

证券研究报告

2025年02月23日

行业报告：行业专题研究

计算机

AI私有云，是一片新蓝海

作者：

分析师 缪欣君 SAC执业证书编号：S1110517080003



天风证券
TF SECURITIES

行业评级：强于大市（维持评级）
上次评级：强于大市

请务必阅读正文之后的信息披露和免责声明

摘要

主流三种私有化部署AI方案：本地部署如裸金属、超融合以及线上部署专属云。对于数据敏感行业（如政府、国企、金融、医疗行业），考虑数据安全高要求，本地部署能规避数据安全风险。专属云泛用性更高，是一种介于公有云和私有云之间的云服务模式。**根据测算，私有云AI硬件+软件总市场空间近2000亿元。**

不只是硬件芯片，软件价值重要，vGPU技术成为集群推理效率关键，同时需要统一纳管平台。vGPU（GPU虚拟化）使得多个应用可以共享同一物理GPU，从而提高资源利用率。以京东云一体机为例，使用vGPU资源利用率提升70%，**推理效率提升30%**。同时随着集群规模扩大，一大问题在于如何高效地分配和管理GPU资源，同时满足不同用户和应用的需求，需要统一纳管平台。

私有云AI，千亿新蓝海，主要利好三类厂商

- **NV算力服务器厂商资源积累较为充足：**服务器厂商拥有较为充足的卡资源，首先受益于此轮浪潮。以智微智能旗下的腾云智算为例，算力资源包含英伟达H200、4090，为私有化部署提供稳定的算力支持。
- **深厚客户基础的IT基础架构云公司/运营商：**直销或者通过渠道拥有丰富客户资源，相关产品迅速推向市场，如深信服、海康威视、新华三等。专属云领域，三大运营商占据专属云近50%市场份额（本身也有极强客户资源）
- **短中长期国产算力显著受益，被市场低估：**国产算力长期受益于此次DS的开源生态，打破了闭源模型如Openai和闭源芯片如NV建立的生态壁垒，短期看若后续NV卡供不应求，国产算力有望量价齐升。

建议关注：

1、算力：

- 1) **国产算力AI芯片四小龙：**寒武纪、海光信息、中科曙光、云天励飞
- 2) **NV&昇腾相关厂商：**智微智能、华勤技术、科大讯飞、浪潮信息、神州数码、紫光股份（与通信组联合覆盖）、烽火通信等

2、云：

- 1) **本地私有云：**深信服、海康威视、优刻得、紫光股份（新华三，与通信组联合覆盖）、云赛智联、品高股份、深桑达等
- 2) **线上私有云（专属云）：**
 - 三大运营商：**中国联通（通信组覆盖）、中国电信（通信组覆盖）、中国移动（通信组覆盖）
 - 字节/华为/阿里云合作伙伴：**汉得信息、金蝶国际、软通动力、中科创达（与电子组联合覆盖）、神州数码、亚信安全、中软国际等

4、专属云上游：

申菱环境（家电组覆盖）、中恒电气、科泰电源、科华数据（通信组和电新组联合覆盖）、数据港（与通信组联合覆盖）、润泽科技（通信组与机械组联合覆盖）、润建股份（通信组覆盖）、锐捷网络、英维克（通信组与机械组联合覆盖）

风险提示：AI算力景气度下降的风险、云计算竞争加剧的风险、政策落地不及预期风险、测算主观性风险、大模型私有化部署渗透率不及预期的风险

一、DS大幅提升本地模型能力， AI私有云浪潮有望开启

1.1、DeepSeek “鲑鱼效应” 下企业端积极接入，央企开启 “AI+” 专项行动

- **2月8日，工信部发文力推最新AI技术普惠应用**，三家基础电信企业均全面接入DeepSeek开源大模型，实现在多场景、多产品中应用，针对热门的DeepSeek-R1模型提供专属算力方案和配套环境，助力国产大模型性能释放。
- **2月19日，国务院国资委召开中央企业“AI+”专项行动深化部署会**，强调国资央企要抓住人工智能产业发展的战略窗口期，强化科技创新，聚焦关键领域加快掌握“根技术”，坚定攻关大模型，积极参与开放生态建设，推动产生更多“从0到1”的原始创新，加速推进成果转化和产业化发展。
- 在DeepSeek引发的“鲑鱼效应”下，AI热潮持续升温。从云计算平台到芯片厂商，从软件企业到互联网巨头，再到运营商等各个领域，纷纷宣布接入DeepSeek。2月13日，腾讯元宝迎来更新，同时支持满血版DeepSeek大模型；**2月15日，微信搜一搜在调用混元大模型丰富AI搜索的同时，正式灰度测试接入DeepSeek**。用户可通过微信对话框顶部搜索入口进入“AI搜索”功能，免费使用DeepSeek-R1满血版模型。

图：国务院国资委部署深化中央企业“AI+”专项行动



国务院国有资产监督管理委员会
State-owned Assets Supervision and Administration Commission of the State Council

2025年2月22日 星期六

[首页](#) [机构概况](#) [新闻发布](#) [国资监管](#) [政务公开](#) [国资数据](#) [互动交流](#) [在线服务](#) [热搜](#)

[首页](#) > [新闻发布](#) > [国资要闻](#) > 正文

国务院国资委部署深化中央企业“AI+”专项行动

文章来源：新闻中心 发布时间：2025-02-21

2月19日，国务院国资委召开中央企业“AI+”专项行动深化部署会，总结国资央企发展人工智能进展成效，研究部署下一步重点工作。国务院国资委党委书记、主任张玉卓出席会议并讲话强调，要深入学习贯彻习近平总书记关于发展人工智能的重要讲话精神，进一步增强责任感、使命感和紧迫感，立足服务国家战略，紧盯前沿发展态势，发挥需求规模大、产业配套全、应用场景多的优势，加快推动人工智能产业高质量发展，实现更多标志性成果和突破性进展，全力塑造产业新优势、培育发展新动能。国务院国资委党委委员、副主任谭作钧主持会议。

会议指出，近年来，国务院国资委坚决贯彻落实党中央、国务院决策部署，全力推动中央企业融入国家算力布局，与龙头企业、科研机构深化合作，一批高价值行业应用场景落地，智能算力供给能力显著提升，数据基建建设稳步推进，大模型构建加速追赶。在人工智能领域取得一系列重要成果。

