



Gemini 3 系列发布, OpenAI 正式推出 GPT5.1

— 全球科技动态月报

2025 年 12 月 8 日

内容摘要

- **十一月美股科技跑输大盘, 主要科技股指均回调:** 2025 年 11 月, 标普 500 指数累计+0.13%, 纳斯达克综合指数-1.51%, 费城半导体指数-2.82%; TAMAMA 科技指数-1.75%; 纳斯达克中国金龙指数-5.81%; 恒生科技指数-5.23%; 计算机-5.64%。
- **美股热门科技股大部分下跌, 谷歌月涨幅超 13%;** 港股计算机公司多数下跌: 据统计, 苹果合计+3.14%, 英伟达-12.59%, 特斯拉-5.78%, 谷歌+13.87%, 亚马逊-4.50%, META-0.06%, 微软-4.98%, ARM-20.17%, 英特尔+1.43%, 高通-7.08%, AMD-15.07%。
- **十一月 10 年期美债累计下降 11bps, 美元兑人民币中间价累计调升 78 个基点:** 11 月 28 日, 美国 10 年期国债利率为 4.02%, 相较 11 月 3 日累计下降 11bps; 11 月 28 日, 美元兑人民币中间价报 7.08, 较 11 月 3 日价累计调升 78 个基点。
- **核心观点:**

11 月 18 日, 谷歌发布新一代 AI 模型 Gemini 3 系列, 推理能力大幅提升, 其整合了 Gemini 系列核心能力, 旨在全方位赋能用户学习、创作与规划各类事务, 以顶尖推理能力为核心, Gemini 3 Pro 在多个权威基准测试中大幅领先前代及竞品。其新增 Deep Think 模式可进一步强化复杂问题推理, 目前面向安全测试人员开放, 后续将推给 Google AI Ultra 订阅用户; 还支持多模态信息整合、智能编码、长周期规划等功能, 同步推出开发者平台 Google Antigravity, 实现端到端软件任务自主规划与验证。

11 月 12 日, OpenAI 正式推出 GPT-5 系列升级版本 GPT-5.1, 包含 GPT-5.1 Instant 和 GPT-5.1 Thinking 两款模型, 旨在提升智能水平与对话体验, 同时优化 ChatGPT 的自定义功能。GPT-5.1 Instant 作为广泛使用的基础模型, 默认语气更温暖具对话感, 新增自适应推理能力, 可根据问题难度动态调整思考流程, 在数学、编程等评测中表现提升, 指令遵循更可靠; GPT-5.1 Thinking 则针对复杂任务优化, 能动态调整思考时间, 简单任务响应速度翻倍, 复杂任务答案更全面, 回复更清晰易懂且富有同理心。据我们测算, 以 GPT-5.1 为例, 在给定指数测算过程中, output tokens 约为 80.9 百万 tokens, 总运行成本约为 859.06 美元, 因此估算 GPT-5.1 单位运行成本约为 10.62 美元/百万 tokens。我们认为, 我们认为, 海外大模型高智能化短期暂时还无法规避成本高问题, 尽管当前 AI Agent 依然面临技术成熟度与商业化节奏的阶段性错配问题, 但加速落地的趋势依然不变。

- **风险提示:** 技术迭代不及预期风险; 科技巨头竞争加剧风险; 法律监管风险; 供应链风险; 下游需求不及预期风险。

计算机行业

推荐 维持评级

分析师

吴砚靖

☎: 010-66568589

✉: wuyanqing@chinastock.com.cn

分析师登记编码: S0130519070001

国内表现

2025-12-05



资料来源: 中国银河证券研究院

全球行情

2025-12-05



资料来源: 中国银河证券研究院

相关研究

【银河计算机】全球科技动态月报_美股三季报披露, 2026 年 Capex 指引整体乐观

【银河计算机】全球科技动态月报_OpenAI 发布 Sora2, 与多巨头合作打造生态闭环

【银河计算机】全球科技动态月报_英伟达财报披露, DeepMind 发布 Genie 3

目录

Catalog

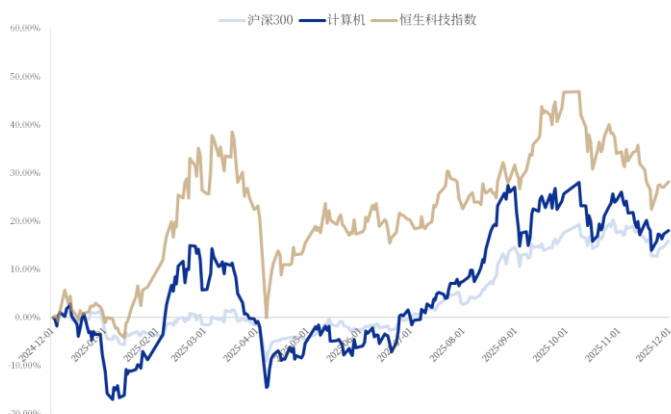
一、全球行情回顾	3
(一) 主要股指变动	3
(二) 热门科技股表现.....	3
(三) 债市及汇率情况.....	6
二、前沿行业动态	6
(一) 前沿政策动态	6
(二) 前沿技术动态	10
三、前沿产品动态	13
(一) 算力及终端	13
(二) 大模型与云应用	18
四、风险提示	26

一、全球行情回顾

（一）主要股指变动

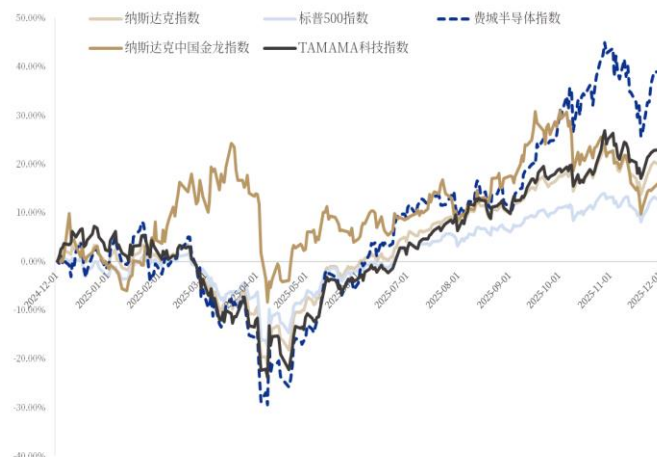
2025 年 11 月，美股科技跑输大盘，主要科技股指均回调。标普 500 指数累计+0.13%，纳斯达克综合指数-1.51%，费城半导体指数-2.82%；TAMAMA 科技指数-1.75%；纳斯达克中国金龙指数-5.81%；恒生科技指数-5.23%；计算机-5.64%。

图1：11 月国内指数表现 TTM



资料来源：WIND，中国银河证券研究院

图2：11 月全球指数表现 TTM



资料来源：WIND，中国银河证券研究院

表1：主要股指月变动

指数代码	指数简称	涨跌幅%				市盈率PE（TTM）
		十一月	2025年初至今	2024年	2023年	
SPX.GI	标普500指数	0.13	16.11	13.31	14.23	29.24
IXIC.GI	纳斯达克指数	-1.51	21.25	28.64	43.42	42.04
SOX.GI	费城半导体指数	-2.82	43.57	19.27	64.90	45.55
8884057.WI	TAMAMA科技指数	-1.75	20.80	46.87	67.81	37.04
HXC.GI	纳斯达克中国金龙指数	-5.81	15.46	4.43	-3.39	23.23
HSTECH.HI	恒生科技指数	-5.23	25.87	18.70	-8.83	23.38
CI005027.WI	计算机	-5.64	26.16	9.91	8.90	131.37

资料来源：WIND，中国银河证券研究院

（二）热门科技股表现

11 月美股热门科技股大部分下跌，谷歌涨幅超 13%；港股计算机公司多数下跌。据统计，苹果合计+3.14%，英伟达-12.59%，特斯拉-5.78%，谷歌+13.87%，亚马逊-4.50%，META-0.06%，微软-4.98%，ARM-20.17%，英特尔+1.43%，高通-7.08%，AMD-15.07%。

表2：重点公司月数据

所属板块	地区	证券代码	公司名称	收盘价 (美元/港元 /新台币)		11月 涨跌幅 (%)	总市值 (亿美元 /亿港元 /亿新台 币)	市销率 PS (TTM)	市盈率 PE (TTM)	PE (2023)	PE (2024)	PE (2025E)
				[11.01]	[11.30]							
							[11.30]					
算力及终端	美股	AAPL.O	苹果(Apple)	270.37	278.85	3.14%	41,203.86	9.90	36.79	30.87	40.18	38.82
			英伟达									
		NVDA.O	(NVIDIA)	202.49	177.00	-12.59%	43,011.00	22.98	43.36	280.04	110.51	0.00
		TSLA.O	特斯拉(TESLA)	456.56	430.17	-5.78%	14,306.68	14.96	281.52	52.67	183.15	270.08
		HPQ.N	惠普(HP)	27.67	24.42	-11.75%	228.25	0.41	9.03	9.14	11.03	0.00
		CSCO.O	思科(CISCO)	73.11	76.94	5.24%	3,039.97	5.27	29.43	16.28	22.85	29.79
		ASML.O	阿斯麦	1,059.23	1,060.00	0.07%	4,174.61	11.09	37.76	34.34	33.90	39.36
			超威半导体									
		AMD.O	(AMD)	256.12	217.53	-15.07%	3,541.48	11.06	107.12	278.85	119.39	79.18
		INTC.O	英特尔(INTEL)	39.99	40.56	1.43%	1,934.71	3.62	977.13	125.43	-4.63	353.70
	港股		高通									
		QCOM.O	(QUALCOMM)	180.90	168.09	-7.08%	1,800.24	4.07	32.49	22.36	16.75	14.85
		ARM.O	ARM	169.82	135.56	-20.17%	1,438.29	32.60	173.29	147.03	424.91	0.00
			安森美半导体									
		ON.O	(ON SEMICON -DUCTOR)	50.08	50.24	0.32%	202.15	3.27	63.35	16.47	16.96	0.00
台股	台股	0909.HK	明源云	3.30	3.15	-4.55%	60.94	4.21	-91.97	-8.66	-25.07	525.08
		9698.HK	万国数据-SW	34.16	33.28	-2.58%	517.56	4.54	8.51	-2.92	9.59	197.99
		1686.HK	新意网集团	6.04	5.14	-14.90%	121.34	4.13	12.39	7.98	11.01	0.00
		2330.TW	台积电	1,500.00	1,440.00	-4.00%	473,365.89	10.28	23.76	18.05	24.07	22.94
	美股	2454.TW	联发科	1,310.00	1,395.00	6.49%	22,374.58	3.83	21.07	21.09	21.30	0.00
			谷歌									
			(ALPHABET)									
		GOOGL.O	-A	281.19	320.18	13.87%	38,664.89	10.03	31.12	23.69	23.09	31.36
			亚马逊									
		AMZN.O	(AMAZON)	244.22	233.22	-4.50%	24,931.72	3.61	32.60	51.61	39.22	33.61
			脸书									
			(META									
		META.O	PLATFORMS)	648.35	647.95	-0.06%	16,331.76	8.62	27.90	23.27	23.79	23.64
		NFLX.O	奈飞(NETFLIX)	1,118.86	107.58	-90.38%	4,558.51	10.51	43.70	39.40	43.77	41.38
		PDD.O	拼多多	134.87	116.08	-13.93%	1,647.93	2.80	11.45	22.94	8.63	11.56

