27 位。重点公司方面，涨幅前三为四维图新（+10.48%，智能汽车），海光信息（+3.40%，AI 芯片）和同花顺（+2.65%，金融 IT）。
- **【估值情况】板块绝对估值处于 0.37%分位，相对沪深 300 估值溢价处于 0.00%分位。**至 2024 年 8 月 16 日，申万计算机指数 PE-TTM（剔除负值）为 31.22 倍，位于近十年的 0.37%分位数；申万计算机指数相对沪深 300 指数的估值溢价（PE-TTM，剔除负值）为 2.79 倍，位于近

十年的 0.00%分位数。

- **风险提示：**AI 技术发展不及预期；行业需求不及预期；宏观经济状况不及预期；政策不及预期。

内容目录

1 行业观点	4
1.1 北美四大云厂商发布 24Q2 财报，合计资本支出同比增长 59%，环比增长 19%，24 年全年资本支出指引维持乐观，后续算力投资仍有保障.....	4
1.2 英伟达发布 NIM 软件包，简化 AI 部署.....	6
2 市场表现	7
2.1 行情回顾.....	7
2.2 估值情况.....	9
3 行业重大资讯	10
> 鸿海：AI 服务器需求持续强劲，英伟达 GB200 服务器四季度如期出货.....	10
> 英业达：7 月营收创历史新高 全年服务器出货有望增长超 20%.....	10
> 英伟达推出 B200A 瞄准 OEM 客群，预估 2025 年高端 GPU 出货量年增 55%.....	10
> 苹果官宣：在谷歌 TPU 上训练其 AI 模型.....	11
> 英伟达发布软件和服务 旨在推动人工智能快速普及.....	12
4 重点公司动态	12
> 中兴通讯：发布 2024 年半年度报告.....	12
> 海康威视：发布 2024 年半年度报告.....	12
> 中科曙光：发布 2024H1 业绩快报.....	13
> 工业富联：发布 2024 年半年度报告.....	13
> 英维克：发布 2024 年半年度报告.....	13
> 新点软件：发布 2024 年半年度业绩预告.....	13
5 风险提示	13

图表目录

图 1：21Q1-24Q2 北美四大云厂商资本开支（亿美元）与合计资本开支同比增速（%）.....	4
图 2：21Q1-24Q2 北美四大云厂商资本开支同比增速（%）.....	4
图 3：21Q1-24Q2 北美四大云厂商经营性现金流、回购股份、股息支付和资本支出情况（分季度，单位：亿元）.....	5
图 4：NVIDIA NIM 工作原理.....	6
图 5：申万一级子行业涨跌幅榜（%）.....	7
图 6：申万计算机二级子行业涨跌幅榜（%）.....	7
图 7：计算机板块重点公司情况.....	8
图 8：近十年申万计算机指数 PE-TTM（剔除负值）.....	9
图 9：近十年申万计算机指数相对沪深 300 指数估值溢价（PE-TTM，剔除负值）.....	9
表 1：2020-2026E 北美四大云厂商经营性现金流、回购股份、股息支付和资本支出（分年度，单位：亿元）.....	5
表 2：计算机板块个股涨跌幅前五名.....	8

1 行业观点

1.1 北美四大云厂商发布 24Q2 财报，合计资本支出同比增长 59%，环比增长 19%，24 年全年资本支出指引维持乐观，后续算力投资仍有保障

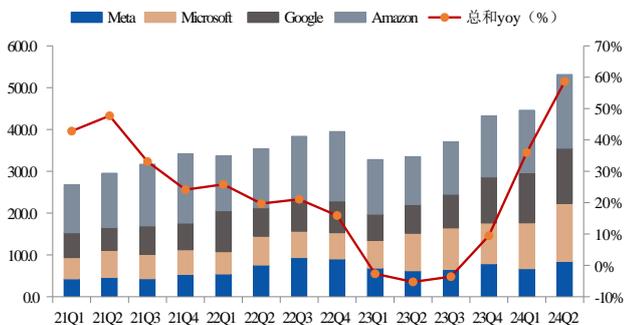
Meta: 24Q2 资本支出为 85 亿美元，同比增长 36%，环比增长 26%。公司上调 2024 年全年资本支出指引至 370-400 亿美元之间（Q1 指引 350-400 亿美元），预计 2025 年的资本支出将大幅增长，用于支持人工智能研究和产品开发工作。

Microsoft: 24Q2 资本支出为 139 亿美元（现金支出），同比增长 55%，环比增长 27%。公司认为 Azure AI 的增长推动了大部分资本支出，考虑到公司对云和 AI 的需求以及现有的 AI 容量限制，预计资本支出将环比增加。

Google: 24Q2 资本支出为 132 亿美元，同比增长 91%，环比增长 10%。公司预计全年每个季度的资本支出将大致保持在或高于 Q1 的 120 亿美元，同时需要注意现金支付的时间可能会影响每个季度报告的资本支出的数额。