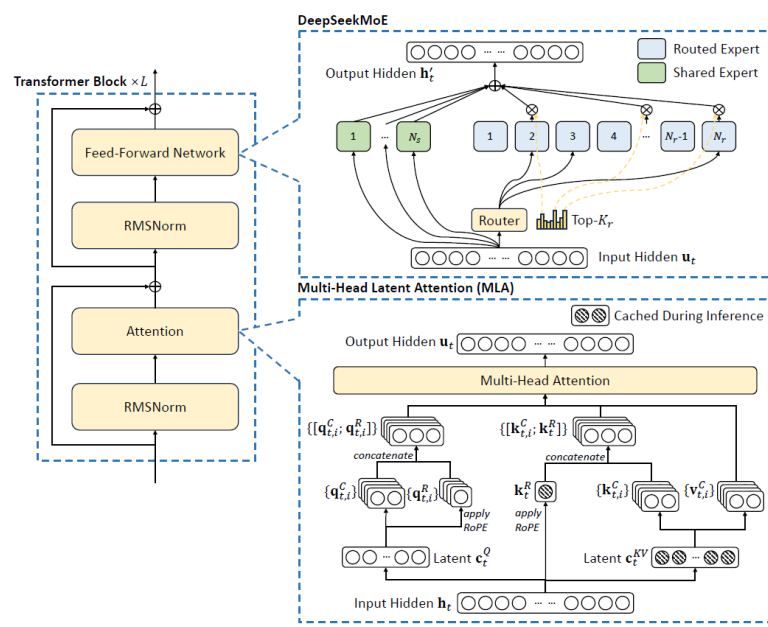
图：已接入DeepSeek厂商统计（部分）

类别	主要厂商	概述
运营商	中国电信	中国电信天翼云全场景上架DeepSeek，为用户提供更便捷、更高效的智能服务。
	中国联通	中国联通联通云已基于“星罗”平台实现国产及主流算力适配多规格DeepSeek-R1模型，兼顾私有化和公有化场景，实现多产品场景调用。
	中国移动	中国移动移动云宣布全面上线DeepSeek，实现全版本覆盖、全尺寸适配、全功能畅用。用户可选择任一智算资源池进行部署、蒸馏、智能体编排等操作。
云计算	华为云	昇腾社区正式上线包括DeepSeek-R1、DeepSeek-V3、DeepSeek-V2、Janus-Pro等系列新模型，支持昇腾硬件平台开箱即用，推理快速部署，带来更快、更高效、更便捷的AI开发和应用体验。
	腾讯云	宣布DeepSeek-R1大模型一键部署至腾讯云HAI上，开发者仅需3分钟即可接入调用。
	阿里云	宣布其PAIModelGallery支持云上一键部署DeepSeek-V3和DeepSeek-R1模型，用户可以在该平台上零代码实现从训练到部署再到推理的全过程。
	火山引擎	宣布全面支持DeepSeek系列大模型，包括V3和R1等不同尺寸模型。
	百度智能云	千帆平台已正式上线DeepSeek-R1和DeepSeek-V3模型，并推出超低价方案以及限时免费服务。
	云轴科技ZStack	宣布其AllInra平台ZStack智塔全面支持企业私有化部署DeepSeekV3/R1/JanusPro三种模型。
软件与互联网	联想集团	与沐曦股份联合发布基于DeepSeek大模型的一体机解决方案，联想个人智能体“小天”已接入DeepSeek。
	安恒信息	宣布旗下恒脑安全垂域大模型集成DeepSeek，发布了首个“DeepSeek版”安全智能体。
	奇安信	宣布其安全智能体已接入DeepSeek平台，通过DeepSeek的优化和蒸馏技术，大幅降低了运营成本，提升了模型在处理威胁检测和响应速度上的性能。
	360	宣布在DeepSeek基于MIT开源协议的生态环境下，公司旗下的部分产品进行了DeepSeek的接入与本地化部署。
	天融信	宣布天问大模型正式完成DeepSeek大模型的接入，赋能安全事件检测分析、威胁情报融合等多个场景。
	万兴科技	宣布已完成推理大模型DeepSeek-R1的适配，并落地融合相关能力于万兴喵影、亿图图示、亿图脑图MindMaster、万兴PDF等产品。
	当虹科技	宣布其BlackEye多模态视听大模型完成DeepSeek深度融合，并完成视听传媒、工业与卫星、车载智能座舱等多行业垂类场景的数据调优训练。
		

1.2、DS通过架构创新等技术提升推理效率，大幅提升低成本本地模型能力

- DeepSeek-V3采用混合专家架构（MoE），包含256个专家模块，每个词元仅激活37B参数，动态选择最相关的专家处理任务，既降低计算资源消耗，又保持了千亿级参数模型的性能。结合多头潜在注意力（MLA）技术，通过低秩压缩机制减少键值缓存需求，**显存占用降低30%-50%，推理效率提升3倍**。这种架构设计在数学、代码生成等任务中表现尤为突出。
- **DeepSeek-R1**在后训练阶段大规模使用了RL（强化学习），在仅有极少标注数据的情况下，极大提升了模型推理能力，降低推理开销。模型表现方面，它在数学、代码、自然语言推理等任务上，**性能平齐OpenAI o1正式版**。
- 通过蒸馏DeepSeek-R1生成的800K样本，小规模模型（1.5B~70B）在推理任务中表现显著。同时，蒸馏小模型对硬件要求大幅降低，有望推动私有化部署加速落地。从硬件部署来看，70B的小模型不考虑并发情况下，本地仅需两张RTX4090。

