

# 计算机

## 推理算力需求创新高叠加制裁可能加剧，看好国产算力

### 1、豆包活跃度数据迅速上升，或推动字节推理算力需求进一步增长

根据量子位公众号及快科技数据，2024年11月份豆包APP DAU接近900万，占据国内整个赛道60%的活跃用户，11月份MAU数据达5998万，位居全球第二，仅次于ChatGPT。此外，截至今年12月18日，豆包大模型的日均tokens调用量超过4万亿，自5月发布以来，整体增长超过33倍；多模态方面，截至9月份日均生成图片5000万张。经测算，仅考虑文字和文生图的情况下，推理峰值所需算力达96.68Eflops，转化为H20所需规模为65.33万张，预计实际所需峰值算力更高，同时我们认为到2025年随着字节在AI的进一步发力，豆包等大模型的应用会进一步对推理算力的需求提升。

### 2、博通与美满电子最新业绩快速增长，或受益推理算力需求迅速升温

博通业绩优秀，对AI业务指引超预期，FY24Q4营收140.54亿美元，同比51.20%，FY24全年营收516亿美元，同比增长43.99%，主要原因为AI业务营收122亿美元，同比220%，驱动半导体业务收入营收创新高。此外，博通预计到2027年，AI XPU与网络领域的可服务市场规模（SAM）将达到600-900亿美元。美满电子FY25Q3营收15.16亿美元，同比增长7%，业绩超预期，主要原因为定制芯片需求较高，推动本财季数据中心业务营收至新高，达11.01亿美元，占总体收入比重达73.1%。我们认为博通与Marvell的业绩也预示着海外AI ASIC与推理算力需求强劲

### 3、英伟达在国内或面临制裁，美国芯片或迎来更严苛制裁，国产算力份额有望快速提升

12月9日，国家市场监督管理总局发布公告，由于英伟达涉嫌违反反垄断法，对英伟达公司开展立案调查。12月2日美国商务部产业与安全局修订美国《出口管理条例》，在过往系列芯片出口管制规则的基础上，进一步加强了针对半导体制造设备、先进计算、超级计算领域相关物项的出口管制。同时美国政府目前正积极制定一套全新的规定，意在限制先进人工智能芯片向包括东南亚和中东地区等在内的特定区域的销售，意图阻断中国大陆通过这些渠道获取高端AI芯片的途径，进一步削弱中国大陆获取此类核心技术的能力。我们认为，随着英伟达被反垄断立案调查以及美国针对AI芯片的进一步出口管控，国产算力份额有望快速提升。

**建议关注：**寒武纪、海光信息、盛科通信（通信覆盖）、神州数码、中科曙光、景嘉微

**风险提示：**供应链进一步被约束的风险、AI需求下降的风险、AI应用发展不及预期的风险、测算存在相应误差的风险

证券研究报告

2024年12月18日

投资评级

行业评级

强于大市(维持评级)

上次评级

强于大市

作者

缪欣君

分析师

SAC 执业证书编号：S1110517080003  
miaoxinjun@tfzq.com

刘鉴

联系人

liujianb@tfzq.com

行业走势图



资料来源：聚源数据

相关报告

- 《计算机-行业深度研究:AI Agent与端侧新入口共筑AI应用未来》  
2024-12-09
- 《计算机-行业点评:资金支撑叠加供需共振，泛信创自主可控有望加速》  
2024-11-11
- 《计算机-行业点评:TSMC事件的长期影响:加速国产算力行业格局集中，利好具备全国产供应链头部厂商》  
2024-11-10

图 1：预计豆包文字+文生图大模型所需推理峰值算力达 96.68Eflops，约需 65.33 万张 H20

文字模型		结果
豆包（云雀）大模型参数量（亿）		1300
12月最新日均Tokens使用量（亿）		40000
假设峰值算力需求/平均需求（倍）		4
推理所需理论算力=2*参数量*Tokens*10（Flops）		4.16*10 <sup>16</sup>
每日推理总时长（秒）		86400
假设单卡利用率		50%
实际每秒所需算力（EFlops）		<b>96.30</b>
H20单卡算力（Tflops）		148
实际需要的卡数（万张）		<b>65.07</b>
图片模型		结果
假设文生图模型参数量（亿）		150
9月最新日均图片生成量（亿张）		0.5
假设每张图片对应Tokens		668
假设峰值算力需求/平均需求（倍）		10
推理所需理论算力=2*参数量*Tokens*10（Flops）		10.02*10 <sup>13</sup>
假设单卡利用率		30%
实际每秒所需算力（EFlops）		<b>0.39</b>
H20单卡算力（Tflops）		148
实际需要的卡数（万张）		<b>0.26</b>
文字+图片		结果
预计所需算力（EFlops）		<b>96.68</b>
预计实际H20集群规模（万张）		<b>65.33</b>

