西南证券
SOUTHWEST SECURITIES

AI专题·从云计算看AI投资的ROI

企业上云具备性价比，云业务具备较高回报率

西南证券研究发展中心
海外研究团队
2024年10月

核心观点

- **小型企业及高成长性企业云化驱动力较强，云上部署相较于私有化部署具备更高性价比。**私有化部署即自建数据中心，云上部署则采用云服务提供商的配套服务。**1）部署考虑因素：**在AI时代下，GPU的供需缺口是导致众多企业难以进行私有化部署的关键，当前，众多中小企业和初创公司难以获得高性能芯片以自建算力集群；且云上部署相较于自建IDC，开通效率更高。从长期来看，企业自建IDC的情况下，业务曲线和资源曲线之间容易产生短缺和浪费，云上部署则可以根据业务情况灵活增减，实现资源按需付费，成本更加可控。**2）企业部署画像：**根据HG Insights数据，初创企业和小微企业多数选择上云为主，而大型企业和中型企业在上云的同时，还会选择自建数据中心。且当业务具备较高成长性时，企业可以选择分阶段、增量式上云。**3）部署成本测算：**基于AI时代下模型预训练的算力需求，我们对企业是否选择云上部署进行成本探讨，根据测算，在各种模型规模下，私有化部署成本远高于云上预训练成本。
- **云服务商加大投入力度，云业务具备较高回报率。1）云服务商投入力度：**亚马逊/微软/谷歌/甲骨文等大厂方面，资本开支持续增加，云计算基础设施加速布局；CoreWeave、Lambda等初创企业方面，近年来积极融资，以寻求更多算力资源。**2）CPU IaaS与GPU IaaS对比：**Semianalysis数据表明，GPU数据中心总拥有成本显著提升，在英伟达DGX H100服务器中，GPU成本中占比约7成，而内存和存储成本占比相较于CPU服务器显著下降；此外，在CPU IaaS时代，云计算通过虚拟化和容器等技术可实现资源的超卖，而在GPU IaaS时代，服务器在模型训练时通常处于满额利用状态，优化MFU成为提升可用算力的有效手段之一。**3）投资回报：**根据各厂商官网数据，H100的租赁价格从2$/h~13$/h不等，其中云服务大厂的算力租赁价格较为稳定。假设数据中心算力使用率为80%、且推出五折优惠，则云厂商每小时对应的实际收益为H100租赁价格的40%。若租赁价格为10$/h、对应实际收入为4$/h，扣除成本0.88$/h，利润率则可达78%，回本周期仅需1年。
- **相关标的：**英伟达(NVDA.O)、微软(MSFT.O)、亚马逊(AMZN.O)、谷歌(GOOGL.O)、甲骨文(ORCL.N)等。
- **风险提示：**市场需求不及预期；行业竞争加剧；投资回报不及预期等风险。

目 录

第一章 企业私有化部署和云上部署对比

资料来源：西南证券

1.1 企业私有化部署和云上部署的考虑因素

- **企业私有化部署对数据中心具有更高控制权，云上部署更具使用弹性。**私有化部署即自建数据中心，云上部署则采用云服务提供商的配套服务。在私有化部署情况下，企业数据不会通过公共网络传输，安全性更高，且不同企业可针对自身特定需求进行定制化部署，具备更大控制权。云上部署则无需投入大量初始资本以及后续运维费用，相关配套服务通常由云服务商统一提供，更加快捷易用，同时可根据业务需求进行扩展或缩减，使用弹性更加灵活。

企业私有化部署和云上部署优劣势对比

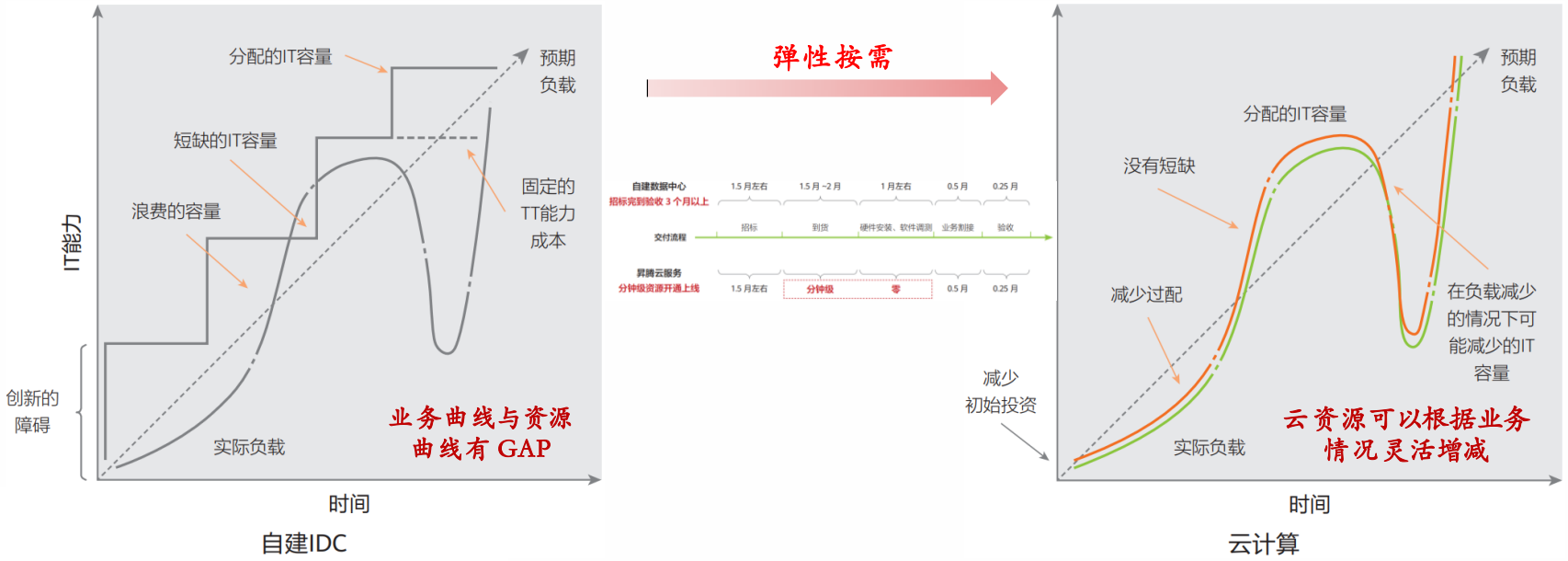
对比	私有化部署	云上部署
定义	私有化部署或自建GPU服务器是指将应用程序或服务部署在组织内部的私有网络中，仅对内部用户开放。	云上部署或采用GPU云服务器是指将应用程序或服务部署在第三方云服务提供商的数据中心，通过网络连接使用。
成本	① 高服务器投资运营成本； ② 设备高功耗，需硬件改造适配； ③ 为保障服务稳定，需运维IT成本；	① 按需购买，不用投入大量资金购置物理服务器； ② 可及时采用最新GPU服务器，无需硬件更新置换； ③ 无需投入服务器运维成本；
安全	① 数据不会通过公共网络传输，降低数据泄露的风险； ② 若不同用户共享资源，数据不隔离，需购买额外的安全防护服务等；	① 不同用户间的资源需要进行隔离，对客户数据采取配套的安全保障措施； ② 通常能够与与服务商的其他云安全产品实现无缝对接，享有云服务器同等的基础云安全基础防护和高防服务；
易用	① 购买装机管理，自行实现硬件扩展、驱动安装； ② 需跳板机登录，操作复杂；	① 与多种云产品接入，内网流量免费； ② 无需跳板机登录，简单易用； ③ 清晰的GPU驱动的安装、部署指引，免去高学习成本；
弹性	① 组织可完全控制部署的硬件和软件，满足内部特定需求； ② 机器固定配置，难以满足未来随着时间变化的需求。	① 云服务可根据业务需求进行扩展或缩减，提高资源利用率。

资料来源：腾讯云，西南证券整理

1.1 企业私有化部署和云上部署的考虑因素

- **GPU供不应求、部署周期较长或成为制约企业私有化部署的关键，使用弹性、成本优势是企业选择云上部署的长远考量。**在AI时代下，GPU的供需缺口是导致众多企业难以进行私有化部署的关键，当前，适用于AI的高性能GPU供不应求，众多中小企业和初创公司难以获得高性能芯片以自建算力集群；此外，根据《华为云昇腾AI云服务》数据，云上部署相较于自建IDC，开通效率更高，通常情况下，自建数据中心从招标到验收需要三个月以上，而云服务可做到分钟级资源开通。而从长期来看，企业自建IDC的情况下，业务曲线和资源曲线之间容易产生短缺和浪费，云上部署则可以根据业务情况灵活增减，实现资源按需付费，成本更加可控。