图：DeepSeekMoe创新架构以及MLA机制大幅降低推理成本



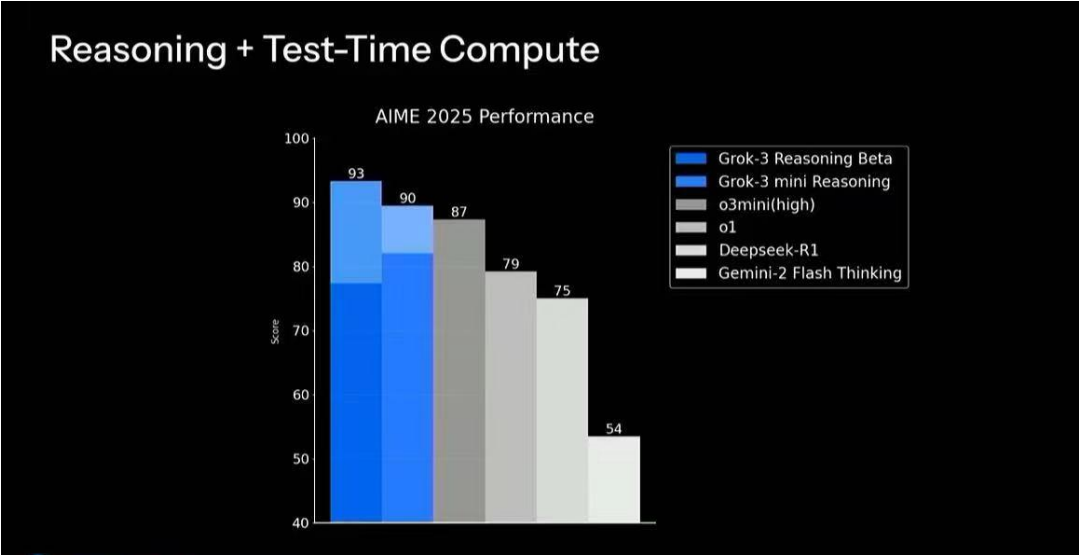
图：Deepseek R1不同参数配置需求

模型	参数量	VRAM 要求	推荐GPU配置
DeepSeek-R1-Zero	671B	~1342 GB	A100 80G*16
DeepSeek-R1	671B	~1342 GB	A100 80G*16
DeepSeek-R1-Distill-Qwen-1.5B	1.5B	~0.75 GB	RTX 3060 12GB 或以上
DeepSeek-R1-Distill-Qwen-7B	7B	~3.5 GB	RTX 3060 12GB 或以上
DeepSeek-R1-Distill-Llama-8B	8B	~4 GB	RTX 3060 12GB 或以上
DeepSeek-R1-Distill-Qwen-14B	14B	~7 GB	RTX 3060 12GB 或以上
DeepSeek-R1-Distill-Qwen-32B	32B	~16 GB	RTX 4090 24GB
DeepSeek-R1-Distill-Llama-70B	70B	~35 GB	RTX 4090 x 2

1.3、私有化部署再添利好：GROK 3免费开放，成熟后立即开源前代模型

- 2月18日，埃隆·马斯克旗下的人工智能初创公司xAI发布最新版Grok 3聊天机器人。根据xAI官方公布的测试结果，Grok 3测试版的推理能力超过了现有的AI模型。在有关推理和测试时间的基准测试中，Grok 3取得了比DeepSeek-R1、OpenAI o1、OpenAI o3 mini- high、Gemini-2 Flash Thinking更优的效果。
- 根据xAI的比较基准显示，Grok 3在科学、编码和数学方面的得分高于Gemini 2 Pro、Deepseek V3、ChatGPT 4o等。并且，在盲测方面，xAI的Grok-3（早期版本）在 ImArena 排行榜中以1402 分的历史最高成绩登顶，成为首个突破1400分的AI模型。这一成绩超越了Google、OpenAI、DeepSeek等业界主要竞品。
- 目前，Grok 3已开放免费使用（直至服务器崩溃），且在Grok 3成熟后将会把上一代Grok 2开源，利好私有化部署。

图：测试结果看，Grok 3 推理能力超越现有的AI模型



图：xGrok-3（早期版本）在 ImArena 排行榜中以1402 分的历史最高成绩登顶

lmarena.ai

xAI Grok-3 #1 in Arena
First model with >1400 score

Rank* (UB)	Model	Arena Score	95% CI	Votes	Organization
1	chocolate_(Early_Grok-3)	1402	+7/-6	7829	xAI
2	Gemini-2.0-Flash-Thinking-Exp-01-21	1385	+5/-5	13336	Google
2	Gemini-2.0-Pro-Exp-02-05	1379	+5/-6	11197	Google
2	ChatGPT-4o-latest_(2025-01-29)	1377	+5/-6	10529	OpenAI
5	DeepSeek-R1	1361	+8/-7	5079	DeepSeek
5	Gemini-2.0-Flash-001	1356	+6/-5	9092	Google
5	o1-2024-12-17	1353	+6/-5	15437	OpenAI
8	o1-preview	1335	+4/-4	33169	OpenAI
8	Qwen2.5-Max	1332	+7/-7	7370	Alibaba
10	DeepSeek-V3	1317	+4/-4	17717	DeepSeek
10	Qwen-Plus-0125	1313	+8/-10	3682	Alibaba
10	Gemini-2.0-Flash-Lite-Preview-02-05	1310	+6/-6	8465	Google
10	GLM-4-Plus-0111	1308	+8/-8	4171	Zhipu
11	o3-mini	1305	+6/-7	9338	OpenAI