港 股	NTES.O	网易	140.10	138.05	-1.46%	871.41	5.56	17.13	14.46	13.68	17.73
	BIDU.O	百度	120.87	116.89	-3.29%	402.28	2.19	33.95	14.52	8.81	22.47
	TCOM.O	携程网	70.65	69.92	-1.03%	490.68	5.93	11.37	17.58	19.77	11.86
	BABA.N	阿里巴巴	170.43	157.30	-7.70%	3,717.20	2.64	21.26	18.64	17.91	0.00
	9988.HK	阿里巴巴-SW	165.10	151.50	-8.24%	28,919.29	2.61	21.05	18.53	17.79	0.00
	0700.HK	腾讯控股	629.00	611.50	-2.78%	55,920.88	7.00	23.43	21.90	18.36	23.01
	80700.HK	腾讯控股-R	578.00	557.50	-3.55%	50,882.41	6.99	23.40	21.86	18.79	22.98
	9999.HK	网易-S	217.00	214.00	-1.38%	6,779.43	5.53	17.06	13.96	13.67	17.95
	9888.HK	百度集团-SW	117.60	113.80	-3.23%	3,129.69	2.19	33.97	14.48	8.91	22.44
	89888.HK	百度集团-SWR	107.90	103.50	-4.08%	2,847.71	2.18	33.84	14.50	9.05	22.38
	1024.HK	快手-W	72.35	67.80	-6.29%	2,930.20	1.93	15.41	32.60	10.72	15.12
	81024.HK	快手-WR	66.40	61.75	-7.00%	2,666.19	1.93	15.38	32.58	10.97	15.13
	9626.HK	哔哩哔哩-W	232.40	206.40	-11.19%	868.60	2.66	102.98	-7.41	-41.13	72.12
	2518.HK	汽车之家-S	49.06	45.86	-6.52%	233.61	3.15	14.58	12.86	14.02	13.65
	9898.HK	微博-SW	84.75	77.40	-8.67%	189.49	1.40	5.26	7.52	7.82	5.39
美 股 软 件 及 应 用	MSFT.O	微软 (MICRO -SOFT)	517.81	492.01	-4.98%	36,568.04	12.45	34.86	38.62	35.56	36.40
	SNOW.N	SNOWFLAKE	274.88	251.24	-8.60%	851.20	20.68	-61.69	-82.25	-60.96	0.00
	ORCL.N	甲骨文 (ORACLE)	262.61	201.95	-23.10%	5,757.18	9.75	46.28	34.08	44.53	0.00
	CRM.N	赛富时 (SALES -FORCE)	260.41	230.54	-11.47%	2,203.96	5.58	33.08	1224.61	77.36	0.00
	ADBE.O	奥多比 (ADOBE)	340.31	320.13	-5.93%	1,340.06	5.78	19.26	50.01	35.27	18.89
	INTU.O	财捷(INTUIT)	667.55	634.08	-5.01%	1,765.28	9.08	42.87	73.39	59.40	0.00
	SNPS.O	新思科技 (SYNOPSYS)	453.82	418.01	-7.89%	776.45	12.07	38.87	63.63	33.15	0.00
	CDNS.O	铿腾电子 (CADENCE)	338.69	311.84	-7.93%	848.83	16.28	80.01	71.17	77.96	0.00
	ADSK.O	欧特克 (AUTODESK)	301.34	303.34	0.66%	643.08	9.34	57.88	63.29	70.14	0.00
	U.N	Unity	37.90	42.52	12.19%	181.95	10.09	-41.78	-18.90	-13.85	-53.78
	0020.HK	商汤-W	2.37	2.12	-10.55%	819.89	17.03	-22.66	-5.46	-11.94	-27.70
	80020.HK	商汤-WR	2.17	1.93	-11.06%	746.02	17.00	-22.62	-5.46	-12.28	-27.58
	3888.HK	金山软件	33.86	28.64	-15.42%	401.08	3.71	24.58	61.69	26.83	22.97
	0268.HK	金蝶国际	14.66	14.12	-3.68%	501.18	6.95	-2,081.78	-178.41	-199.39	374.95

9878.HK	汇通达网络	13.37	10.01	-25.13%	56.31	0.10	18.13	34.57	36.70	13.73
3650.HK	KEEP	4.65	4.04	-13.12%	21.24	1.05	-4.76	6.00	-5.21	147.77
0354.HK	中国软件国际	5.82	5.23	-10.14%	142.89	0.74	24.01	22.39	25.60	19.00
1357.HK	美图公司	8.62	8.60	-0.23%	392.69	10.11	39.85	38.62	15.56	37.34
3896.HK	金山云	6.51	6.39	-1.84%	289.58	2.93	-27.17	-3.18	-10.68	-22.67
2013.HK	微盟集团	2.27	1.99	-12.33%	78.81	5.76	-5.94	-9.62	-5.91	-205.01
1675.HK	亚信科技	9.02	8.04	-10.86%	75.56	1.10	16.97	13.69	10.05	14.38
2121.HK	创新奇智	6.70	5.84	-12.84%	32.91	2.22	-6.60	-8.15	-5.13	-22.95
2400.HK	心动公司	67.45	70.65	4.74%	349.63	5.43	23.42	-53.48	14.09	19.80
0777.HK	网龙	12.14	11.34	-6.59%	60.25	1.07	-93.12	10.84	16.17	11.09

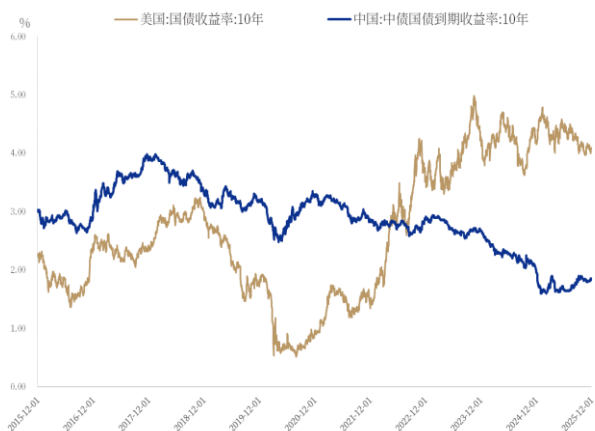
资料来源：WIND，中国银河证券研究院

（三）债市及汇率情况

11月28日，美国10年期国债利率为4.02%，相较11月3日累计下降11bps；中债10年期国债收益率为1.84%，相较11月3日累计上升4.68bps。

11月28日，美元兑人民币中间价报7.08，较11月3日价累计调升78个基点。

图3：国债收益率（10年期）



资料来源：WIND，中国银河证券研究院

图4：美元兑人民币汇率（中间价）



资料来源：WIND，中国银河证券研究院

二、前沿行业动态

（一）前沿政策动态

1. 美国启动“创世纪任务”，巩固美国在全球AI领域的技术主导权和战略领导力

11月24日，美国总统特朗普签署第14363号行政命令，启动“创世纪任务”（Genesis Mission）这一全国性专项协调行动，并于11月28日正式发布。该使命对标曼哈顿计划，旨在打造集成AI平台，整合美国联邦数十年积累的全球最大规模科学数据集，结合国家实验室、企业、高校

等资源，利用高性能计算、AI 模型与工具等基础设施，加速 AI 在科学发现中的应用。其核心目标是解决先进制造、生物技术、关键材料等至少 20 项国家重要科技挑战，进而强化国家安全、保障能源主导地位、提升劳动力生产率，巩固美国在全球 AI 领域的技术主导权与战略领导力。

“创世纪任务”将促进科学进步，加强国家安全，确保美国能源主导地位，提高劳动力生产力，并使纳税人在研发方面的投资回报倍增，从而促进美国的技术主导地位 and 全球战略领导地位。

图5：美国启动“创世纪任务”

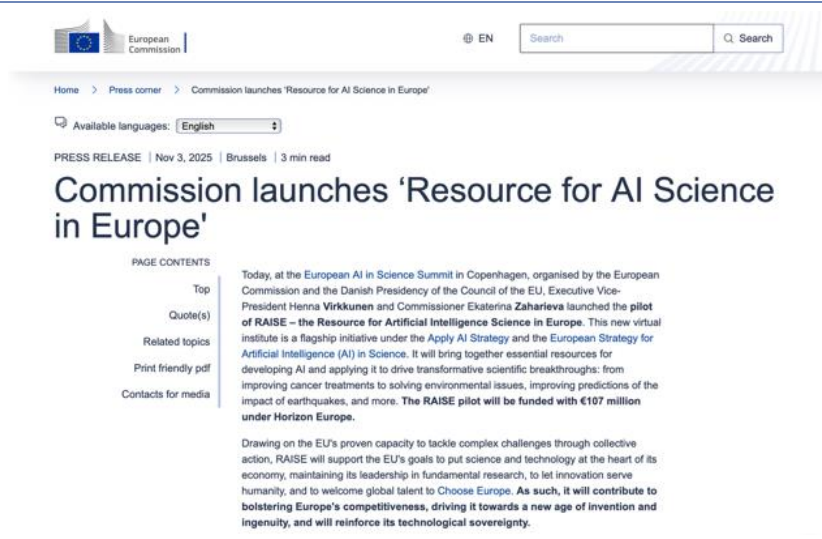


资料来源：美国联合公报官网，中国银河证券研究院

2. 欧盟启动“欧洲人工智能科学资源”计划，汇集核心资源推动前沿 AI 科学进步

11 月 3 日，欧盟委员会在欧洲人工智能科学峰会上启动“欧洲人工智能科学资源”（RAISE）计划。RAISE 是一个虚拟欧洲研究所，这是应用人工智能战略和欧洲人工智能科学战略下的旗舰计划。RAISE 汇集和协调了核心人工智能资源，包括欧盟成员国和私营部门的计算能力、数据、人才和研究资金，以推动前沿人工智能的发展和人工智能驱动的科学进步。RAISE 将通过集体行动来应对复杂挑战，支持欧盟将科学和技术置于经济核心位置，保持其在基础研究领域的领导地位，让创新服务于人类，并吸引全球人才。它将有助于提高欧洲的竞争力，推动其走向发明和独创性的新时代，并将加强其技术主权。RAISE 试点将获得 Horizon Europe 1.07 亿欧元的资助。

图6：欧盟启动“欧洲人工智能科学资源”计划



资料来源：欧盟官网，中国银河证券研究院

3. 欧盟批准捷克援助 Onsemi 建立集成芯片制造厂，加强半导体技术的自主权

11月21日，欧盟委员会根据欧盟国家援助规则，批准了捷克的一项4.5亿欧元（120亿捷克克朗）的援助措施，支持美国芯片制造商Onsemi在Rožnov pod Radhoštěm建立一个新型的碳化硅（“SiC”）电源设备集成芯片制造工厂，这对实现电动汽车、快速充电站、可再生能源发电和其他高效、紧凑、高性能电力电子产品至关重要。该援助将采取向Onsemi提供约4.5亿欧元的直接赠款的形式，以支持该公司16.4亿欧元的投资。预计到2027年，新设施将开始商业运营，这将是欧盟首创的设施。该工厂将使用创新的制造技术，并制造目前欧洲市场上没有的性能优异产品。根据《欧洲芯片法》和欧盟委员会2024-2029年政治指导方针中规定的目标，该措施将有助于增加欧盟在半导体技术方面的自主权，还将有助于加速数字化和绿色转型。