企业私有化部署和云上部署优劣势对比



资料来源：《华为云昇腾AI云服务》，西南证券整理

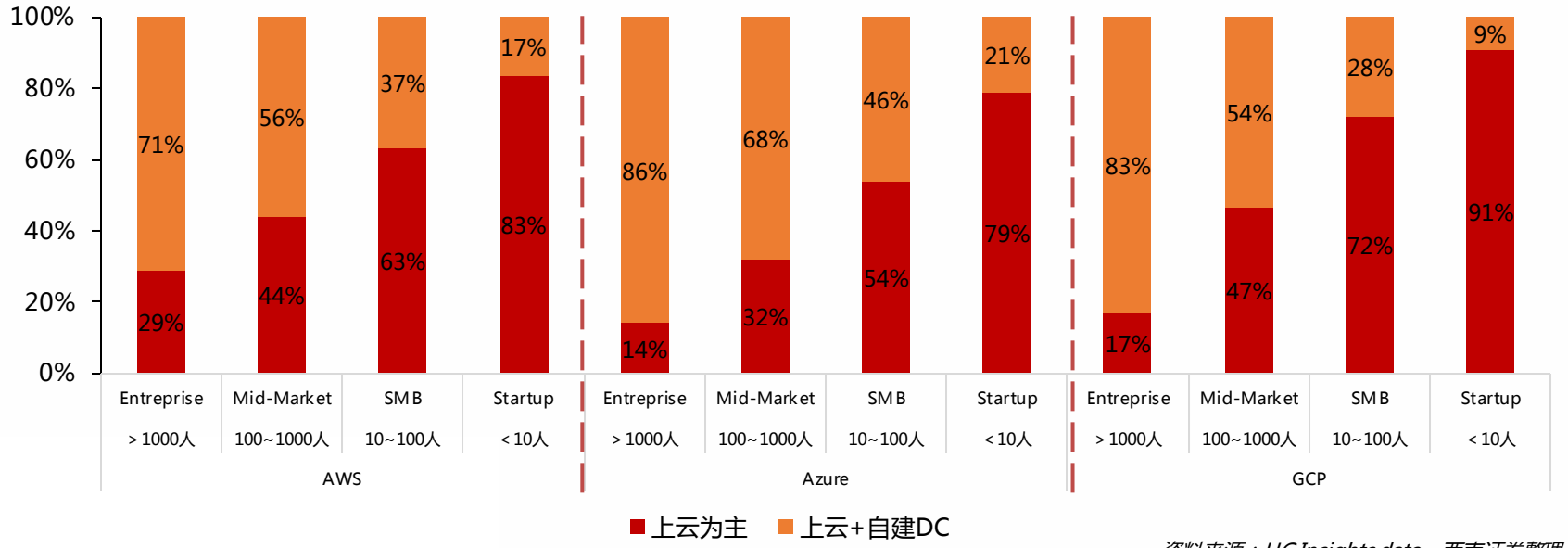
1.2 企业私有化部署和云上部署的客户画像

◆ **微小型企业选择上云为主，大型企业配备私有化部署**

① **大型企业自建数据中心，业务生态有望整合协同。**根据HG Insights数据，大型企业和中型企业在上云的同时，还会选择自建数据中心。三大云厂商的下游客户中，超过70%的大型企业（员工数量超过1000人）均会选择“上云+自建数据中心”的方案。由于大型企业具备较强的资金或资源实力，通常会围绕自身核心业务配备数据中心，以实现业务的生态协同和更高的成本效益。

② **小型企业选择上云为主，按需购买实现弹性易用。**根据HG Insights数据，小微企业（员工数量小于100人）以上云为主。由于小型企业云服务需求相对较小，且上云方案更加简单易用，只需按需订阅，因此初创企业和小微企业多数选择上云为主。

2024年云厂商下游客户上云和自建数据中心占比



资料来源：HG Insights data，西南证券整理

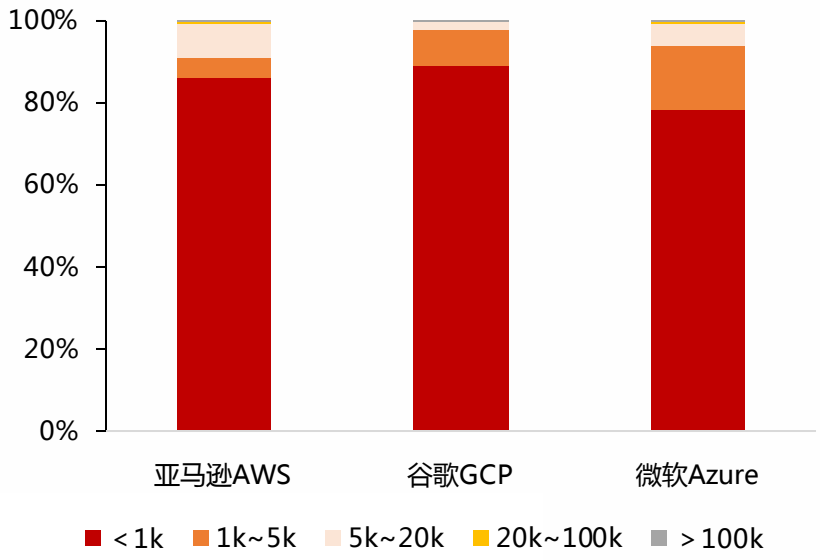
1.2 企业私有化部署和云上部署的客户画像

◆ **微型支出客户占比约八成，北美地区客户分布较多**

① **从客户支出水平来看：**根据HG Insights数据，**微型支出客户（月均支出小于1k美元）在各家云厂商中的占比可达75%~90%**；整体来看，谷歌的微型支出客户占比更高，亚马逊和微软的中大型支出客户占比更高。

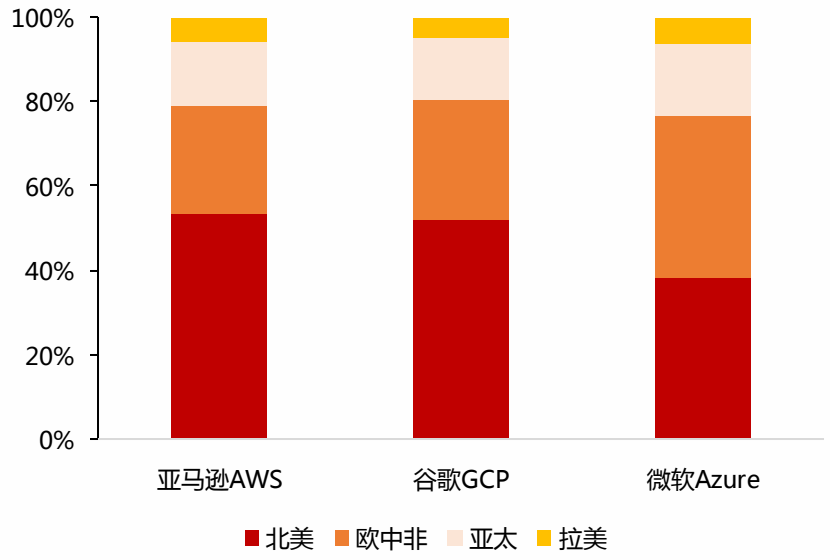
② **从客户地区分布上看：**根据HG Insights数据，**亚马逊AWS和谷歌GCP北美客户占比过半**，分别为53%和52%，高于微软的38%；而**微软Azure在欧中非地区具备相对优势**，客户占比为39%，亚马逊AWS和谷歌GCP分别仅为26%和28%。

2024年云厂商各支出水平客户占比（$/月）



资料来源：HG Insights data，西南证券整理

2024年云厂商全球各地区客户占比



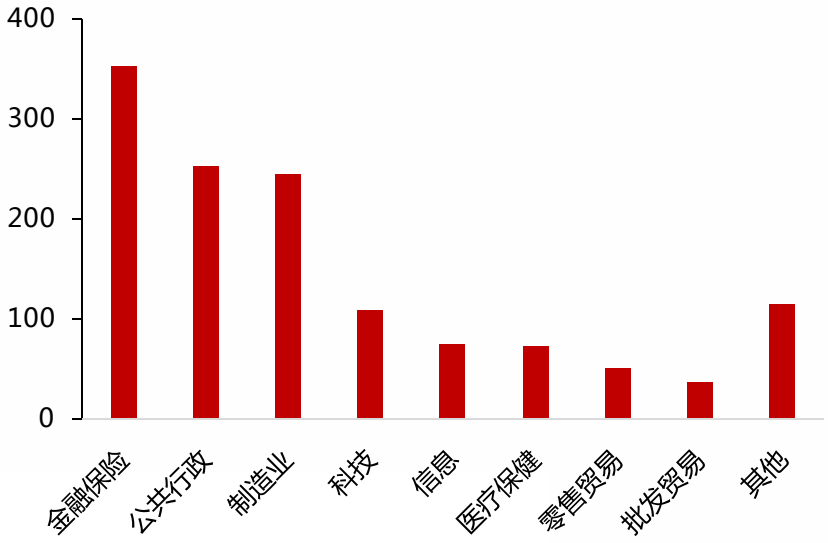
资料来源：HG Insights data，西南证券整理

1.2 企业私有化部署和云上部署的客户画像

- **稳态企业可选择自建机房，高成长性企业云化驱动力较强**

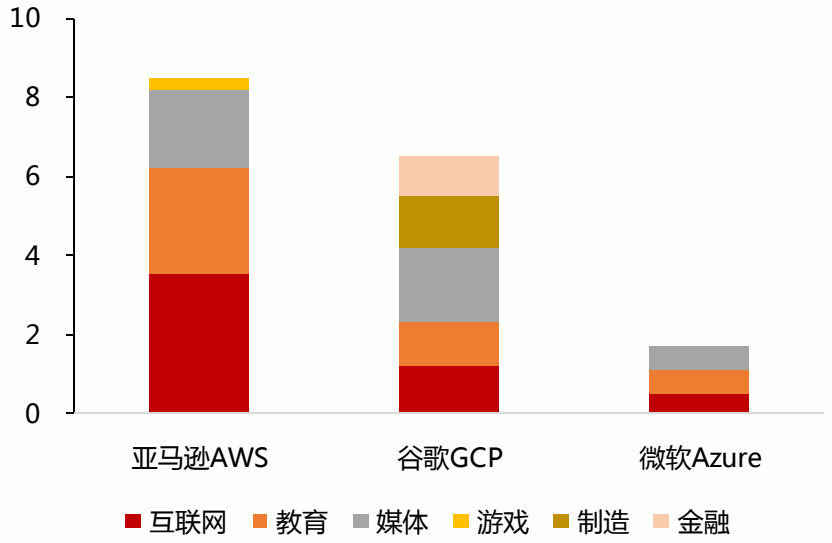
- **稳态企业可选择自建机房，高成长企业云化驱动力较强。**当企业业务长期稳定在某一存量水平，或核心业务需要基础设施配套协同时，自建机房不失为一种合理的选择；而当业务具备较高成长性时，企业可以选择分阶段、增量式上云，如互联网、游戏等行业，在开发阶段算力需求旺盛、系统需要快速上线，项目可能呈现爆发式增长，为满足业务需要，上云更具灵活性。根据HG Insights数据，金融、科技、医疗等行业的IT支出排名较为靠前，支付意愿较强。随着大语言模型的持续发展，金融、科技（代码）、医疗、法律以上四大行业可以通过混合专家模型等技术手段，在基座模型之上进行垂类专业能力的学习强化，推出具备更高成本效益的AI工具，为B端企业降本增效。