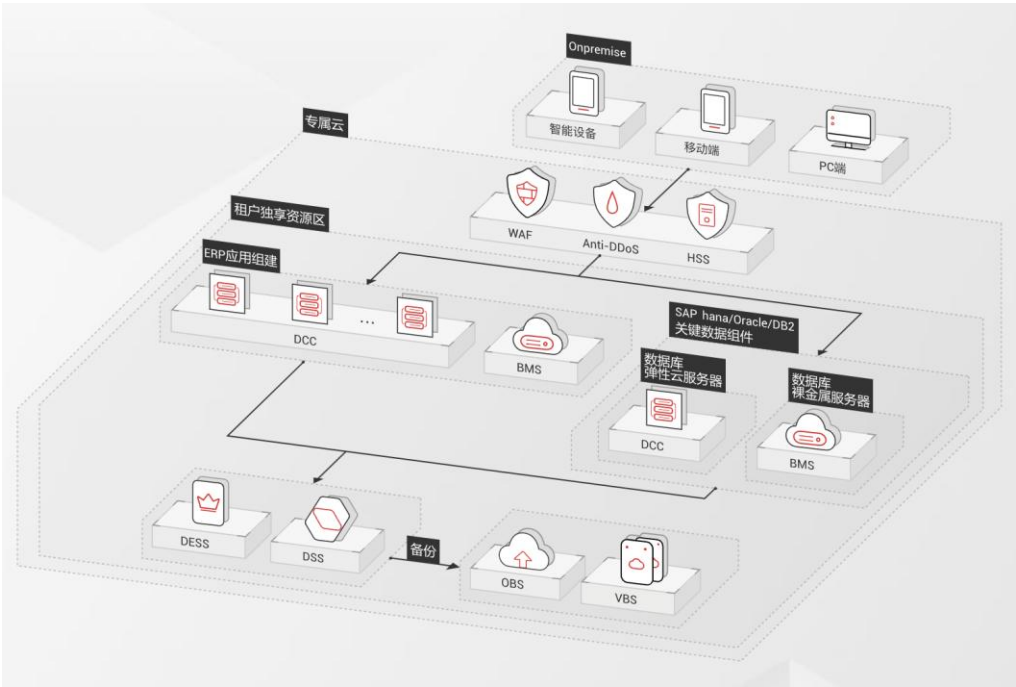
二、三类私有云AI部署方式，千亿市场空间有望打开

2.1、三种私有化部署方式：裸金属、超融合、专属云

目前私有化部署主要分为三种形式：

- **裸金属**：即物理服务器，软件需自行部署；
- **超融合**：一体机形态，硬件+软件平台集成，开箱即用；
- **专属云**：介于私有云和公有云之间的形态，在云端提供资源隔离，保证数据安全。

图：专属云介于公私有云之间的形态，在云端提供资源隔离



图：裸金属即传统物理服务器，软件自行部署



图：超融合一体机包含软件+硬件，本地部署开箱即用

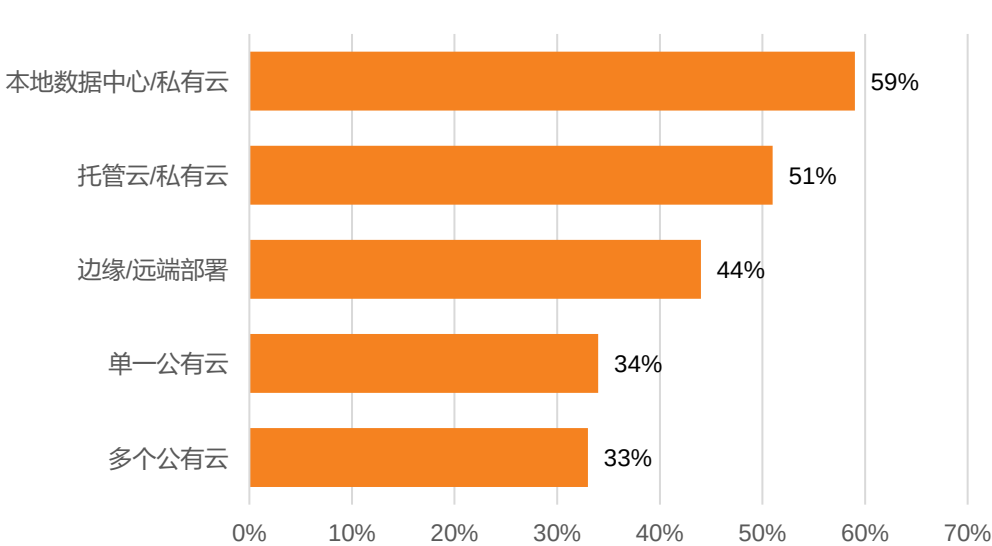


2.2、占比看：本地数据中心占59%，托管云（专属云）占51%

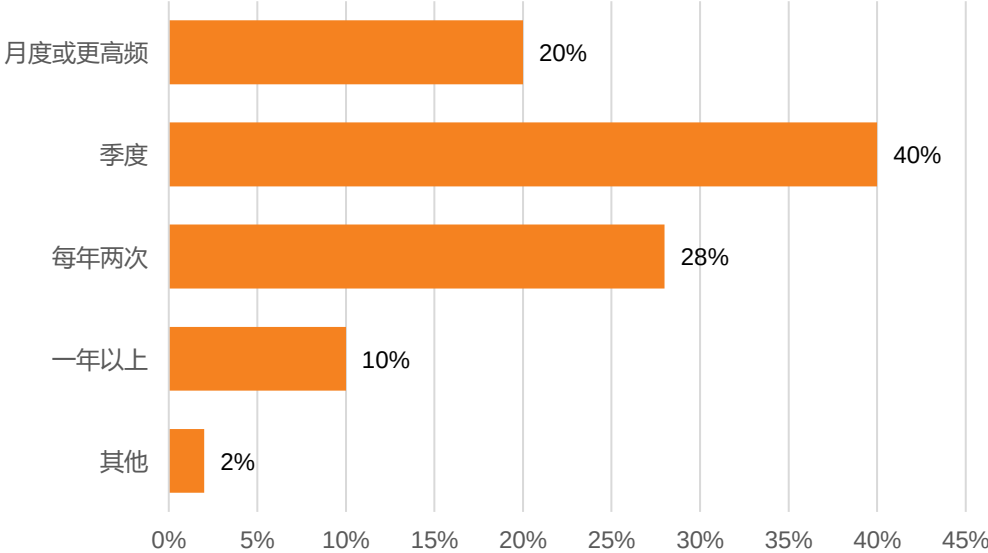
根据Nutanix发布的企业AI现状报告：

- **部署环境：**受数据本地化、所分析数据的监管要求，近60%企业选择在本地数据中心或私有云和边缘位置部署AI推理模型，51%的企业选择在托管云（专属云）部署。
- **AI模型微调频率：**60%的企业计划不低于每季度更新一次其人工智能模型。