图7：欧盟批准捷克援助 Onsemi 新半导体制造设施建设



资料来源：欧盟官网，中国银河证券研究院

4. 上海静安区首个数据跨境服务中心揭牌，推动数据高效便利安全跨境流动

11月29日，上海中心城区首个数据跨境服务中心在静安区静安国际会客厅揭牌。作为上海“五个中心”建设的重要承载区，静安区为响应日益迫切的数据跨境需求，落实国家统筹发展与安全、推动数据高效便利安全跨境流动及数字经济发展战略，成立该中心。中心立足“政策枢纽”“服务平台”“创新前哨”三大定位，设置服务专窗并配备人员，为企业提供数据跨境流动和个人信息处理相关的政策解读、合规咨询、协助申报等“一站式”专业服务，助力企业降低合规成本、提升国际竞争力。活动现场还发布了涵盖服务折页、政策法规汇编等实用内容的“静安区数据跨境服务包”，后续将结合最新政策与企业关切持续迭代升级。

图8：上海静安区启用数据跨境服务中心



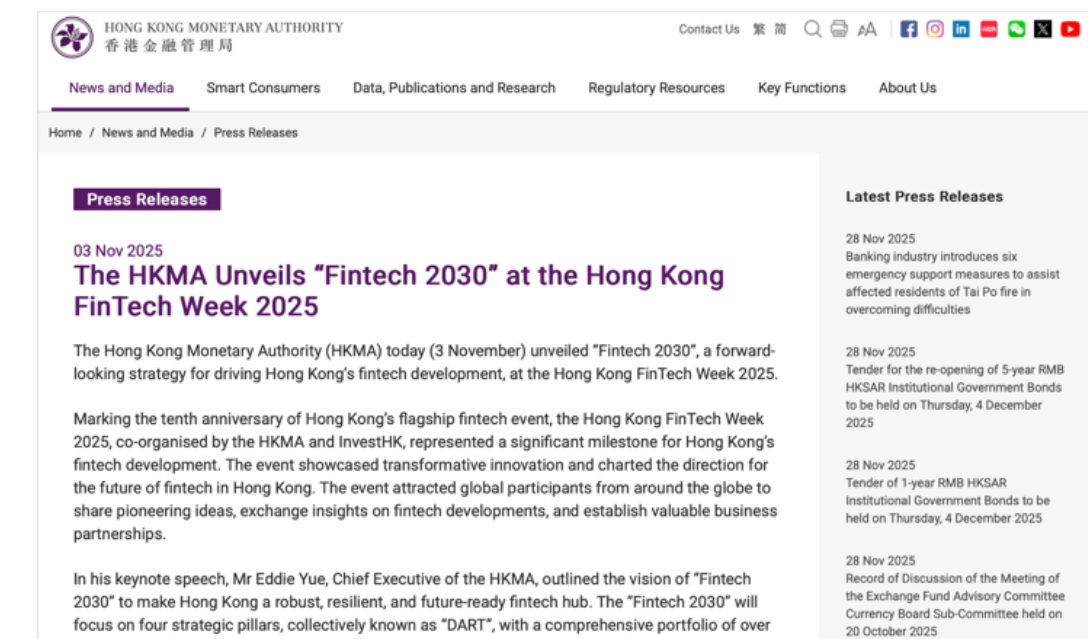
上海中心城区首个数据跨境服务中心揭牌

资料来源：上海市政府官网，中国银河证券研究院

5. 香港金管局公布“金融科技 2030”战略，借助 AI 等新技术构建持续共融金融科技生态

11月3日，香港金融管理局在由其与投资推广署合办的香港金融科技周 2025 上公布“金融科技 2030”战略。该战略以“重塑金融蓝图，助力香港未来”为愿景，旨在将香港打造为稳健、有韧性且具前瞻性的国际金融科技枢纽，聚焦 DART 四大重点领域，涵盖超 40 个具体项目：构建新一代数据及支付基建（Data and Payment Infrastructure），支援安全高效的数据共享与跨境支付互联；支持业界全方位应用人工智能（Artificial Intelligence），推出 AI 策略，推动金融业全面负责任应用人工智能，构建共享可扩展的 AI 基础设施及专用模型；强化业务及科技韧性，并为量子电脑时代做好准备（Resilience），制定金融科技网络安全新认证架构、建立预警系统，并为量子时代布局抗量子能力基础设施；促进金融代币化（Tokenisation），推进现实世界资产代币化，恒常化代币化政府债券发行、探讨外汇基金票据及债券代币化可行性，推出 Ensemble 项目试点计划。金管局总裁余伟文强调，将依托公私营及跨境合作，推动构建更可持续、共融的金融科技生态。

图9：香港金融管理局公布“金融科技 2030”战略



资料来源：香港金管局官网，中国银河证券研究院

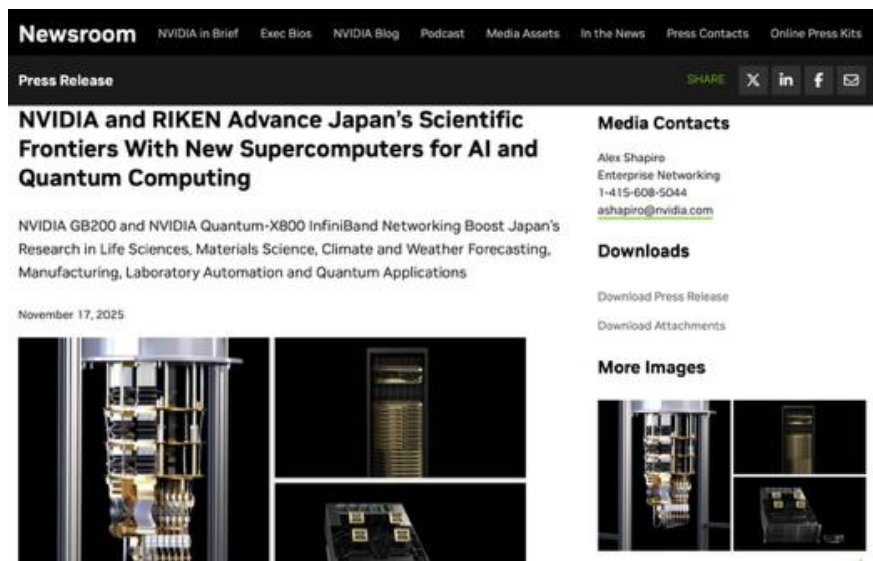
（二）前沿技术动态

1. 英伟达和 RIKEN 合作，以新人工智能和量子计算超级计算机促进日本科学发展

11 月 17 日，英伟达宣布与日本顶尖国家研究机构理化学研究所（RIKEN）达成合作，推出两台全新超级计算机，均搭载 NVIDIA GB200 NVL4 系统及 Quantum-X800 InfiniBand 网络技术，旨在助力日本推进科学 AI 与量子计算领域发展，支持日本人工智能战略，并扩大国内安全基础设施，在科学、工业和技术领域发挥领导作用。其中一台配备 1600 块 NVIDIA Blackwell GPU，聚焦科学 AI 研究，将应用于生命科学、材料科学、气候天气预报、制造业及实验室自动化等领域；另一台搭载 540 块同款 GPU，专注量子计算，将加速量子算法、混合模拟及量子经典计算方法的研究。这两台新超级计算机预计 2026 年春季投入运行，此次合作既是 NVIDIA 与 RIKEN 合作的进一步拓展，也将支撑日本的主权 AI 战略，助力其在科学、工业和技术领域构建领先的本土安全基础设施。

这两台超级计算机还将作为 FugakuNEXT 超级计算机的代理机，用于相关软硬件及应用的协同设计与开发，而 FugakuNEXT 计划采用富士通 MONAKA-X CPU 与英伟达技术结合，预计 2030 年投入使用，应用性能将较现有 CPU 基超级计算机提升 100 倍，未来还将整合量产级量子计算机。此外，双方还将合作开发浮点仿真软件，并借助 NVIDIA CUDA-X 平台的优化库等工具提升高性能计算应用效能。

图10: 英伟达为 RIKEN 提供系统支持



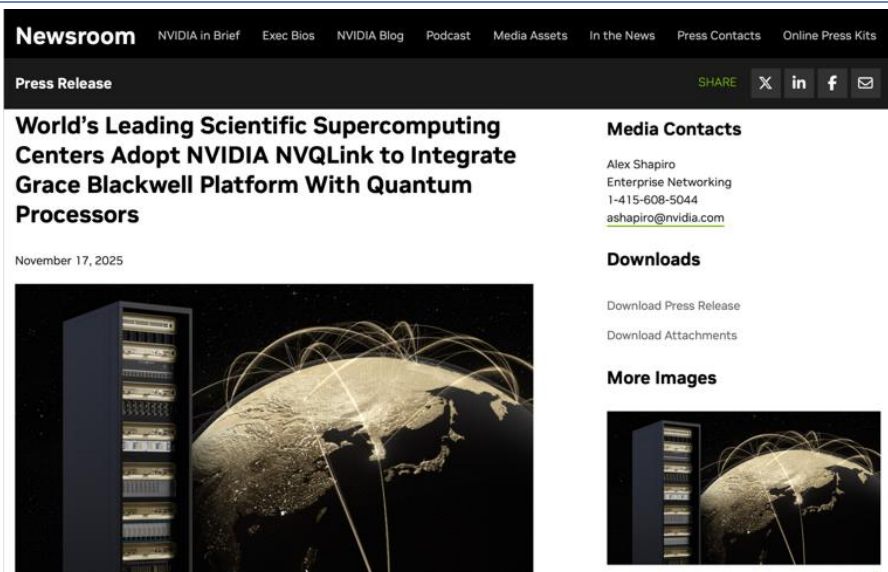
资料来源: 英伟达官网, 中国银河证券研究院

2. 世界顶尖科学超级计算中心采用英伟达技术, 推动量子计算领域发展

11月17日, 英伟达宣布全球顶尖科学超级计算中心将广泛采用其首创的通用互连技术 NVIDIA NVQLink, 该技术可将 Grace Blackwell 平台与量子处理器整合, 助力量子计算领域发展。NVQLink 采用开放系统架构, 通过 NVIDIA CUDA-Q 软件平台实现量子处理器与 GPU 超级计算的紧密集成, 具备低延迟、高吞吐量的优势, 可提供 40 petaflops 的 FP4 精度 AI 性能, 能攻克量子计算中的控制与纠错难题, 支持大规模量子经典混合工作流开发。

目前, 亚洲、欧洲、中东、美国等地超十几家超级计算中心及国家研究机构已确认采用该技术, 包括日本理化学研究所、韩国科学技术信息研究所、意大利 CINECA、德国于利希超级计算中心、美国布鲁克海文国家实验室等。量子计算公司 Quantinuum 已借助 NVQLink 将其最新 Helios 量子处理器与 NVIDIA GPU 整合, 成功实现全球首个可扩展的实时量子纠错解码, 反应速度远超需求; 此外, NVQLink 支持以太网扩展, 便于研究人员根据量子处理器和应用的拓展灵活扩容经典计算资源。