2024年主要行业IT支出预测 (亿美元)



资料来源：HG Insights data，西南证券整理

2024年云厂商行业买家监测数量 (万个)



资料来源：HG Insights data，西南证券整理

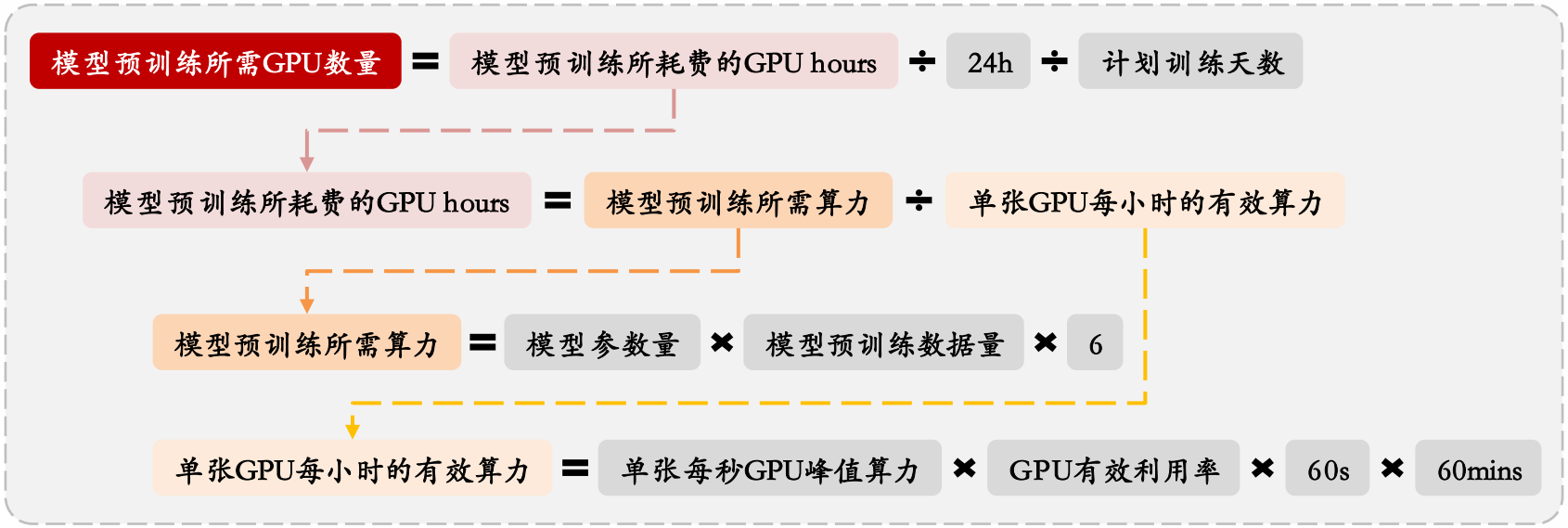
1.3 企业私有化部署与云上部署的成本探讨

- **当前，集群扩展的主要驱动因素在于千亿或万亿参数模型的预训练需求。在LLMs预训练阶段，需要将大量的训练数据并行至足够量的GPU上，而GPU的显存成为制约训练更大模型的主要条件之一。因此，我们基于AI时代下模型预训练的算力需求，对企业是否选择云上部署进行成本探讨。**

① **公式一：**模型预训练所需计算次数=6*模型参数量*预训练数据量。根据OpenAI《Scaling Laws for Neural Language Models》，训练Transformer模型的理论计算量为C≈6N*D，其中，N为模型参数量大小，D为训练数据量大小。

② **公式二：**模型预训练所用GPU hours=预训练所需计算次数/（单颗GPU每秒计算次数*60s*60mins*GPU算力的有效利用率）。

大语言模型预训练算力需求测算思路



资料来源：西南证券

1.3 企业私有化部署与云上部署的成本探讨

◆ **影响因素一：模型大小**

❑ **企业AI模型多为业务场景而设计，部署中等大小模型即可满足一般需求。**参考Meta的Llama系列模型，可分为大/中/小三种尺寸模型。从模型参数量来看，可分为7B/70B/400B梯队；从预训练数据量来看，可分为1T~2T和15T级别。从B端企业视角来看，AI模型主要为业务场景而设计，多数客户无需对模型规模进行无限扩展，70B参数大小即可满足一般需求，而7B小模型通常为端侧场景设计、千亿参数模型大多用作通用基座模型。因此，我们**基于70B的模型参数量、预训练数据量分别选取2.5T、5T、7.5T，对下游企业的模型预训练成本进行测算。**

Meta Llama系列模型预训练情况

模型发布日期	模型参数量(B)	预训练数据量(B tokens)	预训练所用GPU型号	预训练所用GPU hours
2023/2/24 Llama-1	7	1,000	A100	82,432
	13	1,000	A100	135,168
	33	1,400	A100	530,432
	65	1,400	A100	1,022,362
2023/7/18 Llama-2	7	2,000	A100	184,320
	13	2,000	A100	368,640
	34	2,000	A100	1,038,336
	70	2,000	A100	1,720,320
2024/4/18 Llama-3	8	15,000	H100	1,300,000
	70	15,000	H100	6,400,000
2024/7/23 Llama-3.1	8	15,600	H100	1,460,000
	70	15,600	H100	7,000,000
	405	15,600	H100	30,840,000

资料来源：Meta官网，西南证券整理

1.3 企业私有化部署与云上部署的成本探讨

◆ **影响因素二：GPU的峰值算力**

❑ **H100为当前云服务可用实例的领先产品，模型训练多采用半精度算力水平。**根据各个云厂商官网公布的云服务可用实例，H100是当前企业用户能够获得的更为先进的GPU产品，相较于A100、V100等芯片产品，H100的预训练效率更高、可扩展数量更多。且GPU采用的浮点精度不同，实际的算力水平也有较大差别，精度越高、可支持的运算复杂程度越高，而在AI模型的训练场景中，通常使用半精度浮点计算（FP16）。因此，我们将**以H100在FP16 Tensor核心性能下的算力水平为基础，对私有化部署和云上部署成本进行测算。**

用于大模型训练的GPU型号及计算性能

GPU型号	V100-SXM-16GB	A100-SXM-80GB	H100-SXM-80GB
设计架构	Volta	Ampere	Hopper
显存容量	16 GB HBM2	80 GB HBM2e	80 GB HBM3
显存带宽	900 GB/s	2039 GB/s	3.35 TB/s
INT8 Tensor 核心性能	/	624 TFLOPS	3958 TFLOPS
BF16 Tensor 核心性能	/	312 TFLOPS	1979 TFLOPS
FP16 Tensor 核心性能	125 TFLOPS	312 TFLOPS	1979 TFLOPS
TF32 Tensor 核心性能	/	156 TFLOPS	989 TFLOPS
FP64 Tensor 核心性能	/	19.5 TFLOPS	67 TFLOPS
FP32 计算性能	15.7 TFLOPS	19.5 TFLOPS	67 TFLOPS
FP64 计算性能	7.8 TFLOPS	9.7 TFLOPS	34 TFLOPS
TDP	Up to 300W (configurable)	Up to 500W (configurable)	Up to 700W (configurable)