资料来源：火山引擎官网、哔哩哔哩技术公众号、智选 AI 等、天风证券研究所

图 2：博通与美满电子最新业绩大超预期，或受益推理算力需求迅速升温

	收入(FY2024)					净利润(FY2024)			
	Q4(亿美元)	YOY	Q1-Q4(亿美元)	YOY	FY25Q1(指引, 亿美元)	Q4(亿美元)	YOY	Q1-Q4(亿美元)	YOY
博通(BROADCOM)	140.54	51.20%	515.74	43.99%	146	43.24	23%	58.95	-58.14%
	半导体解决方案: 82.30	59%	300.96	58%					
	基础设施软件: 58.24	41%	214.78	42%					
业绩亮点	Broadcom在2024财年实现总收入516亿美元，同比增长44%，创下历史新高。其中，VMware的成功整合推动了基础设施软件收入增长至215亿美元，占总收入的41.7%。同时，AI业务在半导体业务中（301亿美元）表现尤为亮眼，收入达到122亿美元，同比增长220%，成为推动半导体业务增长的核心动力。主要系依托领先的AI.XPU和以太网网络产品组合，Broadcom巩固了其在AI硬件领域的行业领导地位，为整体收入增长提供了强劲支撑。								
	收入(FY2025)					净利润(FY2025)			
	Q3(亿美元)	YOY	QOQ	FY25Q4(指引, 亿美元)		Q3(亿美元)	YOY	QOQ	
美满电子(MRVL)	15.16	7%	19%	18.04		-6.76	-312%	-250%	
	数据中心: 11.01	98%	25%						
	企业网络: 1.51	-44%	0%						
	运营商基础设施: 0.85	-73%	12%						
	消费者市场: 0.97	-43%	9%						
	汽车/工业: 0.83	-22%	9%						
业绩亮点	公司宣布与AWS达成5年多代合作，受强劲的AI应用需求、AI硅芯片领域巨大进展及高量产AI定制项目推动，公司Q3自定义与互联收入超预期（数据中心收入同比98%·环比25%），Q4预计持续高增长及以太网交换机&互联组合显著增长。公司量产交付5nm 1.6T PAM4 DSPs（其中800G系列产品预订需求强劲）、发布全球首款3nm 1.6.Tbps.PAM4.DSP·全球首发PCIe Gen 7技术。								