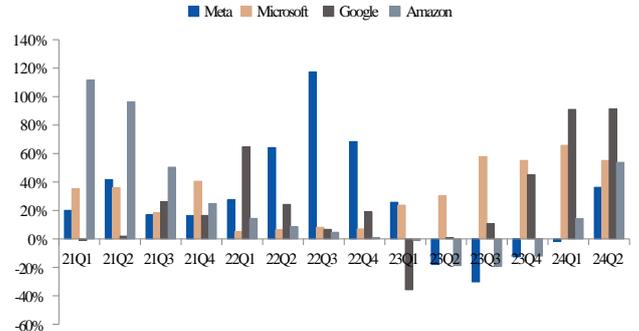
Amazon: 24Q2 资本支出为 176 亿美元，同比增长 54%，环比增长 18%。公司预计 2024H2 的资本支出将会更高，这其中大部分是对于 AWS 基础设施的投资，包括生成式人工智能和非生成式人工智能工作负载需求的满足。

图 1: 21Q1-24Q2 北美四大云厂商资本开支（亿美元）与合计资本开支同比增速（%）



资料来源: Meta、Microsoft、Google 和 Amazon 财报及电话会，财信证券

图 2: 21Q1-24Q2 北美四大云厂商资本开支同比增速 (%)

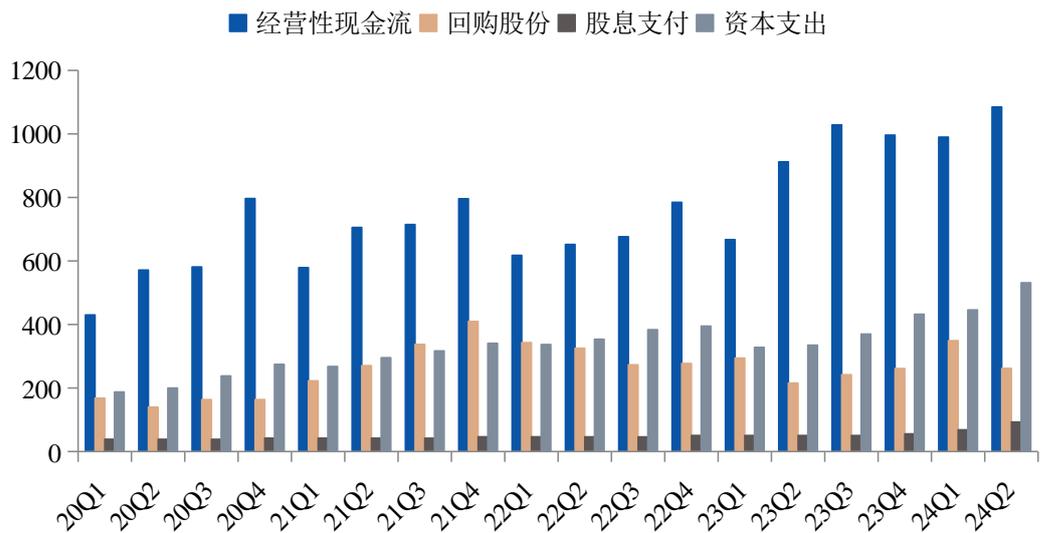


资料来源: Meta、Microsoft、Google 和 Amazon 财报及电话会，财信证券

Meta、Microsoft、Google 和 Amazon 24Q2 的资本支出合计 531 亿元，同比增长 59%，环比增长 19%，同比增速创近 4 年新高，24 年全年资本开支指引继续维持乐观。AI 算力景气度持续得到验证，虽然近期有投资者担心算力投入和现金流成长的匹配度问题，但我们认为 1) 在此阶段算力投资的军备竞赛仍是第一要务，根据 Google 的 CEO 所言，公司认为在现阶段对于 AI 基础设施投资不足的风险远远大于投资过剩的风险，即使最终证明投资过多，这些基础设施也有广泛的用途，并且具有较长的使用寿命，公司可以在整个业务中应用这些基础设施，并解决任何过剩问题。2) 当下 AI 业务暂未产生较多正

现金流的情况下，各云厂商的现金流也足以覆盖未来 1-2 年的算力投资高增速，24Q2 四大云厂商合计经营性现金流 1085 亿元，回购股份 262 亿元，股息支付 93 亿元，资本支出 531 亿元，三者合计占用经营性现金流的 82%，现金流仍然比较充沛。考虑未来几年资本开支的可持续性：假设 2024 年四大云厂商经营性现金流同比 2023 年增长 15%，达到 4145 亿元；回购和股息支付比例不变，而资本开支维持 50% 的增速达到 2198 亿元，2024 年全年仍将维持 500 亿元左右的现金流冗余。根据各云厂商的业绩指引，假设 2025/2026 年四大云厂商的经营性现金流维持 15% 同比增速，且回购股份和股息支付比例不变，而资本开支维持 30% 的增速，预计 2025 年的现金流冗余为 295 亿元；2026 年的现金流冗余为 -89 亿元，回购+股息+资本支出的现金流与经营性现金流基本持平。根据以上测算，从保障一定现金流冗余以及股份回购+股息支付的稳定性等角度出发，北美四大云厂商在 2024 年还有足够的能力和意愿维持算力投资的高增速，我们在当期仍然看好 AI 高强度投资给 AI 算力产业链带来的业绩增长；后续在 2025-2026 年的某个阶段，各云厂商的经营策略、包括市场的考察重心将转向 AI 带来的收入增速和投资强度的匹配。