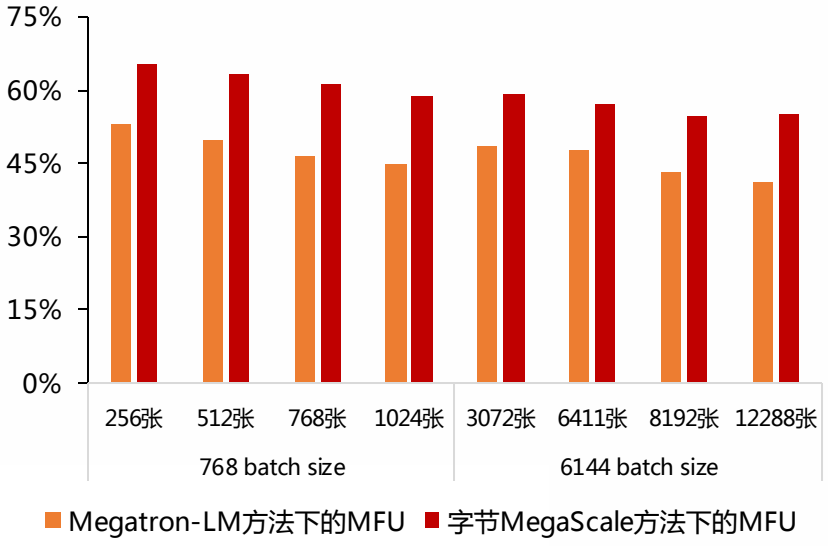
资料来源：英伟达官网，西南证券整理

1.3 企业私有化部署与云上部署的成本探讨

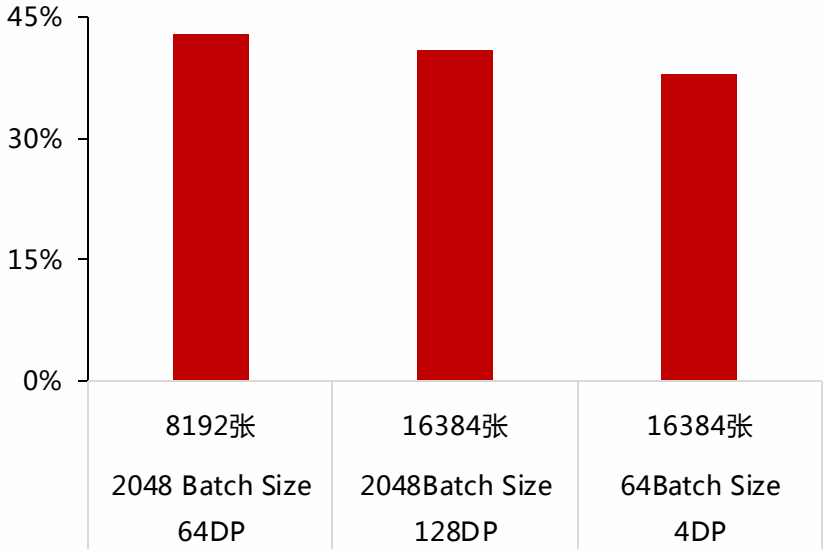
◆ **影响因素三：算力利用率（MFU）**

❑ **万卡集群MFU可达40%，GPU数量越少、MFU越高。**GPU的算力利用率（MFU）即GPU的实际吞吐量与理论峰值吞吐量的比。由于大语言模型预训练并非简单的并行任务，而是需要将模型分布在多个GPU之上，GPU之间需要频繁通信，叠加操作符优化、数据预处理、内存消耗等因素，GPU的MFU在实际训练中难以达到理论上的算力峰值水平。根据Meta Llama-3.1官方披露数据，在8K张和16K张GPU集群下，MFU分别可达到43%和41%的水平。根据字节的MegaScale论文，通过仔细调整并行性配置、硬件和软件，在BF16精度下，MegaScale方法可实现50%以上的算力利用率。此外，随着GPU数量的增加、算力集群的扩大，GPU算力的有效利用率呈现下降趋势。

字节MegaScale方法推动MFU提升



LLAMA-3.1训练GPU利用率



资料来源：《MegaScale: Scaling Large Language Model Training to More Than 10,000 GPUs》，Meta，西南证券整理

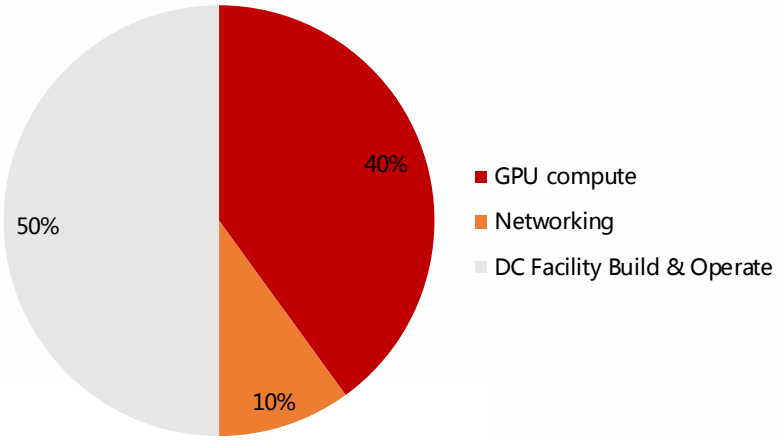
1.3 企业私有化部署与云上部署的成本探讨

◆ **影响因素四：GPU硬件成本、GPU租赁价格**

❑ **私有化部署成本：单张H100成本在2万美金至3.5万美金之间不等，GPU成本约占集群总拥有成本中约40%。**根据SemiAnalysis数据，单张H100成本约2.5万美金。按照英伟达官方对一个1.6万卡超大规模智算中心按照运营4年的计算，成本端需要投入10亿美元建设成本和运营成本，其中数据中心基建投资和运营维护成本约占一半、GPU购置成本为40%，网络成本约10%。

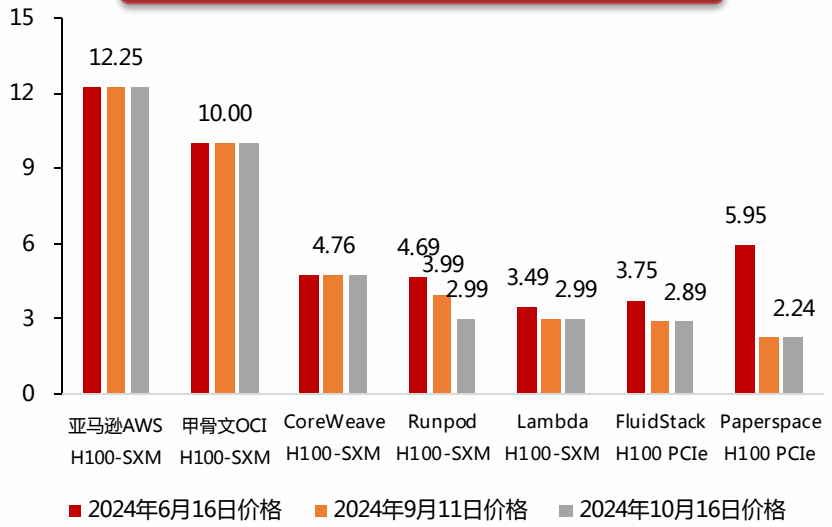
❑ **云上部署成本：H100云租赁价格从2$~13$/GPU/h不等。**根据各个云厂商单张GPU的每小时定价，头部厂商得益于基础设施更加安全完善、PaaS层配套工具更加全面易用，具备更高议价能力，H100云租赁定价基本在10$/GPU/h以上。而算力租赁初创企业定价基本位于2$~5$/GPU/h区间，具备低价优势。

16k张H100集群4年生命周期成本拆分



资料来源：英伟达官网，西南证券整理

云厂商H100租赁价格（$/GPU/h）



资料来源：各公司官网，西南证券整理

1.3 企业私有化部署与云上部署的成本探讨

◆ 测算结果

❑ 基于以上因素假设，根据测算，**私有化部署成本远高于云上预训练成本。**在各种模型规模下，私有化部署中的GPU购置成本为大模型云上预训练成本的数倍。

测算	指标	假设1	假设2	假设2
大模型预训练算力需求测算	模型参数量（B）		70	
	预训练数据量（B tokens）	2500	5000	7500
	预训练所需FLOPs	1.1E+24	2.1E+24	3.2E+24
	GPU型号		H100-SXM-80GB	
	FP16每秒计算次数（TFLOPS）		1979	
	GPU有效利用率（MFU）	*48%*	*45%*	*40%*
	模型预训练所需GPU hours	**307,043**	**655,026**	**1,105,356**
	GPU数量	*512*	*1024*	*2048*
	模型预训练所需天数	**25**	**27**	**22**
私有化部署成本测算	H100价格（万）		$2.5	
	GPU总成本（万）	**$1,280**	**$2,560**	**$5,120**
	GPU成本占总拥有成本比例	*40%*	*40%*	*40%*
	私有化部署机房总拥有成本（万）	**$3,200**	**$6,400**	**$12,800**
云上预训练成本	AWS租赁价格（$/GPU/h）		$12.25	
	Oracle租赁价格（$/GPU/h）		$10.00	
	CoreWeave租赁价格（$/GPU/h）		$4.76	
	Lambda租赁价格（$/GPU/h）		$2.99	
	平均租赁价格（$/GPU/h）		*$7.50*	
	预训练所需金额（$万）	**$230**	**$491**	**$829**