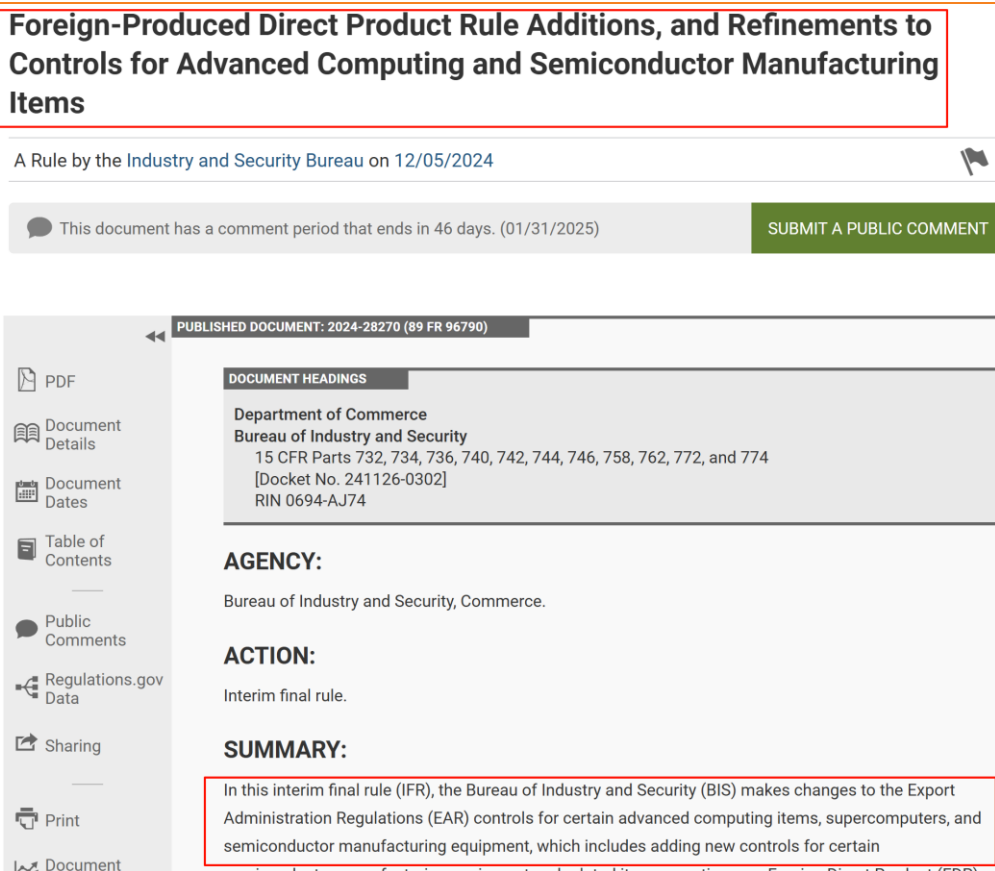
资料来源：博通官网、美满电子官网、天风证券研究所

图 3：英伟达涉嫌违反反垄断法，被立案调查



资料来源：国家市场监督管理总局、天风证券研究所

图 4：美国商务部修订美国《出口管理条例》，AI 芯片限制进一步升级



资料来源：The Daily Journal of the United States Government、天风证券研究所

## 分析师声明

本报告署名分析师在此声明：我们具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，本报告所表述的所有观点均准确地反映了我们对标的证券和发行人的个人看法。我们所得报酬的任何部分不曾与，不与，也将不会与本报告中的具体投资建议或观点有直接或间接联系。

## 一般声明

除非另有规定，本报告中的所有材料版权均属天风证券股份有限公司（已获中国证监会许可的证券投资咨询业务资格）及其附属机构（以下统称“天风证券”）。未经天风证券事先书面授权，不得以任何方式修改、发送或者复制本报告及其所包含的材料、内容。所有本报告中使用的商标、服务标识及标记均为天风证券的商标、服务标识及标记。

本报告是机密的，仅供我们的客户使用，天风证券不因收件人收到本报告而视其为天风证券的客户。本报告中的信息均来源于我们认为可靠的已公开资料，但天风证券对这些信息的准确性及完整性不作任何保证。本报告中的信息、意见等均仅供客户参考，不构成所述证券买卖的出价或征价邀请或要约。该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。客户应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专家的意见。对依据或者使用本报告所造成的一切后果，天风证券及/或其关联人员均不承担任何法律责任。

本报告所载的意见、评估及预测仅为本报告出具日的观点和判断。该等意见、评估及预测无需通知即可随时更改。过往的表现亦不应作为日后表现的预示和担保。在不同时期，天风证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。天风证券的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。天风证券没有将此意见及建议向报告所有接收者进行更新的义务。天风证券的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

## 特别声明

在法律许可的情况下，天风证券可能会持有本报告中提及公司所发行的证券并进行交易，也可能为这些公司提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。因此，投资者应当考虑到天风证券及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突，投资者请勿将本报告视为投资或其他决定的唯一参考依据。

## 投资评级声明

类别	说明	评级	体系
股票投资评级	自报告日后的 6 个月内，相对同期沪深 300 指数的涨跌幅	买入	预期股价相对收益 20%以上
		增持	预期股价相对收益 10%-20%
		持有	预期股价相对收益 -10%-10%
		卖出	预期股价相对收益 -10%以下
行业投资评级	自报告日后的 6 个月内，相对同期沪深 300 指数的涨跌幅	强于大市	预期行业指数涨幅 5%以上
		中性	预期行业指数涨幅 -5%-5%
		弱于大市	预期行业指数涨幅 -5%以下

## 天风证券研究

北京	海口	上海	深圳
北京市西城区德胜国际中心 B 座 11 层	海南省海口市美兰区国兴大道 3 号互联网金融大厦 A 栋 23 层 2301 房	上海市虹口区北外滩国际客运中心 6 号楼 4 层	深圳市福田区益田路 5033 号平安金融中心 71 楼
邮编：100088	邮编：570102	邮编：200086	邮编：518000
邮箱：research@tfzq.com	电话：(0898)-65365390 邮箱：research@tfzq.com	电话：(8621)-65055515 传真：(8621)-61069806 邮箱：research@tfzq.com	电话：(86755)-23915663 传真：(86755)-82571995 邮箱：research@tfzq.com