图 3：21Q1-24Q2 北美四大云厂商经营性现金流、回购股份、股息支付和资本支出情况（分季度，单位：亿元）



资料来源：iChoice, Meta、Microsoft、Google 和 Amazon 财报及电话会，财信证券

表 1：2020-2026E 北美四大云厂商经营性现金流、回购股份、股息支付和资本支出（分年度，单位：亿元）

	2020	2021	2022	2023	2024H1	2024E	2025E	2026E
经营性现金流	2379.63	2795.71	2731.08	3604.52	2074.86	4145.20	4766.98	5482.02
回购股份	635.49	1240.36	1218.63	1013.23	611.10	1165.21	1340.00	1541.00
股息支付	158.27	172.93	189.64	207.38	161.51	238.49	274.26	315.40
资本支出	900.34	1221.19	1469.57	1465.48	977.58	2198.22	2857.69	3714.99
现金流冗余	685.53	161.23	-146.76	918.43	324.67	543.28	295.03	-89.36
回购+股息+资本支出 (%)	71.19%	94.23%	105.37%	74.52%	84.35%	86.89%	93.81%	101.63%
回购+股息 (%)	33.36%	50.55%	51.56%	33.86%	37.24%	33.86%	33.86%	33.86%

资料来源：iChoice, Meta、Microsoft、Google 和 Amazon 财报及电话会，财信证券

建议关注 AI 算力产业链：紫光股份、浪潮信息、海光信息、中科曙光、景嘉微、润泽科技等。

1.2 英伟达发布 NIM 软件包，简化 AI 部署

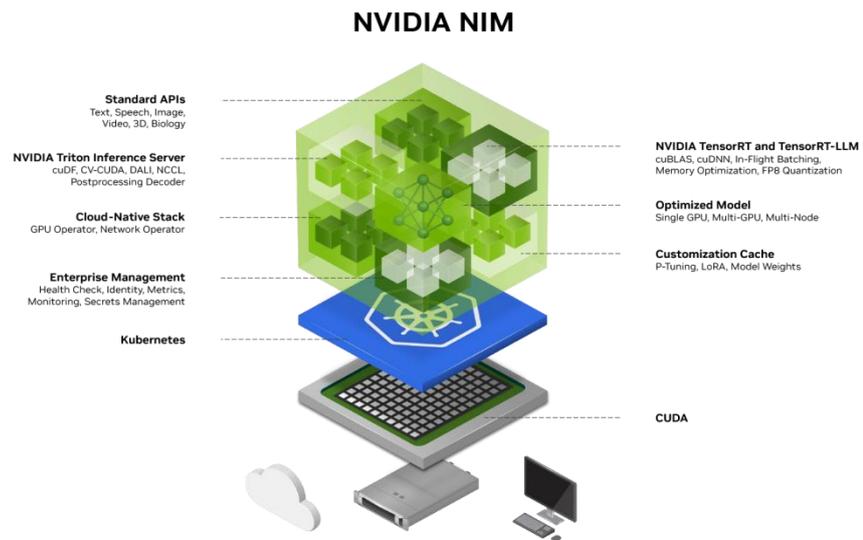
英伟达发布全新软件包 NIM。财联社 7 月 30 日电，英伟达公司正式发布了 Nvidia inference micro services(NIM)，这个软件包可以解决将 AI 用于特定目的所需的许多后勤问题。生成式 AI 是聊天机器人、语音识别和其他自动化人机交互背后的技术，通常需要协调众多硬件、软件和信息检索。许多公司不具备这方面的专业知识，所以英伟达试图就此提供服务并收取费用。英伟达创始人兼首席执行官黄仁勋周一在丹佛举行的 Siggraph 大会上展示并宣传了该最新产品。

什么是 NIM: NVIDIA NIM 是一套经过优化的云原生微服务，旨在缩短 AI 上市时间，并简化生成式 AI 模型在云、数据中心和 GPU 加速工作站的任何位置的部署。借助行业标准 API 和抽象化 AI 模型开发，开发者可以使用几行代码借助 NIM 部署 AI 模型。

NIM 的核心优势: 1) 随时随地部署；2) 使用行业标准 API 进行开发；3) 利用特定领域的模型；4) 在优化的推理引擎上运行；5) 支持企业级 AI；6) 加速 AI 模型随时可供部署。

NIM 作为英伟达在 AI 应用层的全新软件工具包，进一步巩固了英伟达在 AI 领域的竞争优势和护城河。国内的企业级 AI 开发者也有望通过英伟达中国的解决方案合作伙伴使用 NIM 服务进行 AI 模型部署。建议关注：亚康股份、紫光股份、浪潮信息。