资料来源：西南证券

目 录

第二章 AI时代下的云基建与投资回报

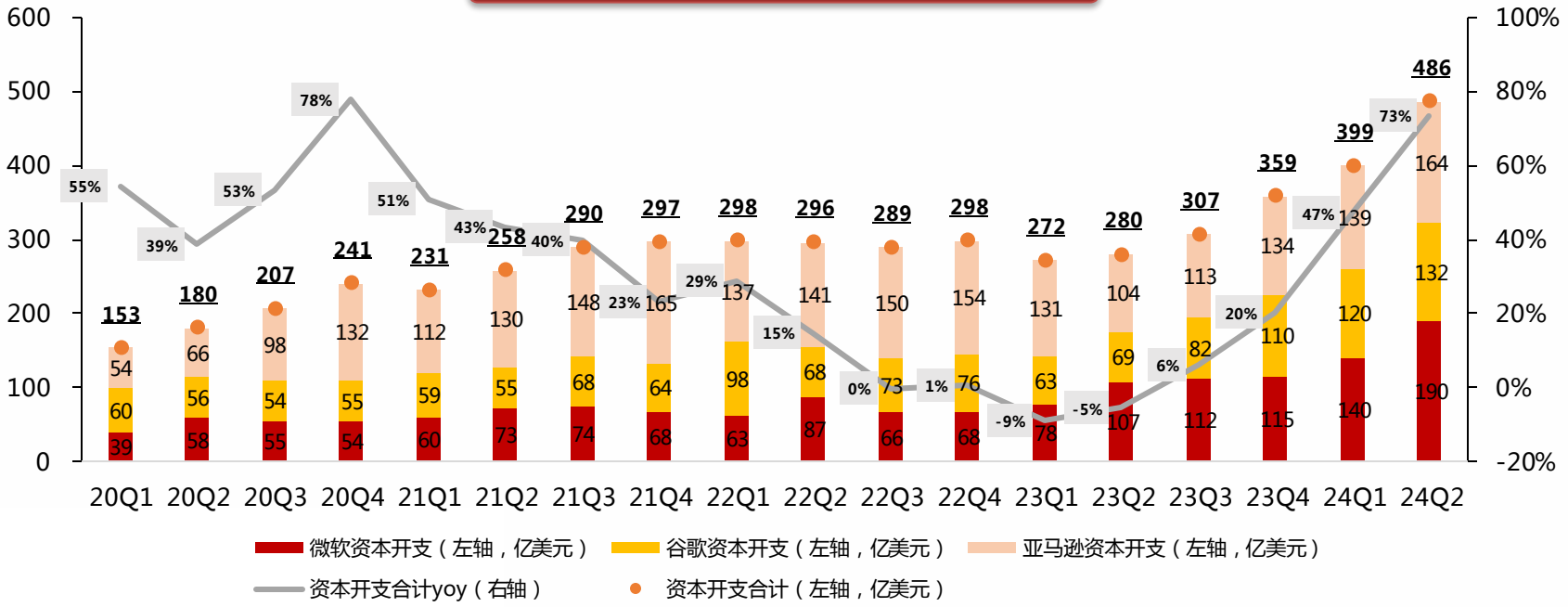
资料来源：西南证券

2.1 云服务商投入力度

◆ **第一梯队——亚马逊AWS / 微软Azure / 谷歌GCP**

❑ **资本开支加速增长，云设施投入占比较高。**23Q3至24Q2，三大云厂商合计资本开支从307亿美元增长至486亿美元，各季度同比增速分别为6%/20%/47%/73%，**呈现加速增长态势。**根据各公司24Q2业绩会，微软FY2024资本开支中GPU和CPU支出占比约一半、基础设施建设占一半；谷歌的资本开支增长主要由基础设施驱动，最大的部分是服务器，其次是数据中心；亚马逊资本开支中的大部分资金用于支持不断增长的AWS基础设施需求；且云厂商的GPU订单约占英伟达FY25Q2数据中心业务收入的45%，表明**云厂商的资本开支主要用于云计算基础设施建设。**

20Q1-24Q2三大云厂商资本开支及增速

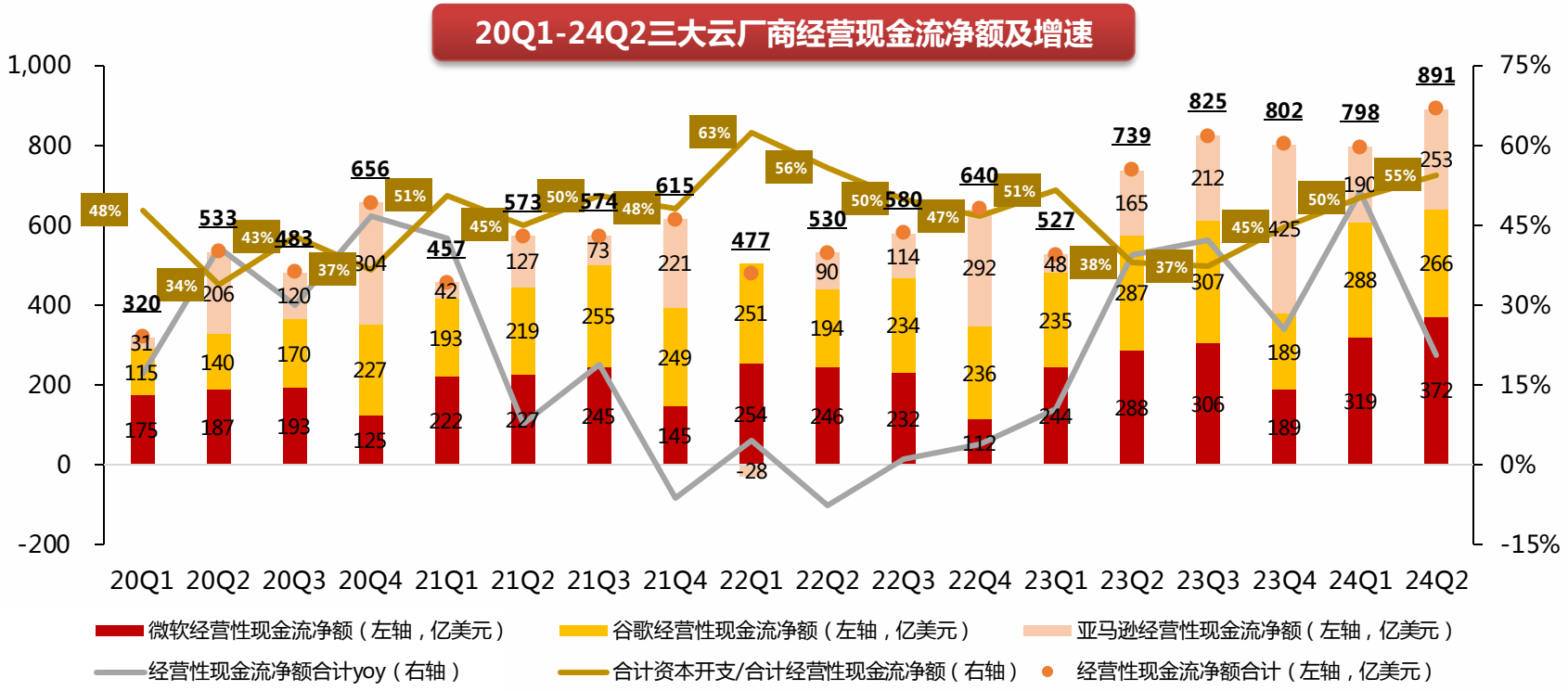


数据来源：各公司公告，西南证券整理

2.1 云服务商投入力度

- ◆ **第一梯队——亚马逊AWS / 微软Azure / 谷歌GCP**

- ❑ **经营性现金流净额持续增长，整体资本开支占比处于正常区间。**20Q1-24Q2，三大云厂商经营性现金流净额从320亿美元增长至891亿美元，合计资本开支占合计经营性现金流净额的比例通常处于30%~65%区间，而24Q2这一比例为55%，处于历史表现区间的正常值。由于以上海外科技厂商的主营业务在各自领域中均处于领先地位、具备较高护城河，每年带来的可观的运营现金流量为其投资AI数据中心提供强大的资金支持，有足够的资金支持其资本开支的提升。

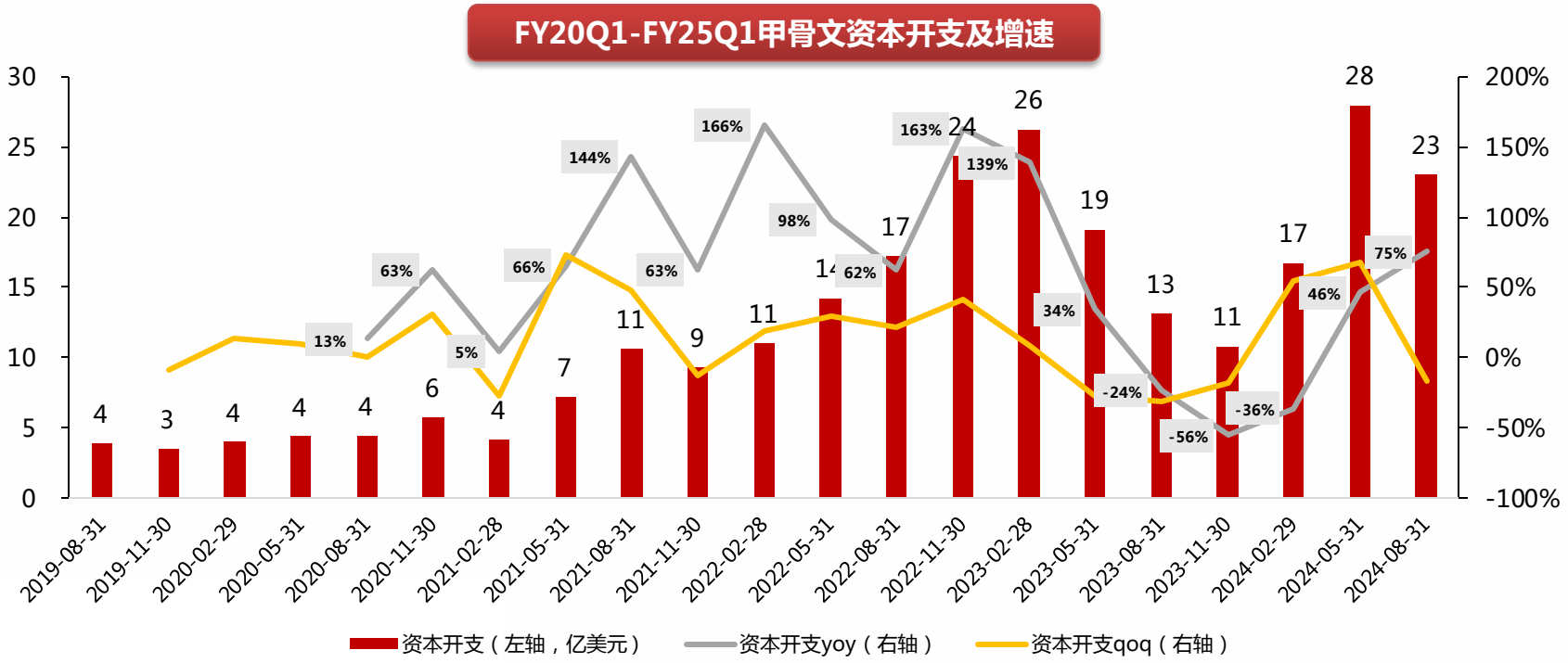


数据来源：各公司公告，西南证券整理

2.1 云服务商投入力度

◆ **第二梯队——甲骨文OCI**

❑ **资本支出有望翻倍，多云合作持续拓宽。**根据公司业绩会，甲骨文FY25Q1（截至2024年8月31日季度）资本开支达23亿美元，yoy+75%，新签42份云订单，价值量超30亿美元。公司表示，鉴于业务需求较为旺盛，**预计FY2025资本开支将是FY2024（69亿美元）的两倍**，未来公司也将基于剩余履约订单（PRO）趋势，谨慎且适当地调整投资步伐。此前，公司已与微软Azure和谷歌GCP建立多云合作，预计很快**将与亚马逊AWS实现合作**，云区域的规模灵活性和部署选项的多样性有望为公司持续强化竞争优势。