图11: 世界领先的科学超级计算中心采用 NVIDIA NVQLink 将 Grace Blackwell 平台与量子处理器集成



资料来源: 英伟达官网, 中国银河证券研究院

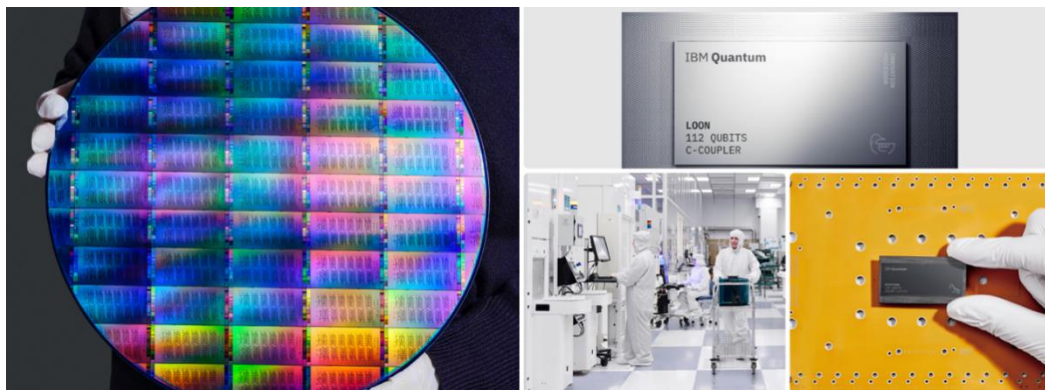
3. IBM 在量子计算领域实现处理器、软件和算法突破

11 月 12 日，IBM 在年度量子开发者大会上公布了量子计算领域的多项关键进展，朝着 2026 年底实现量子优势、2029 年达成容错量子计算的目标稳步推进。

硬件方面，IBM 推出最先进的量子处理器 IBM Quantum Nighthawk，配备 120 个量子比特和 218 个下一代可调耦合器，量子比特连接性提升 20% 以上；可支持复杂度提高 30% 的电路执行；该架构将使用户能够探索更苛刻的计算问题，这些问题需要多达 5000 个双量子比特门。预计 2028 年可支持 1.5 万个双量子比特门。软件与算法层面，量子软件栈 Qiskit 新增动态电路功能，100+ 量子比特规模下精度提升 24%，通过 HPC 加速的误差缓解技术使精准结果提取成本降低超 100 倍，还将推出 C++ 接口及机器学习、优化等领域的计算库。

同时发布实验性处理器 IBM Quantum Loon，首次集齐容错量子计算所需的全部关键组件，为高效量子纠错的实现与扩展奠定基础。IBM 已证明提前一年实现基于 qLDPC 码的实时量子纠错解码（延迟不足 480 纳秒），这将与 Loon 一起展示了 IBM 量子计算机的核心。此外，IBM 将量子处理器晶圆制造迁移至纽约 300 毫米晶圆制造设施，研发速度翻倍，量子芯片物理复杂度提升 10 倍，且，进一步加速量子计算规模化发展。

图12: IBM Quantum Nighthawk 处理器、IBM Quantum Loon 处理器和量子晶圆

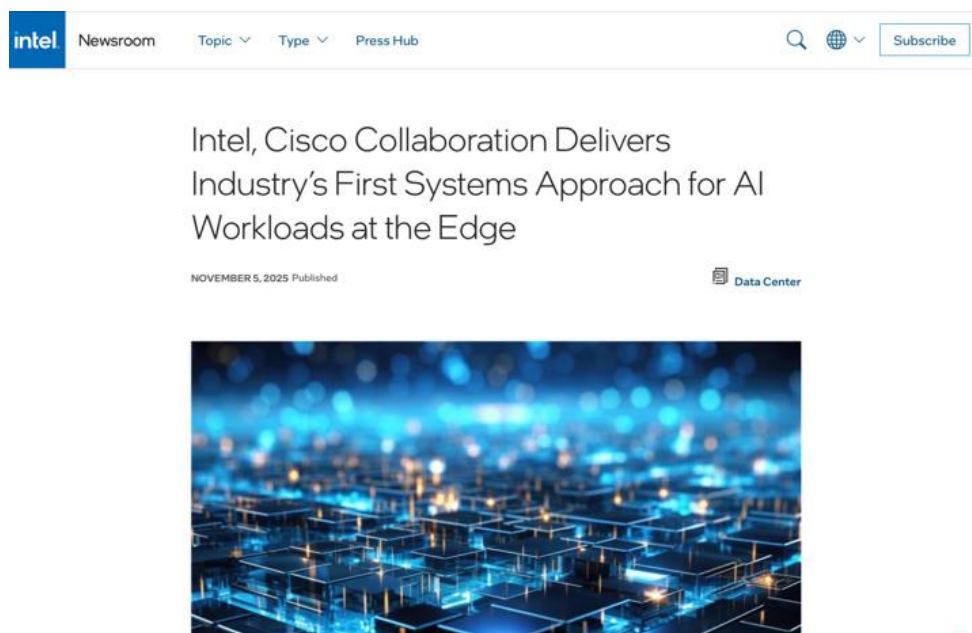


资料来源：IBM 官网，中国银河证券研究院

4. 英特尔和思科合作推出面向未来的 AI 基础设施，在提高性能和安全性的同时减少网络流量

当地时间 11 月 5 日，英特尔与思科宣布合作推出业界首个面向边缘 AI 工作负载的集成系统方案——搭载英特尔至强 6（Xeon 6）系统级芯片（SoC）的思科统一边缘（Cisco Unified Edge）平台。该方案秉持软硬件与开放生态深度融合的系统级 AI 基础设施理念，将计算、网络、存储和安全功能部署于数据生成源头的边缘端，可支持实时 AI 推理及智能体工作负载，提供一个统一、安全和可扩展的基础设施，旨在处理未来十年的复杂人工智能工作负载，支持在最需要的地方产生实时智能。借助统一边缘，客户可以最大限度地提高其基础设施的整体灵活性，从而高效且具有成本效益地实施广泛的人工智能应用程序。同时依托双方广泛的软件开发者生态，支持预验证应用快速部署以缩短价值实现周期，助力企业从数据中心到边缘端全方位拓展性能、简化操作并增强信任，满足日益增长的智能体及物理 AI 工作负载需求。

图13：英特尔、思科合作为边缘人工智能工作负载提供了业界首强的系统方法



资料来源：英特尔官网，中国银河证券研究院

三、前沿产品动态

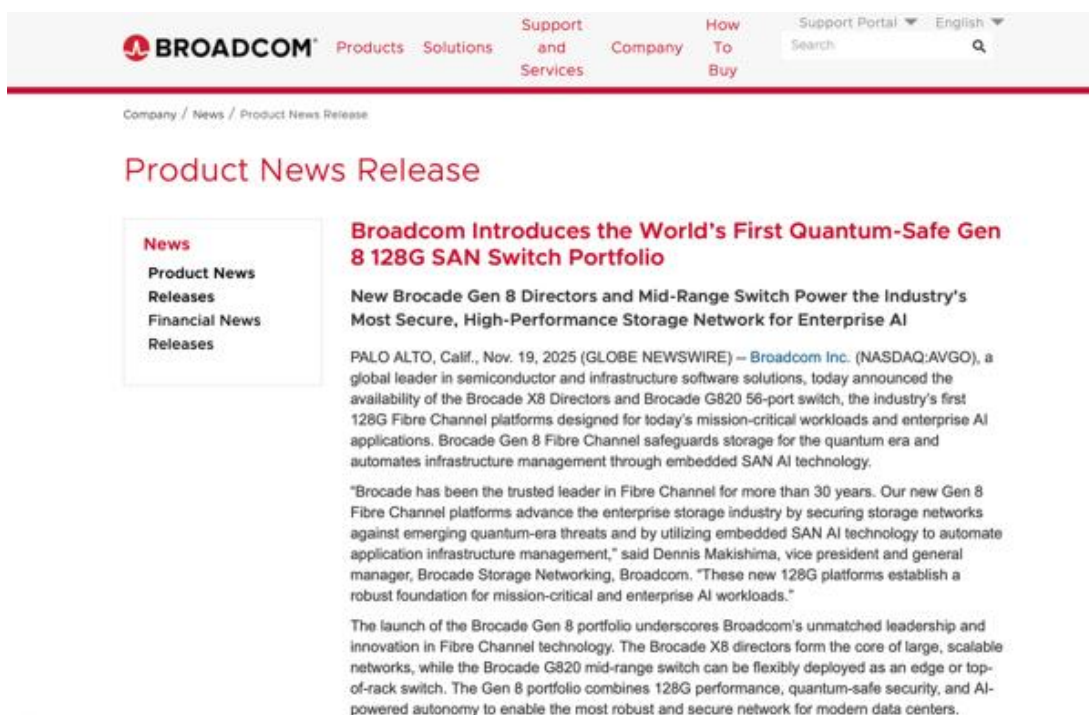
（一）算力及终端

1. 博通推出业界首款量子安全第 8 代 128G SAN 交换机产品，为企业 AI 存储网络提供动力

当地时间 11 月 19 日，博通宣布推出业界首款量子安全的第 8 代 128G SAN 交换机产品组合，包括 Brocade X8 Director 和 Brocade G820 56 端口交换机。该产品组合基于 128G 光纤通道平台打造，专为关键任务工作负载和企业 AI 应用设计，融合了超高性能、量子安全防护与 AI 驱动的自主管理能力，为现代数据中心构建稳健且安全的存储网络。

其中，Brocade X8 Director 为模块化平台，最多支持 384 个 128G 端口，具备高带宽、超低延迟特性，可满足大规模存储环境的扩展与整合需求；Brocade G820 交换机为 1U 紧凑型设计，配备 56 个 128G SFP+ 端口，兼顾高性能与空间利用率，适用于边缘或机架顶部部署。产品采用 256 位加密及后量子密码算法实现量子安全，通过最小权限架构强化访问控制以缩减攻击面；嵌入的 SAN AI 技术可实现基础设施自动化管理，SAN Fabric Intelligence 提供端到端可视性以加速故障排查，目前两款产品均已上市，OEM 合作伙伴将在未来几个季度陆续出货，戴尔科技、H3C、IBM、联想、NetApp 等众多行业伙伴均对该产品组合表示支持，其与各类存储系统的协同将为企业 AI 训练、数据密集型应用等场景提供可靠高效的基础设施支撑。

图14: 博通推出全球首款量子安全第 8 代 128G SAN 交换机产品组合

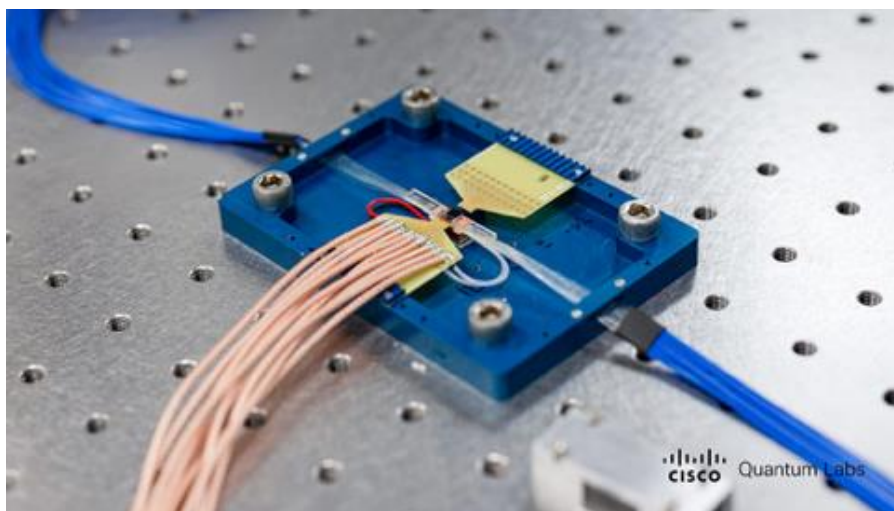


资料来源：博通官网，中国银河证券研究院

2. IBM 与思科合作打造大规模容错量子计算机网络，主力突破量子计算规模化瓶颈

11 月 20 日，IBM 与思科宣布达成合作，计划携手打造大规模容错量子计算机网络，目标最早于 2030 年代初实现，为 21 世纪 30 年代末量子计算互联网奠定基础。双方将结合 IBM 在量子计算机研发领域的领先优势与思科的量子网络创新技术，突破量子计算规模化瓶颈，解决量子计算互联网构建的核心难题。根据规划，未来五年内将完成首个概念验证演示，实现多台位于不同低温环境的大规模容错量子计算机联网，支持数万至数十万个量子比特协同运算，处理需万亿级量子门的复杂任务，如大规模优化、复杂材料及药物设计等。为达成这一目标，双方将研发新型连接技术、配套软件栈及量子网络架构，通过 IBM 打造的量子网络单元（QNU）实现量子信息在量子处理单元（QPU）与网络间的转换传输，借助思科的高速软件协议框架动态配置网络路径；同时探索长距离量子信息传输技术，计划先实现数据中心内多台 QPU 互联，未来拓展至跨数据中心联网。此外，IBM 还将与费米国家加速器实验室领导的超导量子材料与系统中心（SQMS）合作，计划三年内演示多台互联 QPU，并联合思科共同资助学术研究及合作项目，助力量子生态发展。该合作有望推动超安全通信、精准气候监测等新型应用落地，拓展量子计算的应用边界。