图：59%企业选择本地/私有云部署大模型，51%选择托管云



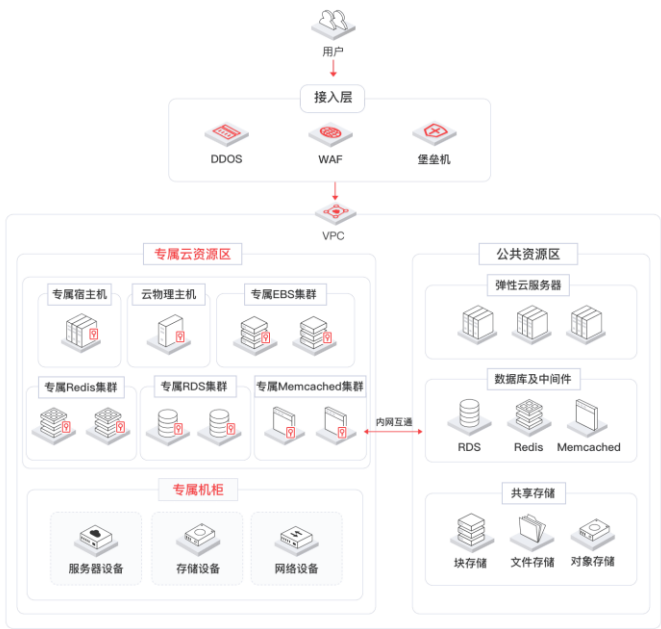
图：60%企业计划保持每月/每季度大模型微调频率



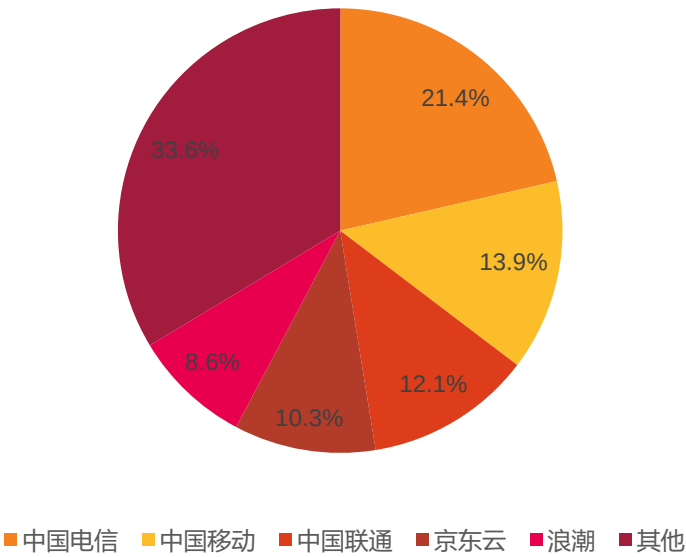
2.3、专属云：多数行业的最优选择，市场由运营商、公有云厂商主导

- **对于70B及以上的超级模型，部署本质上是一场系统工程。**不仅需要构建多卡并行计算架构，还要配套液冷散热系统和冗余电源保障。我们认为，未来大多数企业（尤其是中小微企业）会优先选择更加灵活的专属云方案，主要原因在于中小微企业机房条件有限，部分数据中心建设在办公场所内，自建自用模式，一般没有专门的数据机房，虽可以通过API形式调用，但考虑api接入方式存在不稳定性（如官方暂停充值或计算资源紧张）以及数据安全问题，最终会选择自行部署。
- 专属云是一种介于公有云和私有云之间的云服务模式，具有资源独享、数据隔离、性能稳定等优点。对于多数行业来说，专属云可以提供更加灵活、高效和经济的云服务，满足其对数据安全、性能和成本的综合需求。
- **专属云头部效应明显，市场由三大运营商主导。**据IDC数据，2024 上半年，中国专属云服务市场前五大服务商市场份额占比合计达66.4%。其中，中国电信、中国移动、中国联通分别占据21.4%、13.9%、12.1%份额。政府行业依然是专属云服务市场营收的主阵地，贡献了整体市场65.3%的营收。泛政务领域IT基础设施的周期性换代、新架构升级依然是专属云服务市场的主要机会点。