数据来源：Wind，西南证券整理

2.1 云服务商投入力度

◆ **第三梯队：初创企业CoreWeave**

❑ **融资额超百亿美元，客户及算力资源丰富。**CoreWeave成立于2017年，于2019年从以太坊挖矿转型为算力租赁业务。根据CrunchBase数据，公司自成立以来共进行10次融资，**合计筹集121亿美元**，其中，公司在业务转型后实现9轮融资，投资方包括Magnetar Capital、Blackstone、BlackRock、Coatue等投资机构。近年来，**公司在AI大客户方面展现出一定优势**，2023年6月**Inflection AI**宣布其与公司在MLPerf上的合作，2024年3月公司与**Mistral AI**在英伟达GTC大会上展示合作伙伴关系。而得益于优质的客户资源，CoreWeave逐步**受到英伟达关注**，2023年4月，英伟达在B轮融资中进行参投，并随后为其提供H100等先进GPU产品，帮助其构建优质计算集群。在客户优势和算力资源的支持下，公司已成为云计算和算力租赁行业的领先独角兽。

初创企业CoreWeave融资历程

时间	融资额 (亿$)	轮次	投资方
2024年5月17日	75	债务融资	由Blackstone管理的基金领投，Magnetar、Coatue、Carlyle、CDPQ等参投
2024年5月1日	11	C轮	由Coatue领投，Magnetar、Altimeter Capital、Fidelity等参投
2023年12月4日	6.42	二级市场	早期股东将6.42亿美元股份出售给投资方；由Fidelity领投，IMCO/JPM/Zoom Ventures等参投
2023年8月3日	23	债务融资	以英伟达芯片作为抵押品；Magnetar Capital/Blackstone/BlackRock等参与融资
2023年5月31日	2	B轮	Magnetar Capital加投4亿美元，B轮共筹集4.21亿美元融资
2023年4月20日	2.21	B轮	由Magnetar Capital领投，英伟达/前GitHubCEO Friedman/前苹果高管Daniel Gross参投
2022年12月6日	1	债务融资	Magnetar Capital
2021年11月10日	0.5	可转换票据	Magnetar Capital
2019年10月1日	0.025	A轮	/
2018年7月30日	0.03	种子轮	/

资料来源：CrunchBase，西南证券整理

2.1 云服务商投入力度

◆ **第三梯队：初创企业Lambda**

❑ **定位AI开发者云，芯片获取优势明显。**Lambda Labs成立于2012年，主营业务包括云服务和深度学习工作站，致力于为工程师和研究人员提供支持。根据CrunchBase数据，Lambda共进行9轮融资，**合计筹集资金9.322亿美元。**其中，Lambda在种子轮/A轮/B轮融资中，均获得由谷歌AI中心基金Gradient Ventures的投资，推动其完善在云计算和软件服务方面的发展。2023年11月，**英伟达表示，Lambda将成为首批允许客户享用其H200芯片的云服务公司之一。**2024年4月5日，公司以芯片资产作为抵押，从Macquarie Group等贷款机构获得5亿美元贷款，并计划利用该笔资金继续采购GPU芯片，以扩展其云计算的规模。2024年7月，Lambda推出1-Click Clusters，使AI开发人员能够立即访问带有英伟达InfiniBand网络的H100 TensorCore GPU集群。公司**客户主要包括科研或学术机构、开源社区**，以满足GPU云租赁**长尾市场的多样化需求。**

初创企业Lambda融资历程

时间	融资额 (亿$)	轮次	投资方
2024年4月4日	5	债务融资	由Macquarie Group领投，Industrial Development Funding参投
2024年2月15日	3.2	C轮	由USIT领投，B Capital/SK Telecom/T. Rowe Price Associate/Crescent Cove等参投
2023年10月27日	/	C轮	由Alumni Ventures领投，Alpha Square Group参投
2023年3月21日	0.44	B轮	由Mercato Partners领投，Bloomberg Beta/Gradient Ventures/1517 Fund等参投
2022年11月17日	0.397	风险投资轮次	/
2021年7月14日	0.15	A轮	由Bloomberg Beta、Gradient Ventures、1517 Fund、Invariantes Fund等投资
2021年7月14日	0.095	债务融资	由Silicon Valley Bank领投
2019年4月28日	0.04	种子轮	由Gradient Ventures领投
2017年12月20日	/	预种子轮	/

资料来源：CrunchBase，西南证券整理

2.2 CPU IaaS VS GPU IaaS

- **数据中心总拥有成本对比：GPU数据中心总拥有成本显著提升，硬件资本开支远高于托管成本。**根据semianalysis数据，通过对比CPU和GPU数据中心总拥有成本，可以发现，在CPU IaaS时代下，单个CPU核心的成本为0.006$/h，而GPU IaaS时代下，单张GPU的成本为1.524$/h。根据semianalysis数据，单台CPU服务器的托管成本（$220/月）与资本成本（$301/月）差距相对较小，而单台GPU服务器的资本成本（$7026/月）远高于托管成本（$1872/月）。

CPU数据中心总拥有成本（TCO）拆分

总拥有成本（Total Cost of Ownership）拆分	CPU服务器
1）单台服务器采购成本	
单台服务器OEM售价（$）	15,000
生命周期（年）	6
资本成本	13%
单台服务器资本成本（$/月）	**301**
2）单台服务器托管成本	
美国工业平均零售电价（$/KWh）	0.087
使用率	80%
PUE	1.25
① 电力成本（$/KWh/月）	63.5
② 托管空间成本（$/KWh/月）	120
合计托管成本（$/KWh/月）	**183.5**
单台服务器最高功耗（KWh）	1.2
单台服务器托管成本（$/月）	**220**
合计｜单台服务器成本	
单台服务器采购+托管成本（$/月）	**521**
单台服务器包含的CPU核心或GPU的数量	128
单位CPU核心或单位GPU的成本（$/h）	**0.006**
服务器托管成本：服务器资本成本	**73%**

资料来源：semianalysis，西南证券整理

GPU数据中心总拥有成本（TCO）拆分

总拥有成本（Total Cost of Ownership）拆分	GPU服务器
1）单台服务器采购成本	
单台服务器OEM售价（$）	350,000
生命周期（年）	6
资本成本	13%
单台服务器资本成本（$/月）	**7,026**
2）单台服务器托管成本	
美国工业平均零售电价（$/KWh）	0.087
使用率	80%
PUE	1.25
① 电力成本（$/KWh/月）	63.5
② 托管空间成本（$/KWh/月）	120
合计托管成本（$/KWh/月）	**183.5**
单台服务器最高功耗（KWh）	10.2
单台服务器托管成本（$/月）	**1,872**
合计｜单台服务器成本	
单台服务器采购+托管成本（$/月）	**8,898**
单台服务器包含的CPU核心或GPU的数量	8
单位CPU核心或单位GPU的成本（$/h）	**1.524**
服务器托管成本：服务器资本成本	**27%**

资料来源：semianalysis，西南证券整理

2.2 CPU IaaS VS GPU IaaS

- **服务器成本构成对比：GPU服务器存储成本占比下降，GPU成本占比约七成。**在AI兴起之前，数据中心主要基于CPU服务器搭建，CPU擅长线性任务，其高频单核性能是关键。然而，随着AI的高速发展，云计算能力面临新的挑战。在深度学习等领域，需要处理大规模的数据、复杂的算法、以及海量的计算，由于GPU擅长同时执行矩阵和向量计算等任务，在AI场景中展现出强大的并行计算能力和运行效率，价值逐步凸显。根据semianalysis数据，通过对比2x Intel Sapphire Rapids CPU标准服务器和Nvidia DGX H100 GPU服务器的成本构成，可以发现，在英伟达DGX H100服务器中，GPU成本中占比约7成，而内存和存储成本占比相较于CPU服务器中的占比显著下降，在AI时代的云计算基础设施中，GPU的重要性实现大幅提升。

2x Intel Sapphire Rapids服务器成本构成

组件	成本（美元）	占总成本比例
GPU	/	/
CPU	1,850	18%
内存(Memory)	3,930	38%
存储(Storage)	1,536	15%
网卡SmartNIC	654	6%
机箱（外壳/背板/电缆）	395	4%
主板	350	3%
散热（散热器+风扇）	275	3%
电源	300	3%
组装及测试	495	5%
Markup	689	7%
成本合计	**10,474**	**100%**