图15: IBM 和思科宣布计划建立一个大型容错量子计算机网络

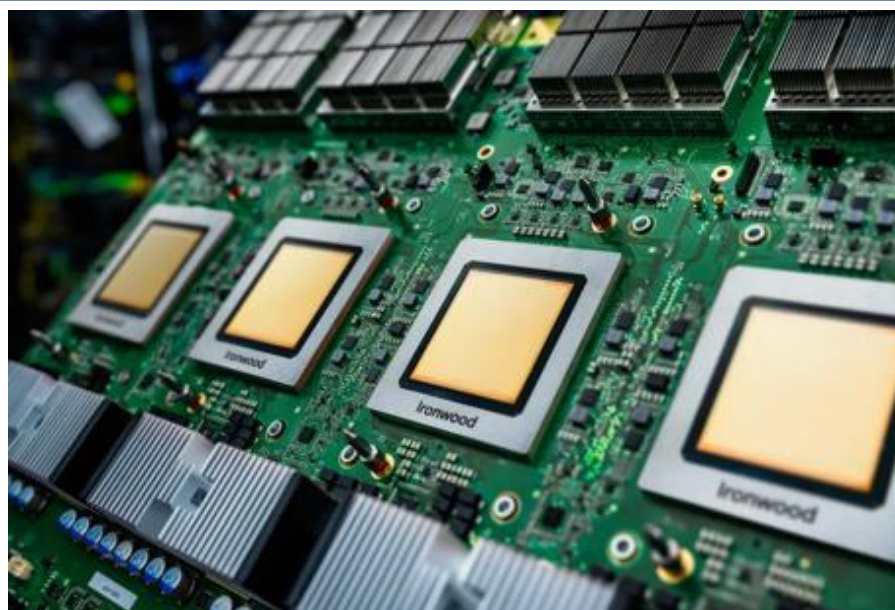


资料来源: 思科官网, 中国银河证券研究院

3. 谷歌宣布第七代张量处理单元 Ironwood 正式面向云客户服务

当地时间 11 月 25 日, 谷歌宣布其第七代张量处理单元 (TPU) Ironwood 正式面向云服务客户推出, 该产品是谷歌迄今为止性能最强、能效最高的定制芯片, 专为大规模运行具备思考和推理能力的 AI 模型打造。Ironwood 聚焦 AI 推理时代需求, 针对高容量、低延迟的 AI 推理及模型部署场景量身定制, 单芯片在训练和推理工作负载上的性能较上一代提升超 4 倍; 作为谷歌 AI 超级计算机的核心组件, 它可组成超大规模集群, 单个超级节点能集成 9216 块芯片, 通过速率达 9.6 Tb/s 的突破性芯片间互连 (ICI) 网络连接, 共享 1.77PB 的高带宽内存 (HBM), 有效突破数据瓶颈, 极大降低培训和运行尖端人工智能服务所需的计算时间和能量。Ironwood 是谷歌连续循环的结果, 研究人员影响硬件设计, 硬件加速研究。谷歌 DeepMind 与 TPU 工程师深度协作, 且借助 AlphaChip 技术, 为包括 Ironwood 在内的过去三代 TPU 生成卓越的布局。

图16: 谷歌第七代张量处理单元 (TPU) Ironwood



资料来源: 谷歌官网, 中国银河证券研究院

4. Q.ANT 公司宣布推出第二代光子处理器 Q.ANT NPU 2，为 AI 和 HPC 提供动力

当地时间 11 月 18 日，Q.ANT 公司宣布推出第二代光子处理器 Q.ANT NPU 2，该产品具备增强的非线性处理能力，旨在为人工智能和高性能计算提供支持。作为一款原生处理单元，Q.ANT NPU 2 依托光子计算技术，通过光进行原生模拟计算，相比传统 CMOS 芯片，可减少多达 30 倍的能耗，同时将复杂 AI 和 HPC 工作负载的性能提升 50 倍。与传统的基于 CPU 的系统相比，Q.ANT 的光子处理器以更少的参数和更少的操作实现更准确的结果，展示现有服务器架构中的现实世界光子加速。

产品以一体化 19 英寸机架式服务器形式交付，内置多块第二代 NPU，可通过 PCIe 接口及 C/C++/Python API 与现有 CPU、GPU 无缝集成，便于在 HPC 和数据中心环境中快速部署。

图17: Q.ANT 第二代原生处理器 NPU 2

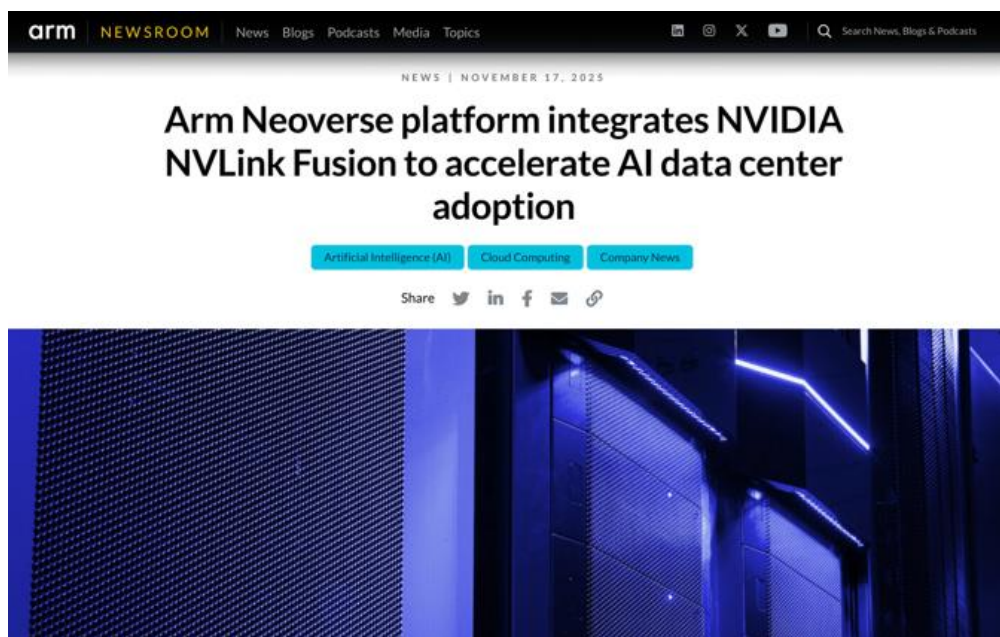


资料来源: Q.ANT 官网, 中国银河证券研究院

5. ARM 与英伟达深化合作，助力 AI 数据中心加速落地

当地时间 11 月 17 日，ARM 宣布其 Neoverse 平台将集成 NVIDIA NVLink Fusion 技术，深化双方在 AI 领域的合作，助力 AI 数据中心加速落地。ARM Neoverse 作为专为高能效、高性能扩展打造的计算平台，已部署超 10 亿个核心，2025 年有望在全球顶级超大规模数据中心占据 50% 市场份额，AWS、谷歌、微软等主流厂商及 Stargate 等下一代 AI 数据中心项目均以其为核心计算平台。此次集成后，生态合作伙伴可将高效的 ARM 架构计算融入 NVIDIA NVLink Fusion 生态，借助 AMBA CHI C2C 协议实现 CPU 与加速器之间连贯且高带宽的连接，消除人工智能系统性能的内存和带宽瓶颈，获得与 NVIDIA Grace Blackwell 平台同级别的性能表现。这一合作既为客户在 ARM 平台的工作负载加速连接方面提供了更多选择，满足 AI 模型和工作负载指数级增长下对能效的严苛需求，推动下一代专业化 AI 基础设施的创新设计，彰显了双方在 AI 时代架构转型中的协同引领作用。

图18: ARM 宣布 Neoverse 平台集成 NVIDIA NVLink Fusion 技术



资料来源: ARM 官网, 中国银河证券研究院

6. 高通推出骁龙 8 Gen 5 移动平台, 为旗舰手机提供卓越性能支撑

当地时间 11 月 26 日, 高通技术公司宣布推出骁龙 8 Gen 5 移动平台, 该平台作为高端产品线的新成员, 以卓越性能和前沿技术重塑 8 系列旗舰体验, 为全球更多消费者带来旗舰级智能与使用感受。核心配置上, 其搭载定制化高通 Oryon CPU, 峰值速度可达 3.8 GHz, CPU 性能提升 36%, 网页浏览响应速度提升 76%; Adreno GPU 采用创新切片架构, 游戏和图形性能提升 11%; Hexagon NPU 性能提升 46%, 搭配高通感知中枢和 AI 引擎, 可实现拿起设备即唤醒 AI 助手、智能感知用户语音意图, 还能支持智能体 AI 助手提供情境化交互与个性化建议, 打造流畅的端侧智能体验。据悉, iQOO、荣耀、魅族、摩托罗拉、vivo 等全球主流 OEM 厂商的旗舰机型将搭载该平台。