图：专属云资源独享、数据隔离，介于私有云和公有云之间



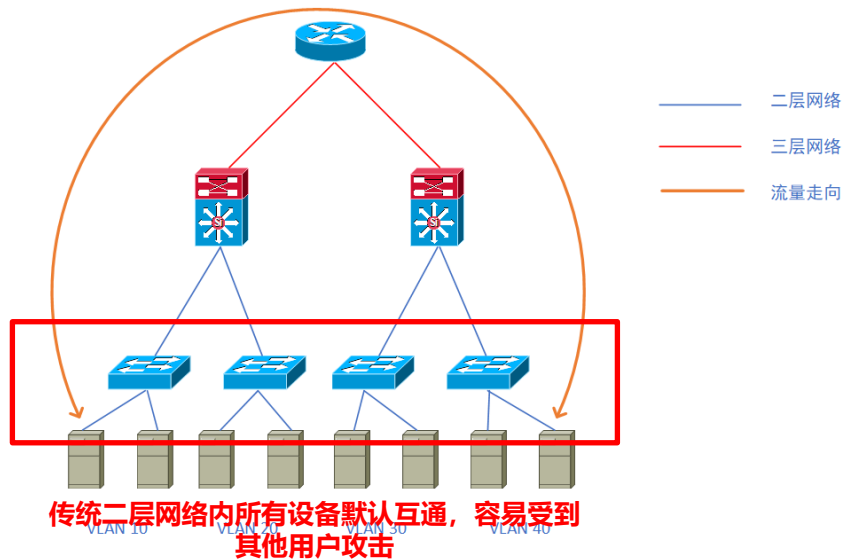
图：专属云市场由三大运营商主导（24H1国内专属云市场份额占比）



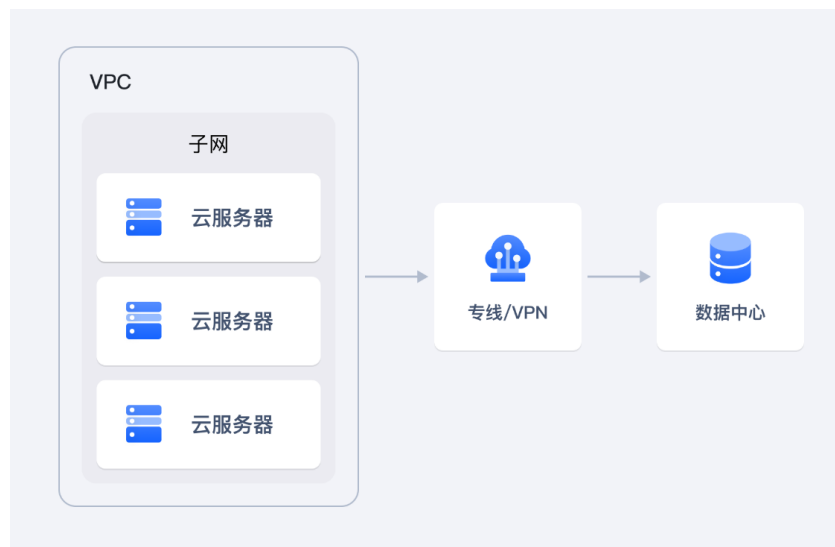
2.3、专属云：VPC提供云端网安，保证数据、工作负载与其他租户相互隔离

- 云计算对虚拟化网络的要求越来越高，要求网络具有弹性、安全性、可靠性和私密性。传统的二层网络模型存在很多问题，其中最大的问题就是安全问题。**二层网络内的所有设备默认是可以互通的，这就导致云主机很容易被同网络的其他用户恶意攻击。**
- 私有网络VPC (Virtual Private Cloud) 是基于公有云基础上构建的虚拟网络环境，能够在公共网络资源上隔离出专属独立网络，保障用户安全和隐私。此外可以灵活定义VPC内主机的IP地址、路由策略等。VPC托管于多租户架构，**逻辑上每个客户的数据和工作负载与其他租户相互隔离，由云供应商负责保障这种隔离。**

图：传统二层网络内所有设备默认互通，容易受到其他用户攻击



图：混合云场景下VPC与数据中心通过专线/VPN进行数据交互



2.4、超融合：部分行业倾向于本地部署，超融合一体机有望成为主要形态

- **本地部署为数据敏感行业首选：**对于金融、医疗等数据相对敏感的行业，考虑数据安全高要求，本地部署能规避第三方服务器泄密风险，避免了网络延迟和隐私泄露的风险，有望成为此类行业部署首选。
- **国产AI芯片基本完成适配：**目前华为昇腾、寒武纪、海光信息、沐曦、天数智芯、摩尔线程等多家国产AI芯片企业宣布完成对DeepSeek模型的适配或服务上架，覆盖从训练到推理的全链条场景。
- **超融合一体机集成了计算、存储、网络和管理功能，部署简单、管理方便且性能高效。**对于数据敏感行业来说，超融合一体机可以提供更加安全、可靠和高效的本地部署方案，有望成为主要形态。

图：昇腾一体机推荐配置，技术成本大幅降低，开箱即用

模型名称	参数	产品	配置	系统吞吐 (Token/s)	多用户并发数 (路)
DeepSeek V3	671B	Atlas 800I A2	1024GB	1911	192
DeepSeek R1	671B	Atlas 800I A2	1024GB	1911	192
DeepSeek-R1 Distill-Llama-70B	70B	Atlas 800I A2	512GB	3300	165
DeepSeek-R1 Distill-Qwen-32B	32B	Atlas 800I A2	256GB	4940	247
DeepSeek-R1 Distill-Qwen-14B	14B	Atlas 800I A2	256GB	7500	300
DeepSeek-R1 Distill-Qwen-14B	14B	Atlas 300I Duo	1*Duo 96GB PCIE	730	80
DeepSeek-R1 Distill-Llama-8B	8B	Atlas 300I Duo	1*Duo 96GB PCIE	956	115
DeepSeek-R1 Distill-Qwen-7B	7B	Atlas 300I Duo	1*Duo 96GB PCIE	956	115
DeepSeek-R1 Distill-Qwen-1.5B	1.5B	Atlas 300V	1*300V 24GB PCIE	432	16