图 4：NVIDIA NIM 工作原理



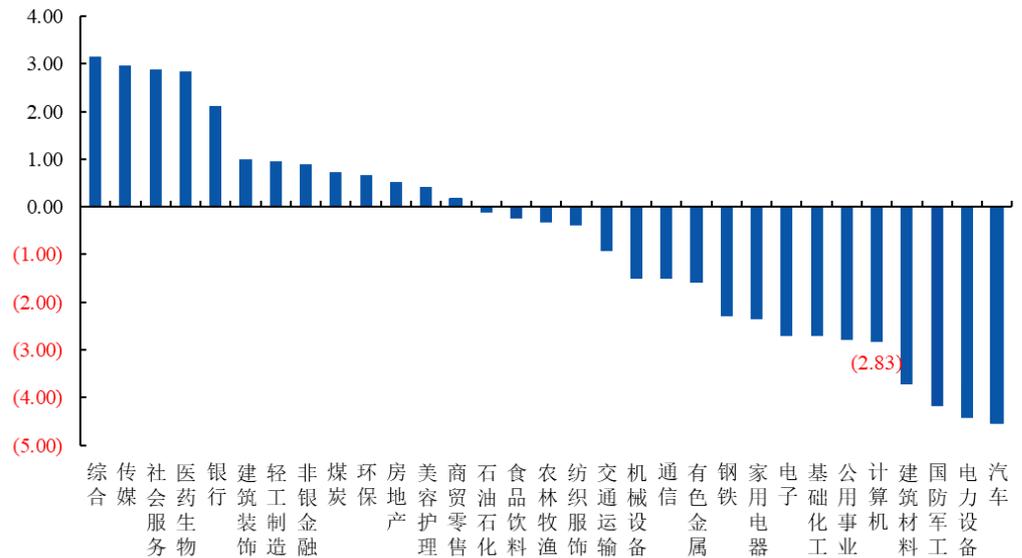
资料来源：英伟达官网，财信证券

2 市场表现

2.1 行情回顾

本报告期内（2024年7月29日-2024年8月16日）市场，上证指数下跌0.40%，沪深300指数下跌1.87%，创业板指下跌4.10%，申万计算机指数下跌2.83%，落后上证指数2.43个百分点，落后沪深300指数0.96个百分点，在申万一级行业中排名第27位。

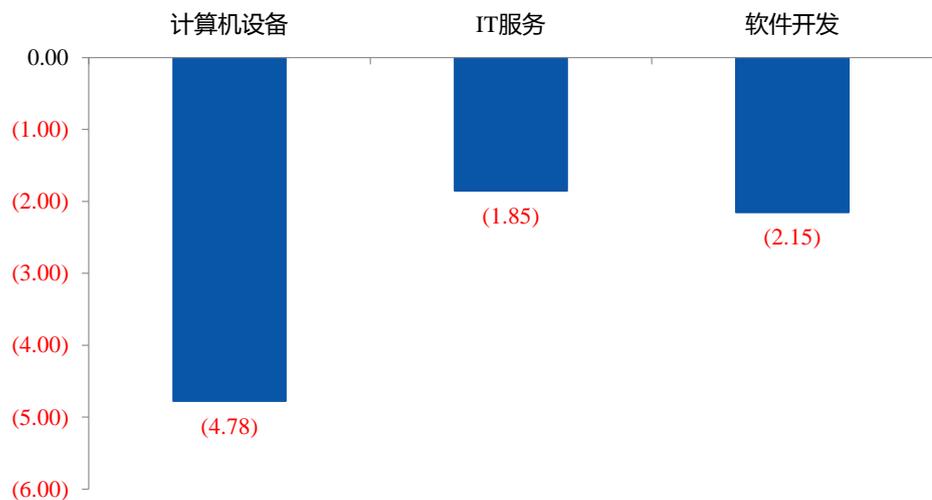
图 5：申万一级子行业涨跌幅榜（%）



资料来源：ifind，财信证券

申万计算机二级子行业指数悉数下跌，其中IT服务指数下跌4.78%，软件开发指数下跌1.85%，计算机设备指数下跌2.15%。

图 6：申万计算机二级子行业涨跌幅榜（%）



资料来源：ifind，财信证券

个股方面，剔除*ST、终止上市及北交所上市公司外，申万计算机板块 323 只股票中 108 只上涨，213 只下跌，2 只持平，个股涨幅中位数为-2.86%。其中涨幅前五的股票分别为启明信息、飞利信、佳缘科技、天迈科技和古鳌科技；跌幅前五名的股票为威创股份、信安世纪、井松智能、中科星图和中科江南。

表 2：计算机板块个股涨跌幅前五名

板块涨幅前五名			板块跌幅前五名		
股票名称	涨跌幅(%)	关注点	股票名称	涨跌幅(%)	关注点
联合光电	32.35	AI眼镜	威创股份	-53.13	ST
天利科技	27.70	金融科技	新炬网络	-27.68	IDC 运维
浩云科技	25.43	低代码平台	任子行	-25.56	网络安全
思创医惠	24.35	医疗 IT	佳缘科技	-20.95	网络安全
汇纳科技	21.32	AI 视频	云从科技	-20.34	AI 应用