图19: 高通骁龙 8 Gen5 移动平台

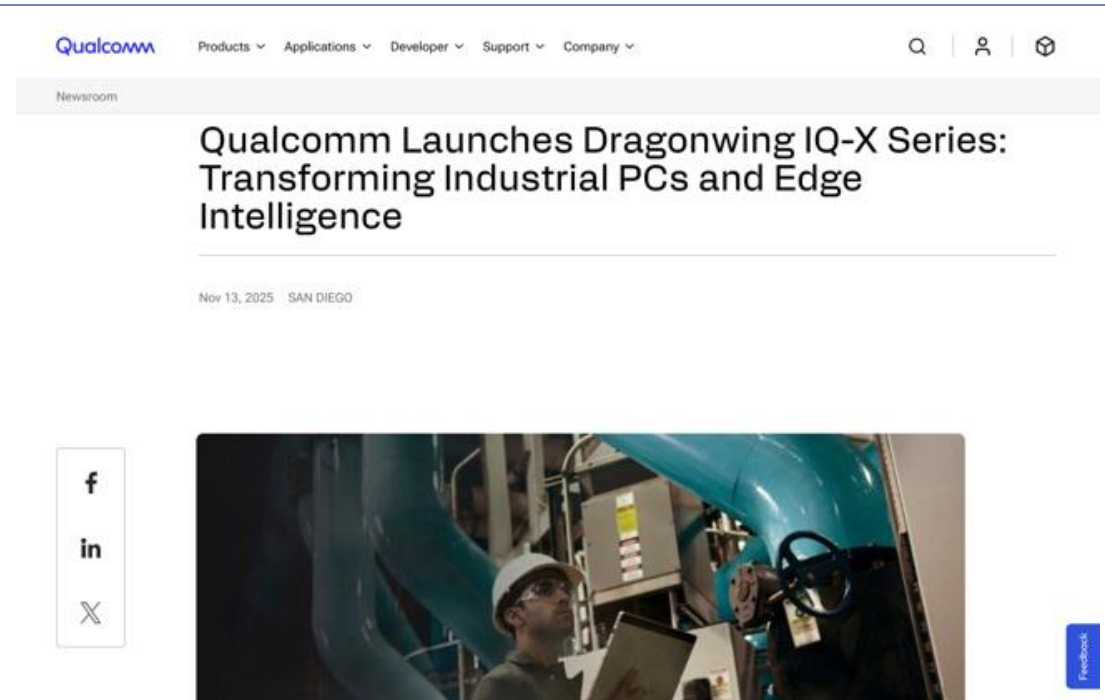


资料来源: 高通官网, 中国银河证券研究院

7. 高通宣布推出 Dragonwing™ IQ-X 系列工业级处理器，满足现代工业需求

当地时间 11 月 13 日，高通宣布推出 Qualcomm Dragonwing™ IQ-X 系列工业级处理器，将一流的单线程和多线程计算与更高效的边缘智能相结合，旨在变革工业计算与边缘智能领域。该系列处理器基于先进 4 纳米工艺的定制化 Qualcomm Oryon CPU，具备一流的单线程和多线程计算性能，核心数可在 8 至 12 核间灵活配置，AI 性能最高达 45 TOPS，且支持 -40°C 至 105°C 的工业级温度范围，适配恶劣工业环境。其支持工业标准 COM 模块规格，可无缝替换现有载板，同时提供丰富的外设支持与 Windows 11 IoT Enterprise LTSC 系统及 Qt、CODESYS 等多款工业软件兼容，还能通过高通 AI 软件栈及 ONNX、PyTorch 等通用运行时，助力预测性维护、缺陷检测等工业 AI 应用的模型移植与开发。该系列专为快速定制和可扩展性而设计，简化了坚固系统的设计，并通过消除对外部人工智能或多媒体模块的需求来减少物料清单，为原始设备制造商（OEM）和原始设计制造商（ODM）缩短上市时间，研华、congatec、新汉、宝威等领先厂商已率先采用，相关商用设备预计将于未来数月推出。

图20：高通宣布推出 Qualcomm Dragonwing™ IQ-X 系列工业级处理器



资料来源：高通官网，中国银河证券研究院

（二）大模型与云应用

1. 谷歌发布新一代 AI 模型 Gemini 3 系列，推理能力大幅提升

11 月 18 日，谷歌正式发布新一代 AI 模型 Gemini 3 系列，整合了 Gemini 系列核心能力，旨在全方位赋能用户学习、创作与规划各类事务，推动 AI 向高效思维伙伴升级。该模型以顶尖推理能力为核心，Gemini 3 Pro 在多个权威基准测试中大幅领先前代及竞品，LMarena Elo 得分达 1501，Humanity's Last Exam 达 37.5%，GPQA Diamond 达 91.9%，数学领域 MathArena Apex 创下 23.4% 的新纪录，多模态方面 MMMU-Pro 达 81%、Video-MMMU 达 87.6%，同时具备 100 万 token 长上下文处理能力和 72.1% 的事实准确性。这意味着 Gemini 3 Pro 具有高度可靠性，能够解决科学和数学等大量主题的复杂问题。

其新增 Deep Think 模式可进一步强化复杂问题推理，目前面向安全测试人员开放，后续将推

给 Google AI Ultra 订阅用户；还支持多模态信息整合、智能编码、长周期规划等功能，同步推出开发者平台 Google Antigravity，实现端到端软件任务自主规划与验证。Gemini 3 现已在 Gemini 应用、AI Studio、Vertex AI、搜索 AI 模式等多平台上线，第三方 IDE 如 GitHub、JetBrains 等也已支持集成，企业用户可通过 Gemini Enterprise 获得专属服务。

图21：谷歌 Gemini 3 模型



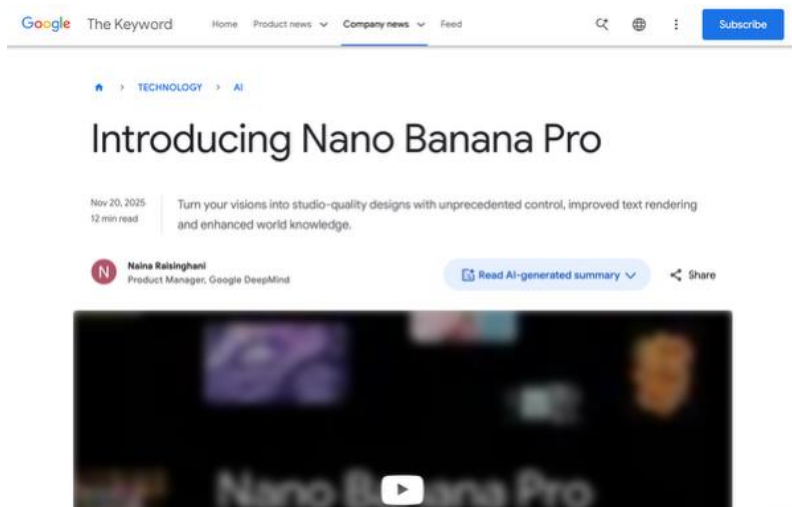
资料来源：谷歌官网，中国银河证券研究院

2. 谷歌发布 Nano Banana Pro，提供更优质的图像创作体验

11 月 20 日，谷歌 DeepMind 推出基于 Gemini 3 Pro 的新一代图像生成与编辑模型 Nano Banana Pro，该模型在原有 Nano Banana（Gemini 2.5 Flash Image）基础上升级，凭借 Gemini 3 Pro 的先进推理能力、丰富世界知识及实时信息接入能力，为用户提供更优质的创作体验。其核心优势包括：能生成包含多语言清晰文本的精准视觉内容，适配原型设计、海报、国际版素材等场景；具备设计风格一致性、高级创意控制等功能，可融合最多 14 张输入图像，保持 5 个人物的一致性；还能实现局部编辑、相机角度调整、色彩分级、光影转换等专业操作。

目前，Nano Banana Pro 已在 Gemini 应用、Google AI Studio、NotebookLM 等多款产品上线，不同用户群体可根据订阅等级获得相应使用配额。所有由该模型生成的媒体内容均嵌入 SynthID 隐形水印，普通用户可通过 Gemini 应用验证图像是否为谷歌 AI 生成，免费及 Google AI Pro 层级用户生成的图像还带有可见的 Gemini 标识水印，Ultra 订阅用户及开发者工具生成的图像则无可见水印，兼顾透明度与专业使用需求。

图22：谷歌推出 Nano Banana Pro 图像生成和编辑模型



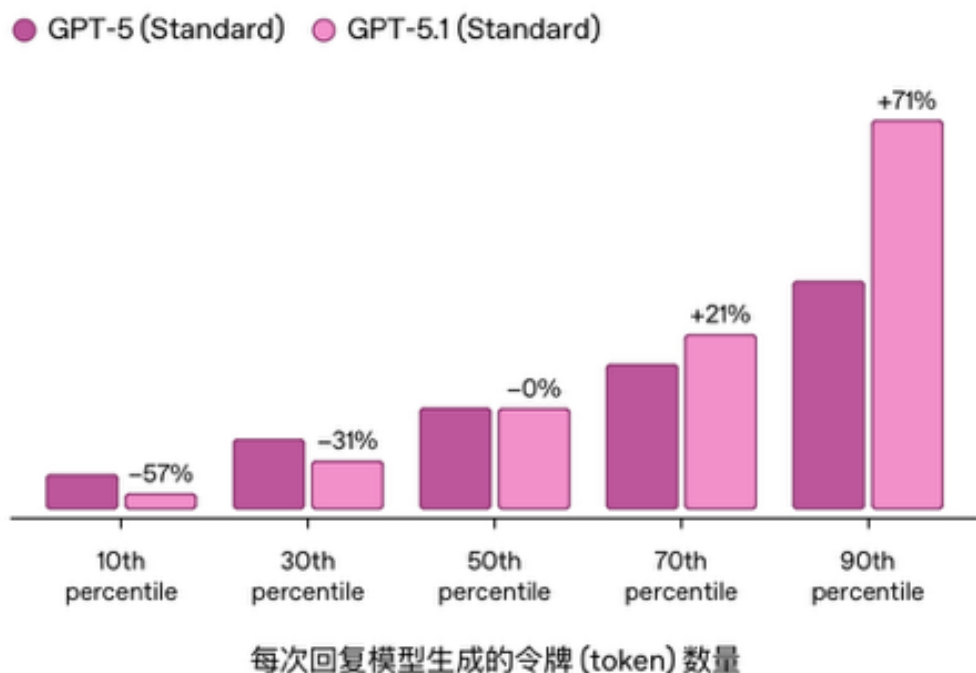
资料来源：谷歌官网，中国银河证券研究院

3. OpenAI 发布 ChatGPT 5.1，优化 ChatGPT 自定义功能，提升智能水平和对话体验

11月12日, OpenAI 正式推出 GPT-5 系列升级版本 GPT-5.1, 包含 GPT-5.1 Instant 和 GPT-5.1 Thinking 两款模型, 旨在提升智能水平与对话体验, 同时优化 ChatGPT 的自定义功能。

GPT-5.1 Instant 作为广泛使用的基础模型, 默认语气更温暖具对话感, 新增自适应推理能力, 可根据问题难度动态调整思考流程, 在数学、编程等评测中表现提升, 指令遵循更可靠; GPT-5.1 Thinking 则针对复杂任务优化, 能动态调整思考时间, 简单任务响应速度翻倍, 复杂任务答案更全面, 回复更清晰易懂且富有同理心。此次更新同步升级了 ChatGPT 的个性化设置, 保留并优化原有 3 种语气预设, 新增“专业可靠”“直言不讳”“天马行空”3 种选项, 还推出实验性性格微调功能, 支持调整回复简洁度、温暖感等细节, 设置更改可即时同步至所有对话。GPT-5.1 先向付费用户开放, 后续扩展至免费用户及未登录用户, Enterprise 和 Edu 套餐将有七天提前体验窗口期, 后续将会接入 API; GPT-5 系列模型将保留三个月过渡期供用户适应, 未来 OpenAI 还将持续迭代优化。