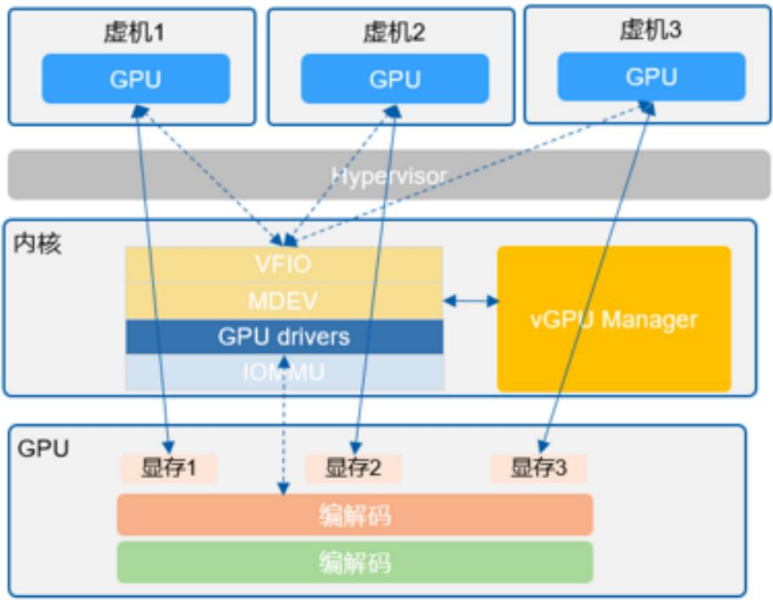
图：多个昇腾合作伙伴已上线DeepSeek超融合一体机



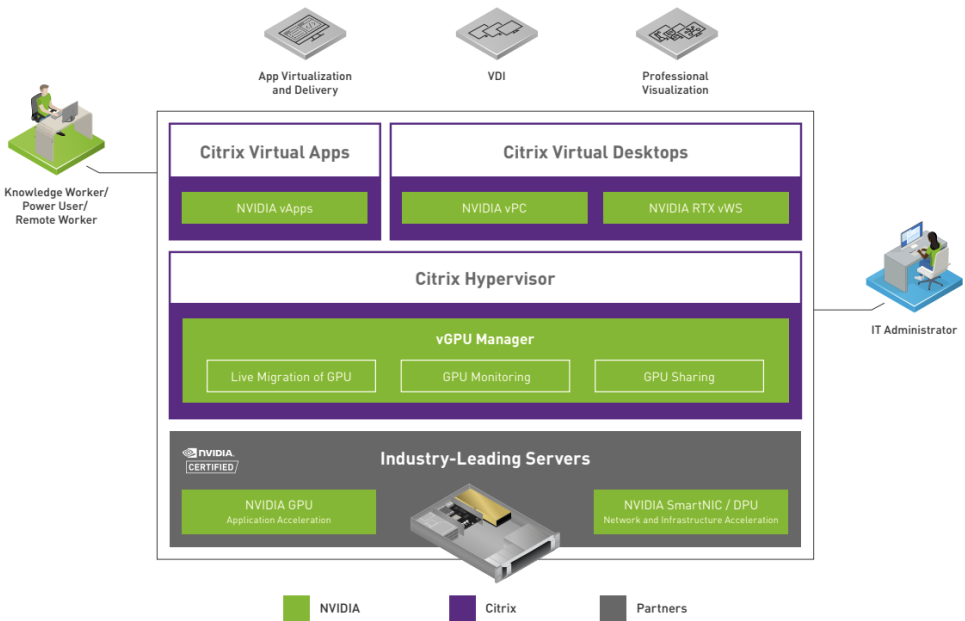
2.5、软件：vGPU技术成为集群性能高效调配关键，交付形态为统一管理平台

- 云计算环境中，随着集群规模扩大，最大的问题在于如何高效地分配和管理GPU资源，确保资源利用率最大化，同时满足不同用户和应用的需求。
- vGPU（GPU虚拟化）通过将物理GPU资源虚拟化为多个逻辑GPU资源，使得多个应用可以共享同一物理GPU，从而提高资源利用率，降低硬件成本，并提供灵活的资源调度能力。以京东云一体机为例，使用vGPU资源利用率提升70%，推理效率提升30%。

图：英伟达vGPU方案，高效管理虚拟机中的GPU资源



图：GPU资源通过虚拟化统一调度，提升利用率



2.5、软件：vGPU技术成为集群性能高效调配关键，交付形态为统一管理平台

- 我们认为，私有云AI随着集群规模扩大，最大的问题在于如何高效地分配和管理GPU资源，同时满足不同用户和应用的需求，需要统一纳管平台。
- 最终交付形式为整体的AI算力平台统一调配：以深信服为例，最终形式为其面向大模型开发场景的AICP平台，提供异构GPU算力管理调度、模型和数据管理等能力，降低技术门槛。
- 深信服通过AICP算力平台的部署方案在多实例、并发推理场景中可实现5-10倍的性能提升。大模型并发度更高、吞吐量更大、推理响应延时更低，让用户在企业级AI应用构建及实际使用时，资源投入更低、体验更优。

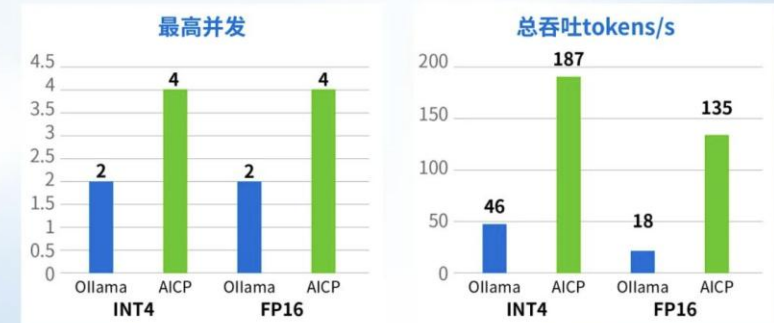
图：32B模型2k上下文，AICP并发是Ollama的8~10倍，总吞吐10倍以上（4*4090）

图：深信服AICP算力平台，提供算力调度、模型和数据管理等功能



注：INT4用2*4090，FP16用4*4090

图：知识库应用场景（4k上下文），AICP并发是Ollama的2倍，总吞吐4~8倍



注：INT4用2*4090，FP16用4*4090

2.6、空间：DS引领模型平权下，私有化部署加速，市场规模有望近2000亿元

前提假设：

- 1) 假设医疗行业（二级以上的医院）、上市券商、央国企、政府、高校会采购AI硬件设备；
 - 2) 假设以上行业对标政务云上云率，渗透率达70%；
 - 3) 假设中型企业以上有能力购买超融合一体机设备且渗透率达40%；
 - 4) 假设软件成本为总体硬件的30%。
- 经测算，大模型私有化部署硬件+软件总体市场空间达1977.16亿元。

图：私有化部署市场空间达1977.16亿元

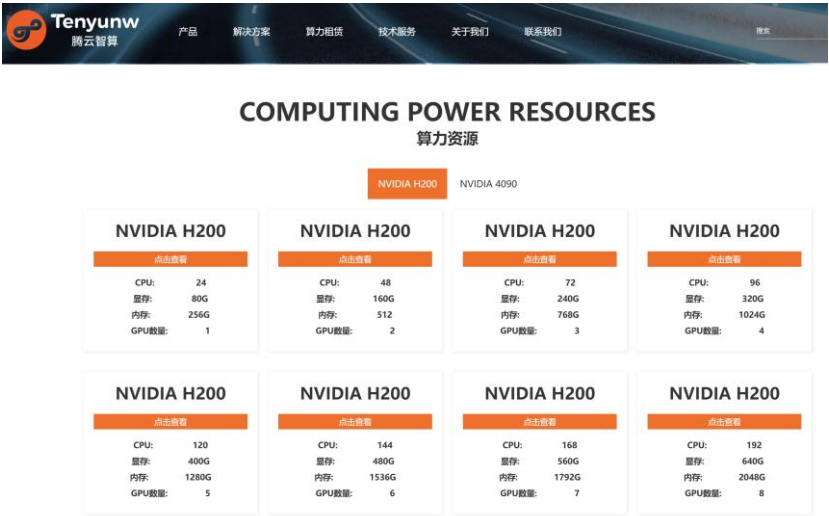
分类		计算
硬件		
医疗行业	三级医院数量	3855家
	二级医院数量（假设此项渗透率20%）	11946家
金融行业	上市券商数量	43家
央国企	央企数量	243家
	省属国企数量	665家
政府	假设各单位以省级为单位部署671B模型	
	省数量	23个
	自治区	5个
	直辖市	4个
	特别行政区	2个
	分支机构（假设每个省级地方一个分支架构）	84家
高校		3117所
大型企业		21872家
假设渗透率		70%
671B满血私有化部署价格	（考虑并发数假设4台8*H20服务器）	520万元
医疗+金融+央国企+高校+大型企业对应市场规模		1312.74亿元
中型企业		
	企业数量	13.01万家
	上云率	40%
假设渗透率		40%
70B蒸馏版部署价格		100万元
中型企业对应市场规模		208.16亿元
未来硬件总市场空间		1520.89亿元
软件		
软件及运维成本（包含适配、部署以及调度平台）	按硬件的30%计算	456.27亿元
总体市场空间		
总体市场空间	硬件+软件	1977.16亿元