资料来源：ifind，财信证券

重点公司方面，涨幅前三为四维图新(+10.48%，智能汽车)，海光信息(+3.40%，AI 芯片)和同花顺(+2.65%，金融 IT)。

图 7：计算机板块重点公司情况

2024-8-16 单位：亿元	股票代码	股票简称	总市值	营业收入			归母净利润			PE			PS			年初至今(%)	本报告期(%)
				2023A	2024E	2025E	2023A	2024E	2025E	2023A	2024E	2025E	2023A	2024E	2025E		
网络安全	601360.SH	三六零	514	91	95	103	(6)	0	2	(0.04)	1093	225	6	5	5	-19.30	-0.69
	300454.SZ	深信服	198	77	83	93	2	3	4	100	68	47	3	2	2	-34.79	-2.87
	688561.SH	奇安信	158	64	70	79	1	2	4	220	71	43	2	2	2	-42.55	2.31
	002439.SZ	启明星辰	162	45	55	66	7	9	12	22	17	14	4	3	2	-50.28	-9.59
	603232.SH	格尔软件	23	6	7	8	0	1	1	62	32	25	4	3	3	-28.39	-3.46
金融 II	300033.SZ	同花顺	546	36	40	46	14	16	19	39	34	29	15	14	12	-34.28	2.65
	600570.SH	恒生电子	311	73	80	90	14	17	20	22	18	16	4	4	3	-42.72	-5.91
医疗 II	300253.SZ	卫宁健康	119	32	37	43	4	5	7	33	22	17	4	3	3	-23.43	1.67
	300168.SZ	万达信息	77	25	25	26	(0)	(0)	(2)	(0)	(25)	(8)	3	3	3	-37.38	-1.30
建筑 II	002410.SZ	广联达	165	65	69	76	1	5	8	143	30	21	3	2	2	-41.89	-3.13
交通 II	301339.SZ	通行宝	81	7	9	12	2	3	3	42	32	24	11	9	7	-1.76	-8.44
	002373.SZ	千方科技	128	78	93	112	5	7	8	24	19	17	2	1	1	-27.33	-2.41
	300552.SZ	万集科技	62	9	12	20	(4)	(0)	1	(16)	(665)	44	7	5	3	-1.89	-4.91
工业软件	600845.SH	宝信软件	875	129	151	179	26	31	38	34	28	23	7	6	5	-23.83	-5.13
	301269.SZ	华大九天	387	10	13	17	2	1	3	193	281	147	38	30	23	-32.59	-4.86
	688083.SH	中望软件	77	8	10	13	1	1	1	125	81	54	9	8	6	-36.34	-0.74
智能汽车	002920.SZ	德赛西威	502	219	281	352	15	21	27	32	24	18	2	2	1	-29.65	-0.53
	300496.SZ	中科创达	187	52	62	75	5	6	8	40	30	23	4	3	3	-48.90	-5.47
	002405.SZ	四维图新	170	31	36	42	(13)	(7)	(2)	(13)	(25)	(83)	5	5	4	-19.44	10.48
安防监控	002415.SZ	海康威视	2570	893	996	1124	141	168	194	18	15	13	3	3	2	-18.57	-6.80
	002236.SZ	大华股份	459	322	355	398	74	39	46	6	12	10	1	1	1	-22.90	-4.13
数据要素	300212.SZ	易华录	116	8	16	21	(9)	0	2	(6)	1883	74	15	7	5	-48.93	0.00
	002368.SZ	太极股份	95	92	101	115	4	5	6	25	21	16	1	1	1	-48.18	-5.94
	688031.SH	星环科技	36	5	7	9	(0)	(2)	0	(12)	(21)	(162)	7	5	4	-54.62	-4.65
A 应用	688111.SH	金山办公	877	46	57	74	13	17	22	67	53	40	19	15	12	-39.85	-2.93
	002230.SZ	科大讯飞	797	197	231	274	7	9	12	121	87	68	4	3	3	-25.54	-10.86
	300624.SZ	万兴科技	82	15	18	22	1	1	2	95	69	51	6	5	4	-37.10	-8.88
A 芯片	688041.SH	海光信息	1786	60	84	114	13	20	27	141	90	65	30	21	16	8.42	3.40
	688256.SH	寒武纪	1021	7	14	23	(8)	(2)	(2)	(120)	(210)	(614)	144	73	45	81.24	-2.44
	300474.SZ	景嘉微	274	7	12	16	1	2	3	459	125	82	38	24	17	-15.37	-7.75
服务器 & 交换机	601138.SH	工业富联	4184	4763	5606	6554	210	255	310	20	16	14	1	1	1	44.84	-8.12
	000938.SZ	紫光股份	570	773	862	967	21	26	32	27	22	18	1	1	1	3.75	-12.36
	000977.SZ	浪潮信息	487	659	777	896	18	22	27	27	22	18	1	1	1	0.06	-10.81
	603019.SH	中科曙光	579	144	166	192	18	23	27	32	26	21	4	3	3	0.56	-7.81
DC	300442.SZ	润泽科技	366	44	60	83	18	23	32	21	16	11	8	6	4	-15.62	-3.40
	300383.SZ	光环新网	142	79	89	102	4	6	8	37	24	19	2	2	1	-17.88	0.64
	603881.SH	数据港	74	15	16	17	1	1	2	60	51	43	5	5	4	-18.53	-0.40