图23: ChatGPT 5.1 在思考时间上的动态调整比 ChatGPT 5 更灵活

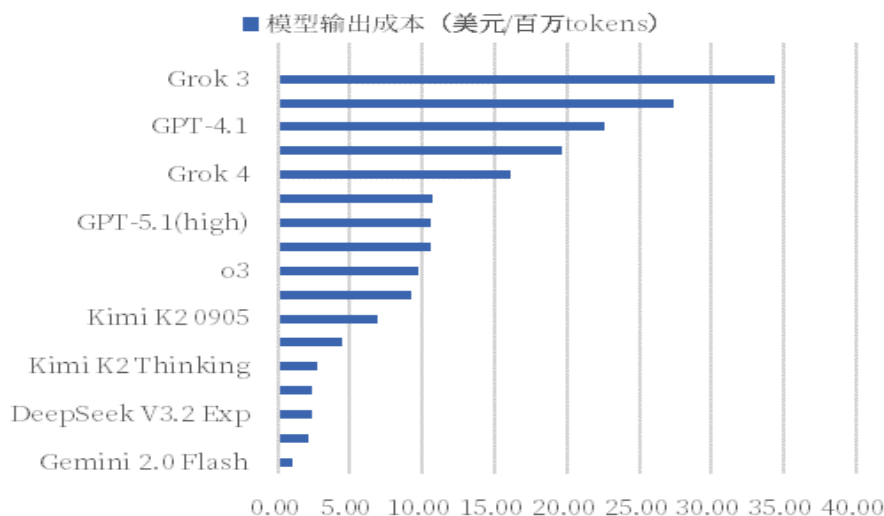


资料来源: OpenAI 官网, 中国银河证券研究院

基于我们发布的 2026 年计算机行业年度策略报告, 根据 Artificial Analysis 给出的 Artificial Analysis Intelligence Index (大模型智能化指数) 标准测算主流大模型, 在给定指数任务中, 观测主流大模型的 output tokens 消耗量和总运行成本。若以此作为定量假设进行测算, 可得出在给定任务中, 不同大模型所需的单位运行成本。据我们测算, 目前主流大模型的输出成本大部分集中在 2 美元/百万 tokens 至 25 美元/百万 tokens 区间内。

以 GPT-5.1 为例, 在给定指数测算过程中, output tokens 约为 80.9 百万 tokens, 总运行成本约为 859.06 美元, 因此估算 GPT-5.1 单位运行成本约为 10.62 美元/百万 tokens。

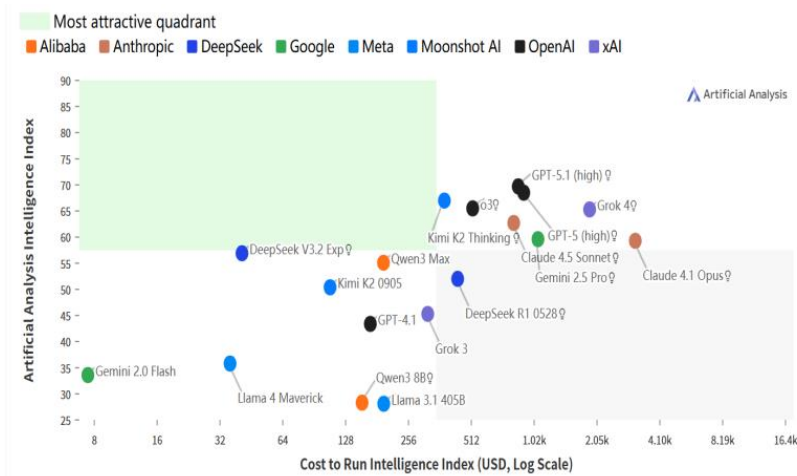
图24: 主流大模型单位运行成本测算 (美元/百万 tokens)



资料来源: Artificial Analysis, 中国银河证券研究院

模型高智能化短期暂时还无法规避成本高问题。基于上述模型,进行其智能化指数与模型运行人工智能指数时产生成本的关系,散点基本分布在“高智能-高运行成本”或“低智能-低运行成本”象限,可见模型实现“高智能化”暂时无法与“低运行成本”同时兼得。当前 AI Agent 本质是技术成熟度与商业化节奏的阶段性错配,预计 2026 年加速落地趋势不变。

图25: 主流大模型智能化指数 vs.运行大模型智能化指数所需成本(美元)对数关系



资料来源: Artificial Analysis, 中国银河证券研究院

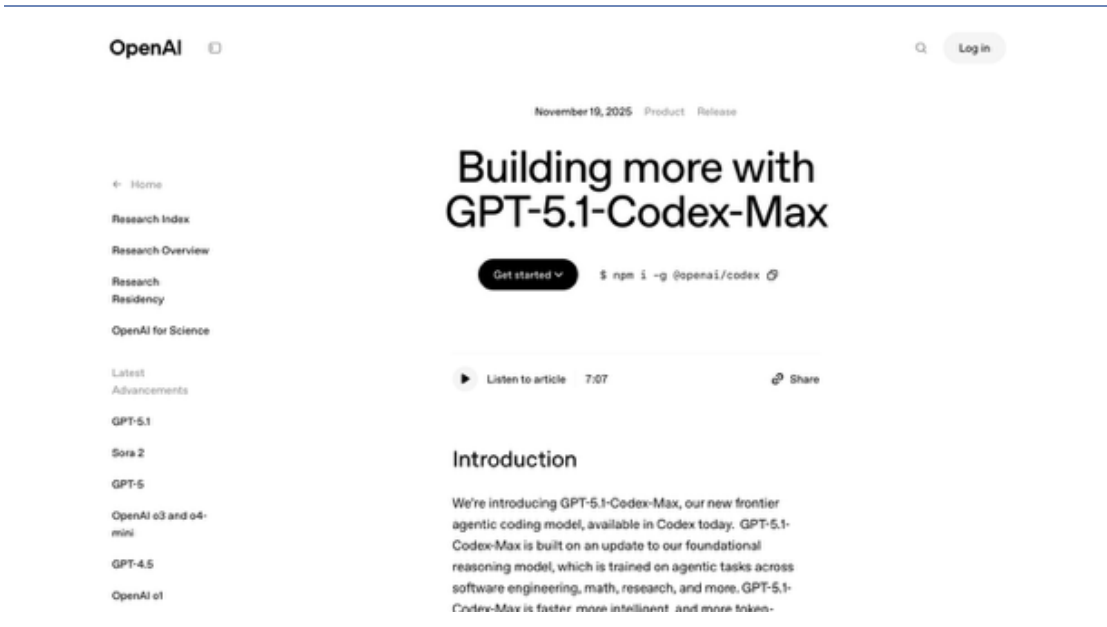
4. OpenAI 推出前沿智能编码模型 GPT-5.1-Codex-Max, 助力工程生产力的提升

当地时间 11 月 19 日, OpenAI 正式推出基于基础推理模型升级的前沿智能编码模型 GPT-5.1-Codex-Max, 该模型专为长时间、精细化编码任务设计, 现已在 Codex 平台上线, API 访问权限即将开放。它训练涵盖 PR 创建、代码审查等真实软件工程任务, 首次支持 Windows 环境, 在 SWE-bench Verified 等多项基准测试中超越前代, 通过一个叫做 compaction 的过程在多个上下文窗口中运行, 中等推理强度下令牌使用量减少 30%并新增超高推理选项, 同时默认运行于安全沙箱, 生成终端日志辅助审核, 采取迭代部署模式防范网络安全能力滥用风险 ChatGPT Plus、Pro 等套餐用户可在 Codex 的 CLI、IDE 扩展等功能中使用, 现已取代 GPT-5.1-Codex 成为平台默认模

型，内部数据显示其能显著提升开发效率。

GPT-5.1-Codex-Max 显示了模型在维持长平线编码任务、管理复杂工作流程以及用更少的令牌生成高质量实现方面取得的成果。该模型与 CLI、IDE 扩展、云集成和代码审查工具的稳步升级相结合，从而提高了工程生产力。

图26: OpenAI 推出前沿智能编码模型 GPT-5.1-Codex-Max

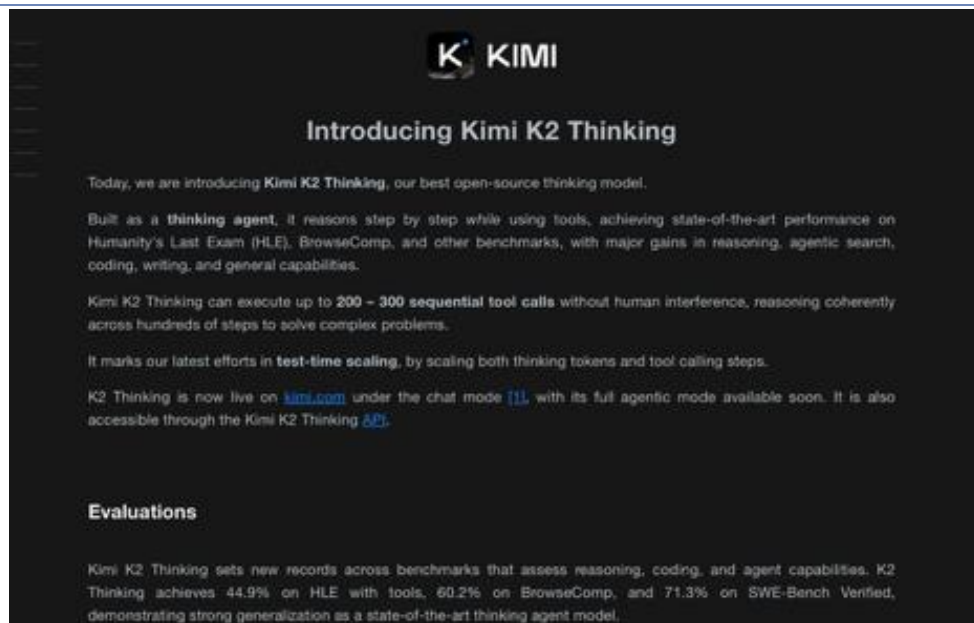


资料来源: OpenAI 官网, 中国银河证券研究院

5. 月之暗面全面开源 Kimi K2 Thinking, 在多项基准测试中表现优异

11月11日, 月之暗面开源思考型大模型 Kimi K2 Thinking, 该模型以思维代理为核心定位, 能在使用工具的同时分步推理, 可执行 200-300 次连续工具调用, 连贯解决复杂问题, 目前已在 kimi.com 聊天模式上线, 全代理模式及 API 访问后续将开放。它在多项基准测试中表现优异, 工具辅助下 Humanity's Last Exam 得分 44.9%、BrowseComp 得分 60.2%、SWE-Bench Verified 得分 71.3%, 在推理、智能搜索、编码等领域实现大幅提升, 可解决博士级数学问题、完成复杂编码任务及深度网页搜索推理。通用能力方面, K2 Thinking 在创意写作更生动有深度, 实用写作逻辑严谨、贴合指令, 回应个人情感问题时更具同理心; 通过量化感知训练 (QAT) 采用 INT4 权重量化, 推理速度提升约 2 倍, 同时保持顶尖性能, 所有基准测试结果均基于 INT4 精度呈现。

图27：月之暗面全面开源 Kimi K2 Thinking



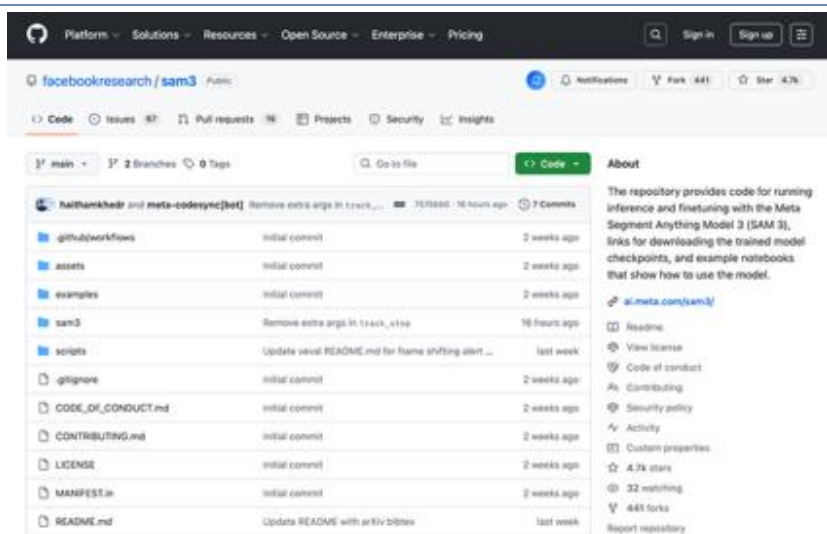
资料来源：月之暗面官网，中国银河证券研究院

6. Meta 发布 SAM 3 和 SAM 3D 模型，突破传统固定标签限制

11月21日，Meta发布第三代“分割一切”模型SAM 3。SAM 3引入可提示概念分割技术，支持文本、示例、视觉等多种提示方式，能统一完成图像与视频的检测、分割、跟踪任务，突破了传统模型依赖固定标签集的限制。在Meta全新构建的SA-Co基准测试中，其图像+视频概念分割性能较现有模型大幅提升，且总体性能始终优于Gemini 2.5 Pro等基础模型。