资料来源：semianalysis，西南证券整理

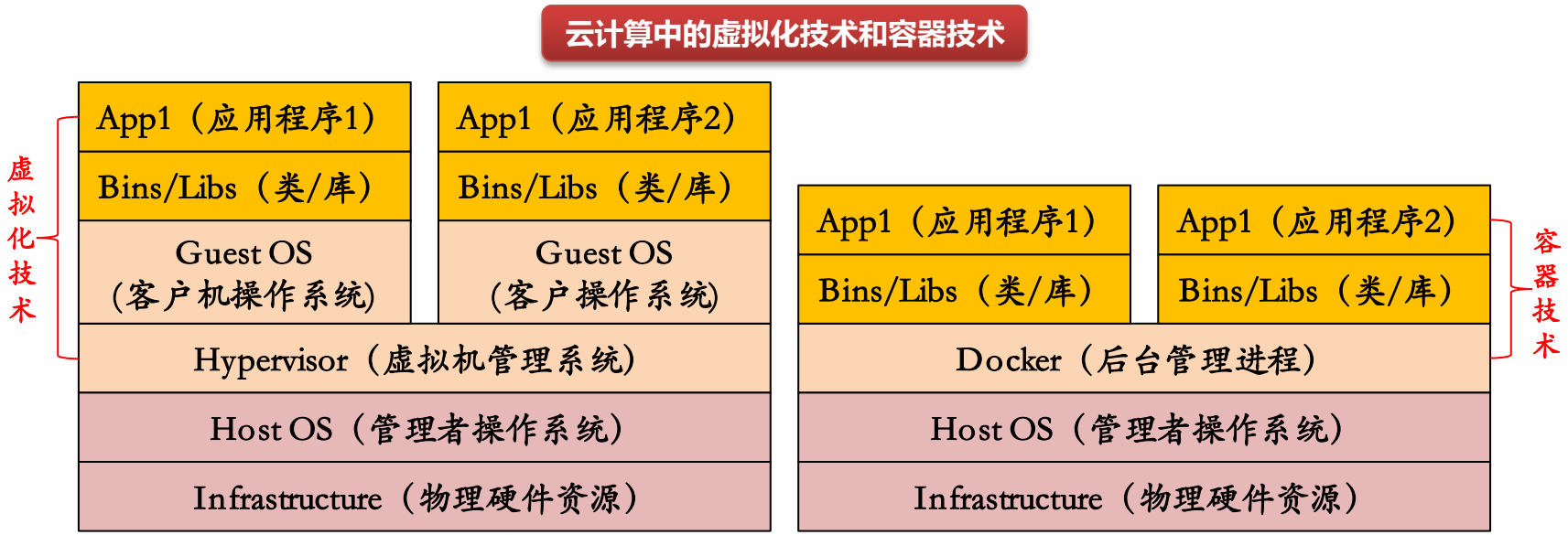
Nvidia DGX H100服务器成本构成 (不含HBM)

组件	成本（美元）	占总成本比例
8 GPU + 4 NV Switch Baseboard	195,000	72%
CPU	5,200	2%
内存(Memory)	7,860	3%
存储(Storage)	3,456	1%
网卡SmartNIC	10,908	4%
机箱（外壳/背板/电缆）	563	0%
主板	875	0%
散热（散热器+风扇）	463	0%
电源	1,200	0%
组装及测试	1,485	1%
Markup	42,000	16%
成本合计	**269,010**	**100%**

资料来源：semianalysis，西南证券整理

2.2 CPU IaaS VS GPU IaaS

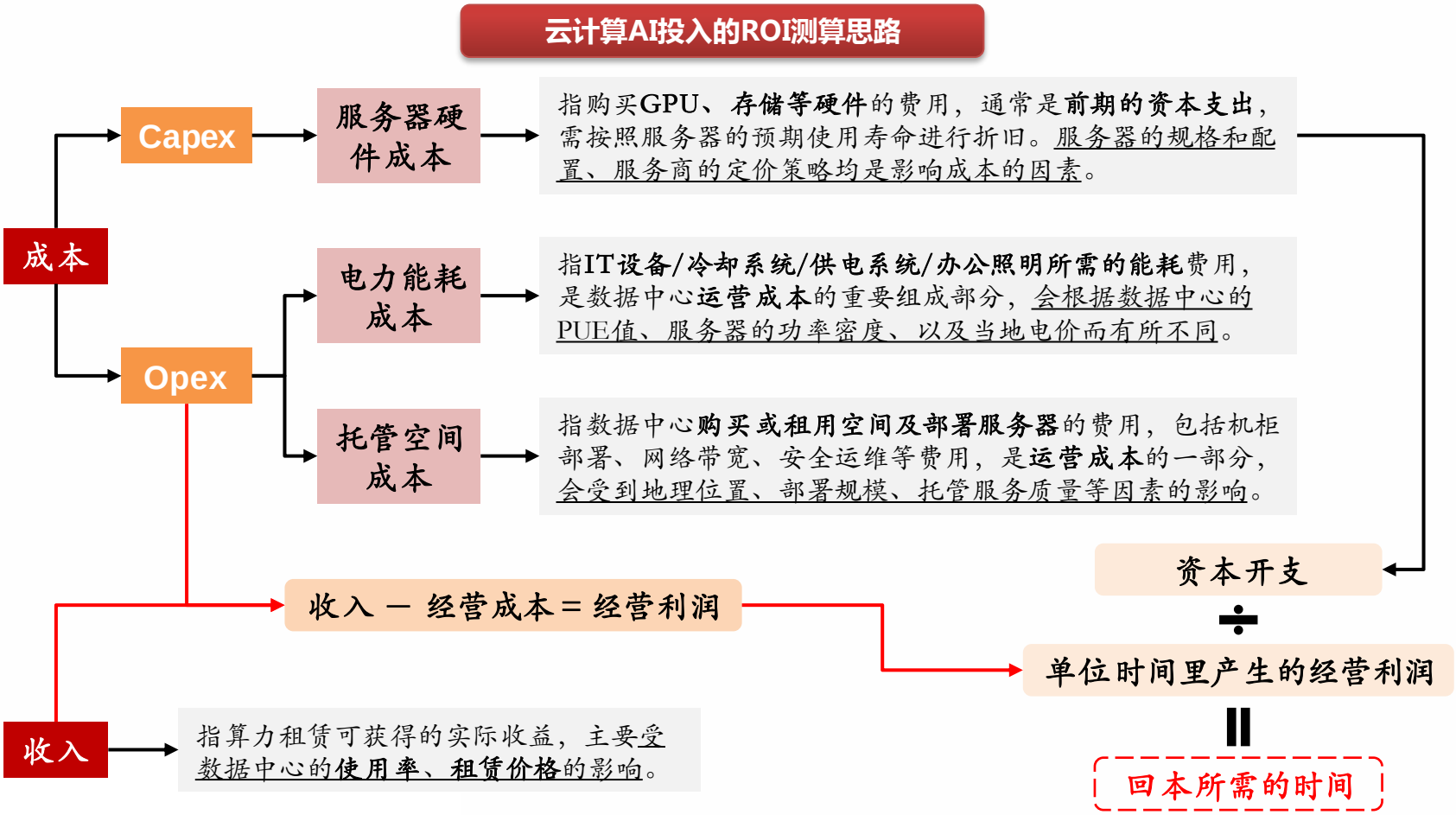
- **CPU超卖率 VS GPU利用率：GPU服务器在AI训练时通常处于满额负载状态，优化使用率（MFU）成为提升可用算力的有效手段之一。**在CPU IaaS时代，云计算通过虚拟化和容器等技术，将处理器、内存、网络等多种硬件资源打包成硬件池，**根据不同客户的使用量及使用时间段，分配给多个客户使用，提升硬件使用效率。**云厂商通过以上技术手段，使云计算服务的使用不再局限于物理资源的限制，能够基于服务器硬件实体实现算力资源的超卖，从而提高业务的经营利润率。而在GPU IaaS时代，AI需要训练算力和推理算力，从训练端来看，通常**服务器在模型训练时处于满额利用状态**，难以实现算力的灵活调用，需要等到训练完成，才能将其所用算力释放出来，基于该情况，云厂商需要通过**提升MFU等方法实现可用算力资源的扩充**，提升集群的收益和利润水平。目前，算力的有效利用率基本处于30%至50%之间，而在技术优化后，MFU有望达到60%以上的水平。**未来，提升算力利用率、实现集群性能调优将是云厂商核心竞争力的体现。**



资料来源：网络技术联盟站，西南证券整理

2.3 GPU云服务投资回报率测算

- 随着云厂商在AI时代对GPU IaaS的投入持续加大，AI投资的回报率逐渐成为市场关注并亟需解答的问题，我们基于以下测算思路，对相关影响因素进行定量分析。



资料来源：西南证券

2.3.1 成本测算

◆ **硬件成本影响因素——服务器折旧年限**

❑ **科技大厂折旧年限延长，服务器折旧年限调整为六年。**近年来，海外科技大厂随着对AI及业务资源投入的增加，逐渐面临成本压力。**各个公司在放缓人员招聘速度、实行组织提效的同时，也在重新评估设备资产的预期使用寿命、调整折旧年限**，通过会计方法减少折旧费用、增加账面利润，以减轻不断增长的成本压力。根据各公司公告，**谷歌于2023年1月开始**，将服务器折旧年限从4年延长至6年，特定网络设备的预期使用寿命从5年延长至**6年**，帮助公司2023年折旧费用减少39亿美元、净利润增加30亿美元；而**微软自2022年7月开始**，即执行**6年**的服务器和网络设备折旧；亚马逊则于2022年1月将服务器寿命从4年延长至5年，使其22年折旧成本减少36亿美元、利润增加28亿美元，此外，根据**亚马逊24Q1**业绩会，公司已再次延长服务器折旧年限，从5年调整为**6年**，预计该举措将进一步优化未来公司的利润表现。