资料来源：ifind，财信证券

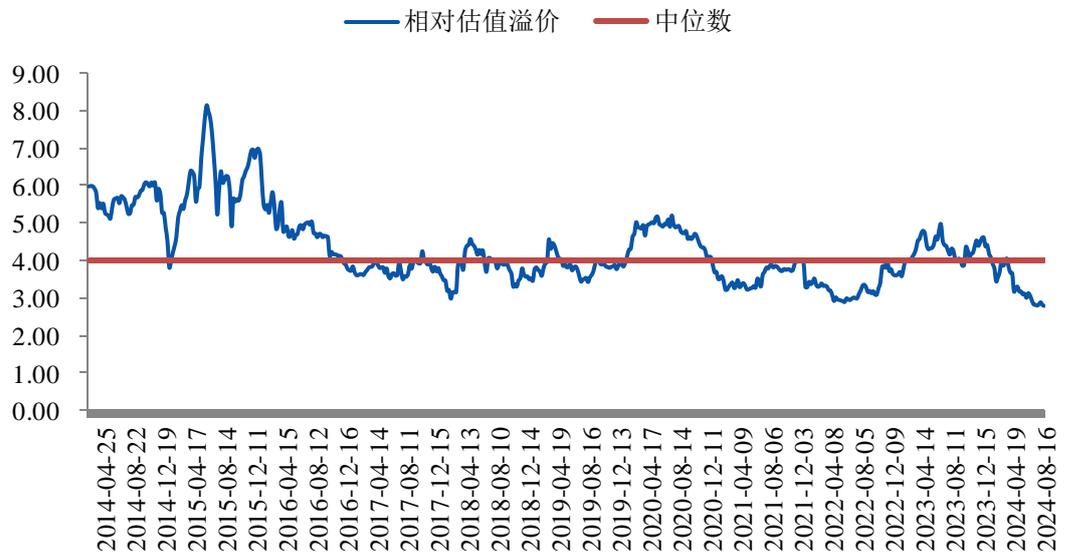
2.2 估值情况

至 2024 年 8 月 16 日，申万计算机指数 PE-TTM（剔除负值）为 31.22 倍，位于近十年的 0.37% 分位数；申万计算机指数相对沪深 300 指数的估值溢价（PE-TTM，剔除负值）为 2.79 倍，位于近十年的 0.00% 分位数。

图 8：近十年申万计算机指数 PE-TTM（剔除负值）



图 9：近十年申万计算机指数相对沪深 300 指数估值溢价（PE-TTM，剔除负值）



3 行业重大资讯

➤ 鸿海：AI 服务器需求持续强劲，英伟达 GB200 服务器四季度如期出货

IT 之家 8 月 14 日消息，鸿海今天举行法人说明会，发言人巫俊毅指出，鸿海将在第四季度开始出货搭载英伟达 GB200 处理器的服务器。

鸿海方面预计，今年全年整体营运将优于 2023 年，其中消费智能产品、电脑终端产品为持平，但云端网络产品、元件及其他产品将大幅成长。巫俊毅指出，AI 服务器能见度显著提升，需求强劲，占整体服务器营收比重超 40%，将成鸿海集团下一个万亿元新台币营收产品。

资料来源：IT 之家

➤ 英业达：7 月营收创历史新高 全年服务器出货有望增长超 20%

同花顺 8 月 8 日消息，代工厂英业达 7 月合并营收 555.33 亿新台币，环比增长 9.04%，同比增长 32.75%，创历史新高。累计 1-7 月合并营收 3401.98 亿新台币，同比增长 16.24%。机构表示，预期第三季笔电出货环比持平，全年预计将有个位数增长；全年服务器出货有望增长超过 20%。

资料来源：同花顺

➤ 英伟达推出 B200A 瞄准 OEM 客群，预估 2025 年高端 GPU 出货量年增长 55%

TrendForce 集邦咨询 8 月 7 日消息，市场近日传出 NVIDIA（英伟达）取消 B100 并转为 B200A，但根据 TrendForce 集邦咨询了解，NVIDIA 仍计划在 2024 年下半年推出 B100 及 B200，供应 CSPs（云端服务业者）客户，并另外规划降规版 B200A 给其他企业型客户，瞄准边缘 AI（人工智能）应用。

TrendForce 集邦咨询表示，受 CoWoS-L 封装产能吃紧影响，NVIDIA 会将 B100 及 B200 产能提供给需求较大的 CSPs 客户，并规划于 2024 年第三季后陆续供货。在 CoWoS-L 良率和量产尚待整备的情况下，NVIDIA 同步规划降规版 B200A 给其他企业客户，并转为采用 CoWoS-S 封装技术。

TrendForce 集邦咨询预期，B200A 的热设计功耗（TDP）将比 B200 低，搭配该芯片的 GB Rack（机柜）可采用气冷散热方案，预计 2025 年较不受设计难度高且复杂的液冷散热影响，从而避免出货延迟等问题。B200A 的存储器规格将采用 4 颗 HBM3e（第五代高带宽内存）12hi（12 层堆叠），总容量为 144GB。预期 OEMs（原始设备制造商）应于 2025 年上半年正式拿到 B200A 芯片，这个供货时间点能让延迟至今年第三季才能放量的 H200 有更多被市场采用的机会，避免产品线相隔太近而产生冲突。