三、利好两类公司，具备深厚客户基础的IT基础架构/云公司，或者国产算力/NV算力服务商

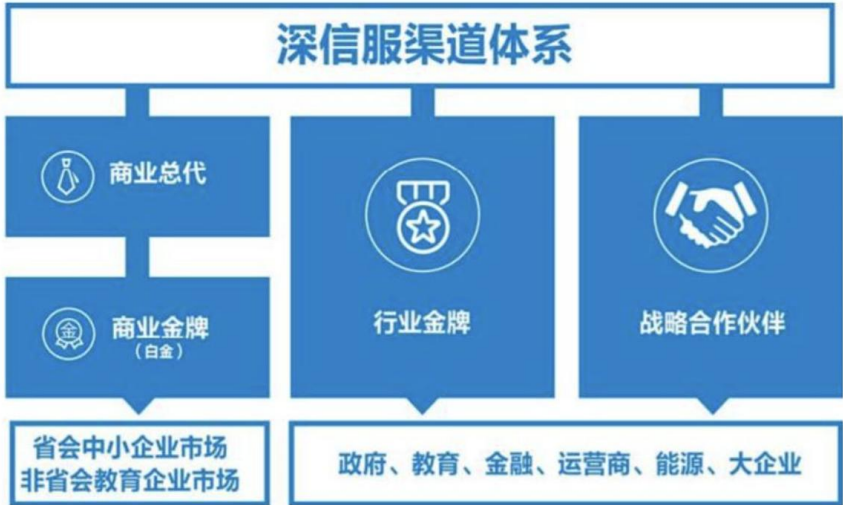
3.1、建议关注利好两类公司，具备深厚客户基础的IT基础架构/云公司，或者国产算力/NV算力服务商

- **NV算力服务器厂商资源积累较为充足：**服务器厂商拥有较为充足的GPU资源，以智微智能旗下的腾云智算为例，算力资源包含英伟达H200、4090，为私有化部署提供稳定的算力支持。
- **深厚客户基础的IT基础架构/云公司：**渠道代理销售先将产品销售给渠道代理商，再由渠道代理商将产品销售给终端用户。拥有大量客户资源，产品迅速推向市场，如华为、深信服、海康威视、新华三等。**专属云领域，三大运营商占据专属云市场份额：**三大运营商在政府+国央企客户方面拥有较大优势，24H1占据专属云近50%市场份额。
- **若后续NV卡供不应求，国产算力有望量价齐升：**国产算力长期受益于此次DS的开源生态，**打破了闭源模型如Openai和闭源芯片如NV建立的生态壁垒**，短期看若后续NV卡供不应求，国产算力有望量价齐升。利好国产推理卡厂商，如昇腾、寒武纪、海光信息、燧原科技、摩尔线程、天数智芯等。

图：智微智能旗下腾云智算算力储备包括H200、4090等GPU



图：深信服渠道体系不同类型代理面向不同行业铺开



3.2、建议关注与风险提示

私有云AI，千亿新蓝海，主要利好三类厂商

- NV算力服务器厂商资源积累较为充足：服务器厂商拥有较为充足的卡资源，首先受益于此轮浪潮。以智微智能旗下的腾云智算为例，算力资源包含英伟达H200、4090，为私有化部署提供稳定的算力支持。
- 深厚客户基础的IT基础架构公司/运营商：直销或者通过渠道拥有丰富客户资源，相关产品迅速推向市场，如深信服、海康威视、新华三等。专属云领域，三大运营商24H1占据专属云近50%市场份额（本身也有极强客户资源）
- 短中长期国产算力显著受益，被市场低估：国产算力长期受益于此次DS的开源生态，打破了闭源模型如Openai和闭源芯片如NV建立的生态壁垒，短期看若后续NV卡供不应求，国产算力有望量价齐升。

建议关注：

1、算力：

- 1) 国产算力AI芯片四小龙：寒武纪、海光信息、中科曙光、云天励飞
- 2) NV&昇腾相关厂商：智微智能、华勤技术、科大讯飞、浪潮信息、神州数码、紫光股份（与通信组联合覆盖）、烽火通信等

2、云：

- 1) 本地私有云：深信服、海康威视、优刻得、紫光股份（新华三，与通信组联合覆盖）、云赛智联、品高股份、深桑达等
- 2) 线上私有云（专属云）：
 - 三大运营商：中国联通（通信组覆盖）、中国电信（通信组覆盖）、中国移动（通信组覆盖）
 - 字节/华为/阿里云合作伙伴：汉得信息、金蝶国际、软通动力、中科创达（与电子组联合覆盖）、神州数码、亚信安全、中软国际等

4、专属云上游：

申菱环境（家电组覆盖）、中恒电气、科泰电源、科华数据（通信组和电新组联合覆盖）、数据港（与通信组联合覆盖）、润泽科技（通信组与机械组联合覆盖）、润建股份（通信组覆盖）、锐捷网络、英维克（通信组与机械组联合覆盖）

风险提示

- 1) **AI算力景气度下降的风险**：算力支出与下游应用息息相关，若AI应用需要更长期才能突破，则算力支出的高景气可能不可持续；
- 2) **云计算竞争加剧的风险**：云计算领域有较多参与者，未来市场竞争可能加剧；
- 3) **政策落地不及预期风险**：政务及部分行业主要依靠各地政策推动，若政策落地不及预期，则可能影响相关公司；
- 4) **测算主观性风险**：就市场空间的测算而言，我们依据不同第三方信息来源进行假设、计算，测算可能会与实际情况偏差较大；
- 5) **大模型私有化部署渗透率不及预期的风险**：各行业私有化部署渗透率与AI应用落地速度以及下游AI预算相关，可能影响到最终渗透率。

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

THANKS