用于3D重建的模型SAM 3D，包含专注通用物体与场景3D重建的SAM 3D Objects和聚焦人体3D重建的SAM 3D Body，前者通过人机循环数据标注引擎为真实世界图像标注3D信息，后者基于高质量图片训练，还开源了全新参数化人体模型MHR，二者均在各类3D基准测试中达到业界领先水平，可将2D图片转化为适用于AR/VR、影视、机器人等领域的3D数据。

图28：Github上的SAM 3模型



资料来源：Github 官网，中国银河证券研究院

7. xAI 发布 Grok 4.1，幻觉率大幅下降，模型速度和质量显著提升

11月18日，马斯克旗下xAI发布前沿模型Grok 4.1，在LMArena的文本排行榜上，其推理模式Grok 4.1 Thinking以1483 Elo分数居首，非推理模式以1465 Elo分数排名第二，且在EQ-Bench3情绪智能测试中两款模式同样位列前二。Grok 4.1较前代在速度和质量上显著提升，重点强化了情感智能，能更敏锐感知用户意图、保持一致个性，回应情感类提示时更具同理心和细节，创意写作也更具文学表达与戏剧张力；同时通过后训练阶段优化，将信息检索提示中的事实性幻觉率从12.09%降至4.22%，减少将近3倍，保留了前代的智能与可靠性，树立了对话智能、情感理解和现实实用性的新标准。

图29：马斯克在X上宣布Grok 4.1模型推出



资料来源：证券时报官网，中国银河证券研究院

8. Anthropic发布Claude Opus 4.5，在多领域达业界顶尖水平

当地时间11月25日，Anthropic发布Claude Opus 4.5，该模型在编码、智能代理、计算机使用等领域达到业界顶尖水平，同时在深度研究、幻灯片及电子表格处理等日常任务中表现显著提升。现已在Anthropic应用、API及三大主流云平台上线。

Claude Opus 4.5针对即时注入攻击的稳健性方面取得了重大进展，比业内任何其他前沿模型都更难以通过及时注入而受到攻击。Opus 4.5还擅长办公任务，在制作专业文档、电子表格和演示文稿方面展现其强大能力，为金融等需要精确的行业提供一致性。此外，模型效率显著优化，API新增努力参数，可平衡速度、成本与性能，中等努力水平下SWE-bench Verified得分追平Sonnet 4.5，但输出token减少76%；在最高努力水平上，Opus 4.5比Sonnet 4.5性能高4.3%，同时token减少48%。

图30: Anthropic 推出 Claude Opus 4.5 模型

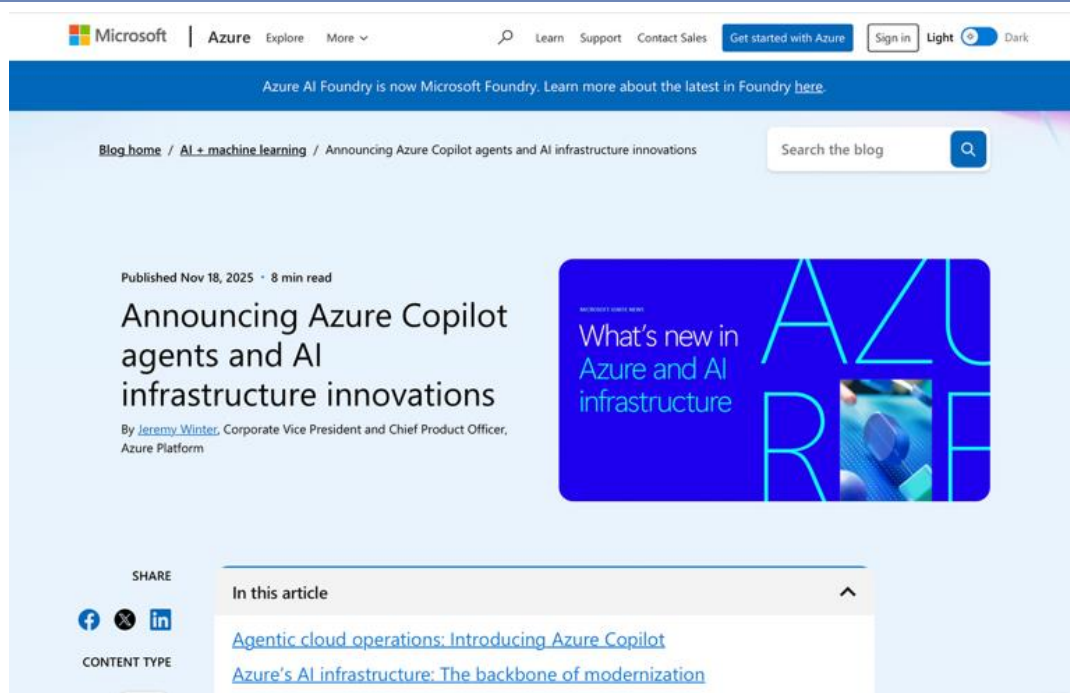


资料来源: Anthropic 官网, 中国银河证券研究院

9. 微软推出 Azure Copilot 智能代理及多项 AI 基础设施创新，助力企业数智化转型

11月18日，微软在 Microsoft Ignite 2025 大会上宣布推出 Azure Copilot 智能代理及多项 AI 基础设施创新，以强化全球云平台的可靠性、安全性与性能，助力企业加速 AI 时代的数字化转型。核心亮点包括：Azure Copilot 作为全新智能代理界面，可协调迁移、部署、优化等六大专项代理，自动化云管理全生命周期任务，遵循 RBAC 权限与 Azure 政策，支持数据自主管控，目前处于受限预览阶段；AI 基础设施方面，微软扩展了 Fairwater AI 数据中心布局，形成亚特兰大与威斯康星州的“AI 超级工厂”，大规模部署 NVIDIA GB300 GPU，通过高密度液冷、扁平化网络及专属 AI WAN 骨干网，实现单机架每秒 110 万 token 处理能力，网络吞吐量提升至 400 Gbps；核心服务升级涵盖 Azure Boost 硬件加速技术、AKS、Azure HorizonDB for PostgreSQL 数据库等。此外，通过与 SAP 深化合作、推出 App Service Managed Instance 等，简化应用与数据现代化迁移流程，AI 代理工具可使 NET 代码迁移效率大幅提升、Java 应用现代化速度提高 4 倍，全方位满足企业从基础设施到业务应用的智能化升级需求。

图31: 微软宣布 Azure Copilot 代理和人工智能基础设施创新



资料来源：微软官网，中国银河证券研究院

四、风险提示

技术迭代不及预期风险；

科技巨头竞争加剧风险；

法律监管风险；

供应链风险；

下游需求不及预期风险。

分析师承诺及简介

本人承诺以勤勉的执业态度，独立、客观地出具本报告，本报告清晰准确地反映本人的研究观点。本人薪酬的任何部分过去不曾与、现在不与、未来也将不会与本报告的具体推荐或观点直接或间接相关。

吴砚靖 TMT/科创板研究负责人，北京大学软件项目管理硕士，10 年证券分析从业经验，历任中银国际证券首席分析师，国内大型知名 PE 机构研究部执行总经理。具备一二级市场经验，长期专注科技公司研究。

免责声明

本报告由中国银河证券股份有限公司（以下简称银河证券）向其客户提供。银河证券无需因接收人收到本报告而视其为客户。若您并非银河证券客户中的专业投资者，为保证服务质量、控制投资风险、应首先联系银河证券机构销售部门或客户经理，完成投资者适当性匹配，并充分了解该项服务的性质、特点、使用的注意事项以及若不当使用可能带来的风险或损失。

本报告所载的全部内容只提供给客户做参考之用，并不构成对客户的投资咨询建议，并非作为买卖、认购证券或其它金融工具的邀请或保证。客户不应单纯依靠本报告而取代自我独立判断。银河证券认为本报告资料来源是可靠的，所载内容及观点客观公正，但不担保其准确性或完整性。本报告所载内容反映的是银河证券在最初发表本报告日期当日的判断，银河证券可发出其它与本报告所载内容不一致或有不同结论的报告，但银河证券没有义务和责任去及时更新本报告涉及的内容并通知客户。银河证券不对因客户使用本报告而导致的损失负任何责任。

本报告可能附带其它网站的地址或超级链接，对于可能涉及的银河证券网站以外的地址或超级链接，银河证券不对其内容负责。链接网站的内容不构成本报告的任何部分，客户需自行承担浏览这些网站的费用或风险。

银河证券在法律允许的情况下可参与、投资或持有本报告涉及的证券或进行证券交易，或向本报告涉及的公司提供或争取提供包括投资银行业务在内的服务或业务支持。银河证券可能与本报告涉及的公司之间存在业务关系，并无需事先或在获得业务关系后通知客户。

银河证券已具备中国证监会批复的证券投资咨询业务资格。除非另有说明，所有本报告的版权属于银河证券。未经银河证券书面授权许可，任何机构或个人不得以任何形式转发、转载、翻版或传播本报告。特提醒公众投资者慎重使用未经授权刊载或者转发的本公司证券研究报告。

本报告版权归银河证券所有并保留最终解释权。

评级标准

评级标准	评级	说明
评级标准为报告发布日后的6到12个月行业指数（或公司股价）相对市场表现，其中：A 股市场以沪深 300 指数为基准，新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准，北交所市场以北证 50 指数为基准，香港市场以恒生指数为基准。	行业评级	推荐：相对基准指数涨幅 10%以上
		中性：相对基准指数涨幅在-5%~10%之间
		回避：相对基准指数跌幅 5%以上
	公司评级	推荐：相对基准指数涨幅 20%以上
		谨慎推荐：相对基准指数涨幅在 5%~20%之间
		中性：相对基准指数涨幅在-5%~5%之间
		回避：相对基准指数跌幅 5%以上

联系

中国银河证券股份有限公司 研究院

深圳市福田区金田路 3088 号中洲大厦 20 层

上海浦东新区富城路 99 号震旦大厦 31 层

北京市丰台区西营街 8 号院 1 号楼青海金融大厦

公司网址：www.chinastock.com.cn

机构请致电：

深广地区：程 曦 0755-83471683 chengxi_yj@chinastock.com.cn

苏一耘 0755-83479312 suyiyun_yj@chinastock.com.cn

上海地区：陆韵如 021-60387901 luyunru_yj@chinastock.com.cn

李洋洋 021-20252671 liyangyang_yj@chinastock.com.cn

北京地区：田 薇 010-80927721 tianwei@chinastock.com.cn

褚 颖 010-80927755 chuying_yj@chinastock.com.cn