根据 TrendForce 集邦咨询对供应链的调查, 2024 年 NVIDIA 的高端 GPU (图形处理器) 出货将以 Hopper 平台产品为主, 除针对北美 CSPs、OEMs 出货 H100、H200 等机种, 针对中国客户则以搭载 H20 的 AI 服务器为主力。预估 H200 在 2024 年第三季才能开始放量、成为 NVIDIA 主流机种, 并延续至 2025 年。

TrendForce 集邦咨询指出, Blackwell 系列于 2024 年仍在前期出货阶段, 进入 2025 年, Blackwell 将成为出货主力, 以效能较高的 B200 及 GB200 Rack 满足 CSPs、OEMs 对高端 AI 服务器的需求。而 B100 属过渡型、主打耗能较低的产品, 在 NVIDIA 出货完既有 CSPs 订单后, B100 将逐渐被 B200 及 GB200 Rack 取代。TrendForce 集邦咨询预估, 2025 年 Blackwell 平台将占 NVIDIA 高端 GPU 逾 8 成, 并促使 NVIDIA 高端 GPU 系列的出货年增率上升至 55%。

资料来源: TrendForce 集邦咨询

➤ 苹果官宣: 在谷歌 TPU 上训练其 AI 模型

财联社 7 月 30 日讯, 美东时间周一, 苹果公司发布技术论文称, 支撑其人工智能 (AI) 系统“苹果智能” (Apple Intelligence) 的两个 AI 模型, 是在谷歌设计的云端芯片上进行预训练的。

分析人士指出, 苹果决定依赖谷歌的云基础设施是一个值得注意的信号, 因为在生产人工智能处理器方面目前还没有公司能超过英伟达, 英伟达占据了大约 80% 的市场份额。这表明, 在训练尖端人工智能方面, 大型科技公司正在寻找英伟达以外的替代品。

苹果在上述技术报告中详细介绍了选择谷歌自主开发的 Tensor 处理单元 (TPU) 进行训练的情况。此外, 苹果还在周一为部分设备发布了 Apple Intelligence 的预览版。苹果指出, 其 Apple Foundation Model (AFM) 和 AFM 服务器是在“云 TPU 集群”上进行训练的。这意味着苹果租用了云服务提供商的服务器来进行计算。“这种系统使我们能够高效且可扩展地训练 AFM 模型, 包括 AFM 设备端、AFM 服务器和更大的模型。”该公司在报告中写道。

具体而言, 苹果指出, 为了构建可以在 iPhone 和其他设备上运行的人工智能模型 (即设备端 AFM), 苹果使用了 2048 颗 TPUv5p 芯片。这是最先进的 TPU, 于去年 12 月首次推出。而在 AFM 服务器中, 苹果部署了 8192 个 TPUv4 处理器。

此外, 在上述报告中, 苹果没有明确表示它未使用英伟达的芯片, 但它对其人工智能工具和功能的硬件和软件基础设施的描述中, 均没有提到英伟达的硬件。

据了解, 与英伟达将其芯片和系统作为独立产品销售不同, 谷歌通过其谷歌云平台销售 TPU 的使用权。有兴趣购买访问权限的客户必须通过谷歌的云平台构建软件才能使用这些芯片。

根据谷歌网站, 最新的 TPU 在订购期限为三年时, 每小时的使用成本不到 2 美元。

谷歌于 2015 年首次推出 TPU 用于内部工作负载，并于 2017 年向公众开放。它们现在是为人工智能设计的最成熟的定制芯片之一。而且，谷歌也是英伟达的主要客户之一。它在训练 AI 系统时使用了英伟达的 GPU 和自家的 TPU，同时也在其云上出售英伟达技术的使用权限。

苹果的工程师在论文中表示，使用谷歌的芯片有可能制造出比论文中讨论的两种型号更大、更复杂的机型。该公司此前还表示，推理过程（即使用预训练的 AI 模型生成内容或做出预测）将部分在其数据中心的自有芯片上进行。

资料来源：财联社

➤ 英伟达发布软件和服务 旨在推动人工智能快速普及

财联社 7 月 30 日电，全球市值最高的芯片制造商英伟达宣布了一系列软件更新，旨在让更多企业更容易使用生成式人工智能。这家芯片公司正式发布了 Nvidia inference micro services(NIMs)，这个软件包可以解决将 AI 用于特定目的所需的许多后勤问题。生成式 AI 是聊天机器人、语音识别和其他自动化人机交互背后的技术，通常需要协调众多硬件、软件和信息检索。许多公司不具备这方面的专业知识，所以英伟达试图就此提供服务并收取费用。英伟达创始人兼首席执行官黄仁勋周一在丹佛举行的 Siggraph 大会上展示了该公司的最新产品。他在这个图形技术大会上试图说服尽可能多的行业开始使用英伟达的技术，在经济中普及人工智能。

资料来源：财联社

4 重点公司动态

➤ 中兴通讯：发布 2024 年半年度报告

2024H1 公司实现营业收入 624.9 亿元，同比增长 2.9%；净利润 57.3 亿元，同比增长 4.8%。分业务看，上半年，中兴通讯运营商网络、政企、消费者业务分别实现营收 373 亿元、91.7 亿元、160.2 亿元，其中政企业务营收同比增速达 56.1%。

➤ 海康威视：发布 2024 年半年度报告

2024H1 公司实现营业收入 412.09 亿元，同比增加 9.68%；归属于上市公司股东的净利润 50.64 亿元，同比减少 5.13%。

其中，营业收入分地区来看，境内收入为 270.29 亿元，同比增加 5.98%；境外收入为 141.80 亿元，同比增加 17.51%。

另外，主要财务数据显示，上半年，海康威视财务费用为-2.50 亿元，上年同期为-5.67 亿元，主要原因系受汇率波动影响，本期汇兑损失增加。

➤ 中科曙光：发布 2024H1 业绩快报

公司 2024 年 1-6 月实现营业收入 57.12 亿元，同比增长 5.77%；归属于上市公司股东的净利润 5.58 亿元，同比增长 2.43%。

公司表示，报告期内，公司主营业务相对稳定，面对国内激烈的市场竞争，公司坚持研发创新，完善产业链布局，推进数字基础设施建设，稳步推进各项业务，报告期内实现营业总收入 57 亿元，比上年同期增长 3.11 亿元，增长比率为 5.77%；归属于上市公司股东净利润 5.58 亿元，比上年同期增长 0.13 亿元，增长比率为 2.43%，归属于上市公司股东的扣除非经常性损益的净利润 3.54 亿元，比上年同期增长 0.49 亿元，增长比率为 15.93%。公司将继续通过不断优化产品结构、强化预算管理、降本增效等多维度措施，巩固和提升核心竞争力。

➤ 工业富联：发布 2024 年半年度报告

2024H1 公司营业收入 2660.9 亿元，同比增长 28.7%，归属于上市公司股东的净利润 87.4 亿元，同比增长 22%，扣非净利润若扣除完整汇兑损益影响后，二季度利润总额同比增长 25%。其中收入增速加快的原因在于品牌客户不断增加 AI 服务器订单。

➤ 英维克：发布 2024 年半年度报告

2024 年半年度公司实现总营收 17.13 亿元，同比增长 38.24%。成本端营业成本 11.82 亿元，同比增长 40.37%，费用等成本 3.55 亿元，同比增长 18.18%。营业总收入扣除营业成本和各项费用后，归母净利润 1.83 亿元，同比增长 99.63%。

➤ 新点软件：发布 2024 年半年度业绩预告

公司预计 2024 年半年度实现归属于母公司所有者的净利润-2,684.99 万元到-1,790.00 万元，与上年同期（法定披露数据）相比，亏损收窄，将减少 5,764.04 万元到 6,659.03 万元，亏损幅度同比减少 68.22%到 78.81%。

5 风险提示

AI 技术发展不及预期；行业需求不及预期；宏观经济状况不及预期；政策不及预期。

评级系统说明

以报告发布日后的 6—12 个月内，所评股票/行业涨跌幅相对于同期市场指数的涨跌幅度为基准。

类别	投资评级	评级说明
股票投资评级	买入	投资收益率超越沪深 300 指数 15% 以上
	增持	投资收益率相对沪深 300 指数变动幅度为 5%—15%
	持有	投资收益率相对沪深 300 指数变动幅度为-10%—5%
	卖出	投资收益率落后沪深 300 指数 10% 以上
行业投资评级	领先大市	行业指数涨跌幅超越沪深 300 指数 5% 以上
	同步大市	行业指数涨跌幅相对沪深 300 指数变动幅度为-5%—5%
	落后大市	行业指数涨跌幅落后沪深 300 指数 5% 以上

免责声明

本公司具有中国证监会核准的证券投资咨询业务资格，作者具有中国证券业协会注册分析师执业资格或相当的专业胜任能力。

本报告仅供财信证券股份有限公司客户及员工使用。本公司不会因接收人收到本报告而视其为本公司当然客户。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发送，概不构成任何广告。

本报告信息来源于公开资料，本公司对该信息的准确性、完整性或可靠性不作任何保证。本公司对已发报告无更新义务，若报告中所含信息发生变化，本公司可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本报告中所指投资及服务可能不适合个别客户，不构成客户私人咨询建议。任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议。在任何情况下，本公司及本公司员工或者关联机构不承诺投资者一定获利，不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。投资者务必注意，其据此作出的任何投资决策与本公司及本公司员工或者关联机构无关。

市场有风险，投资需谨慎。投资者不应将本报告作为投资决策的唯一参考因素，亦不应认为本报告可以取代自己的判断。在决定投资前，如有需要，投资者务必向专业人士咨询并谨慎决策。

本报告版权仅为本公司所有，未经书面许可，任何机构和个人（包括本公司客户及员工）不得以任何形式复制、发表、引用或传播。

本报告由财信证券研究发展中心对许可范围内人员统一发送，任何人不得在公众媒体或其它渠道对外公开发布。任何机构和个人（包括本公司内部客户及员工）对外散发本报告的，则该机构和个人独自为此发送行为负责，本公司保留对该机构和个人追究相应法律责任的权利。

财信证券研究发展中心

网址：stock.hnchasing.com

地址：湖南省长沙市芙蓉中路二段 80 号顺天国际财富中心 28 层

邮编：410005

电话：0731-84403360

传真：0731-84403438