2020-2024年海外云厂商折旧政策调整

公司	时间	楼房建筑	办公设备	机器设备	服务器	网络设备	无形资产
谷歌	2020年12月及以前	7~25年	/	/	3年	3年	2-3年
	2021年1月至2022年12月				4年	5年	
	2023年1月至今				**6年**	**6年**	
微软	2022年6月及以前	5~15年	1~10年	3~20年	4年	4年	3-7年
	2022年7月至今				**6年**	**6年**	
亚马逊	2021年12月及以前	＜40年	5年	10年	4年	5年	2年
	2022年1月至2023年12月				5年	**6年**	
	2024年1月至今				**6年**	/	

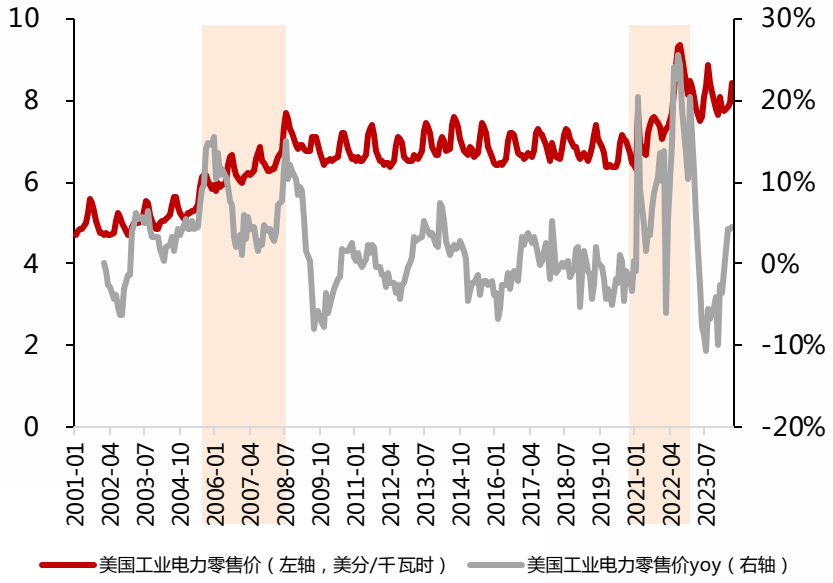
资料来源：各公司公告，各公司业绩会，西南证券整理

2.3.1 成本测算

- **能耗成本影响因素——电力价格与PUE**

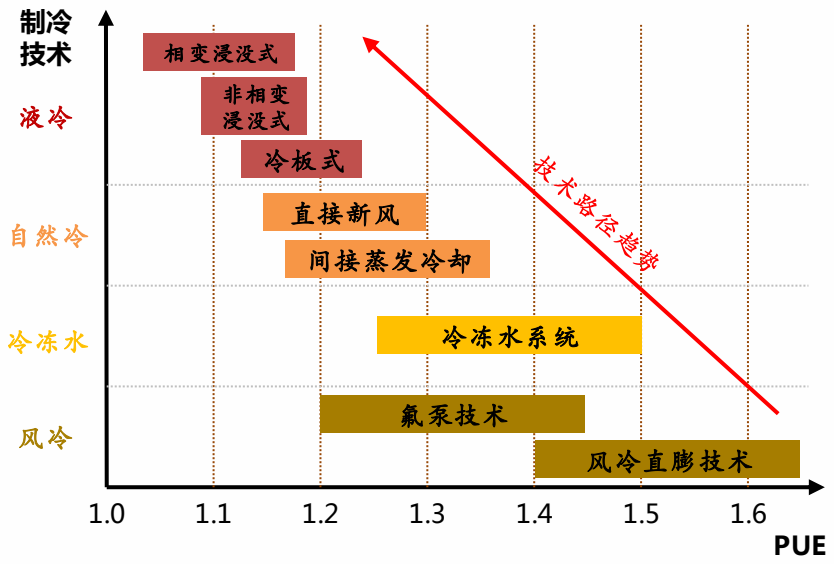
- **美国工业电价近年趋于稳定，数据中心PUE有望持续优化。**2001年1月至2024年6月，美国工业电力主要经历两轮涨价，第一轮涨价从2005年均价5.72美分/千瓦时上涨至2008年的6.95美分/千瓦时，第二轮涨价从2020年的6.66美分/千瓦时上涨至2022年的8.29美分/千瓦时，而后工业电价逐步趋于稳定，**24H1平均约为8美分/千瓦时。**而在典型的数据中心能耗系统中，根据《中兴通讯液冷技术白皮书》数据，IT设备/制冷系统/供电系统/办公照明的能耗占比分别为65%/24%/8%/3%，制冷系统是除IT设备以外的辅助能源占比中最大的部分，因此，**降低制冷系统能耗能够推动数据中心更加节能。**未来，**随着人工智能数据中心的制冷系统向液冷技术升级，PUE有望持续降低。**

2001年至今美国工业电力零售价



资料来源：Wind，西南证券整理

数据中心制冷技术对应PUE范围



资料来源：《中兴通讯液冷技术白皮书》，西南证券整理

2.3.1 成本测算

- **成本测算：中性假设下，每小时单张H100分摊的硬件、能耗、托管空间的合计成本约为0.88美元。**

① **服务器硬件成本：**根据semianalysis数据，一台8卡的英伟达DGX H100服务器总成本约26.9万美元，假设使用寿命6年，则**每小时单张H100分摊的硬件成本为0.64美元。**

② **服务器电力成本：**根据英伟达官网，DGX H100服务器系统的功耗峰值为10.2KW，假设有80%的使用率、PUE为1.25、美国工业电力均价为8美分/千瓦时，则单张H100**能耗成本为$0.1/h。**

③ **托管空间成本：**根据substack数据，托管行业的标准市场价格约为125美元/千瓦/月，若扣除托管提供商的毛利率，则海外大型云厂商数据中心园区的托管空间成本可能接近80美元/千瓦/月，基于单台服务器的系统功耗，则分摊至单张H100的**托管空间成本为$0.14/h。**

英伟达 H100 GPU IaaS 总拥有成本（TCO）测算

一、服务器硬件成本	自建数据中心的情况
CPU	5,200
8 GPU +4 NVSwitch Baseboard	195,000
Memory	7,860
Storage	3,456
SmartNIC	10,908
Chassis (Case, backplanes, cabling)	563
Motherboard	875
Cooling (Heatsinks+fans)	463
Power Supply	1,200
Assembly and Test	1,485
Markup	42,000
单台英伟达DGX H100服务器成本（$）	**269,010**
生命周期（年）	*6*
单台服务器硬件成本（$/h）	**5.12**
单张H100硬件成本（$/h）	**0.64**

二、服务器能耗成本	悲观假设	中性假设	乐观假设
单台英伟达DGX H100服务器最高功耗（Wh）		10200	
使用率	*70%*	*80%*	*90%*
单台服务器使用功耗（Wh）	7140	8160	9180
单台服务器能源消耗（MWh/年）	63	71	80
PUE	*1.25*	*1.25*	*1.25*
总能源消耗（MWh/年）	78	89	101
美国工业平均零售电价（$/KWh）	*0.08*	*0.08*	*0.08*
总能耗成本（$/KWh/年）	**6,255**	**7,148**	**8,042**
单张H100能耗成本（$/h）	**0.09**	**0.10**	**0.11**
三、托管空间成本		**假设**	
单台服务器空间成本（$/KWh/月）		*80*	
单台英伟达DGX H100服务器最高功耗（Wh）		10200	
单张H100空间成本（$/h）		**0.14**	
单张H100“硬件+能耗+空间”成本（$/h）	**0.87**	**0.88**	**0.90**

资料来源：semianalysis，substack，西南证券整理

2.3.1 成本测算

- **自建数据中心TCO VS 外购服务器及外包托管空间TCO**：对比自建数据中心的总拥有成本，若云厂商向OEM外采服务器、向托管提供商外包空间及服务，在数据中心服务器80%的利用率下，总拥有成本约1.15$/h/GPU，高于自建数据中心的0.88$/h/GPU。

自建数据中心TCO VS 外购服务器及外包托管空间TCO

一、服务器硬件成本	自建数据中心的情况			外购服务器和外租托管空间的情况		
单台英伟达DGX H100服务器成本（$）	269,010			350,000		
生命周期（年）	*6*			*6*		
单台服务器硬件成本（$/h）	5.12			6.66		
单张H100硬件成本（$/h）	0.64			0.83		
二、服务器能耗成本	**悲观假设**	**中性假设**	**乐观假设**	**悲观假设**	**中性假设**	**乐观假设**
单台英伟达DGX H100服务器最高功耗（Wh）	10200			10200		
使用率	*70%*	*80%*	*90%*	*70%*	*80%*	*90%*
单台服务器使用功耗（Wh）	7140	8160	9180	7140	8160	9180
单台服务器能源消耗（MWh/年）	63	71	80	63	71	80
PUE	*1.25*	*1.25*	*1.25*	*1.25*	*1.25*	*1.25*
总能源消耗（MWh/年）	78	89	101	78	89	101
美国工业平均零售电价（$/KWh）	*0.08*	*0.08*	*0.08*	*0.08*	*0.08*	*0.08*
总能耗成本（$/KWh/年）	6,255	7,148	8,042	6,255	7,148	8,042
单张H100能耗成本（$/h）	0.09	0.10	0.11	0.09	0.10	0.11
三、托管空间成本	**假设**			**假设**		
单台服务器空间成本（$/KWh/月）	*80*			*120*		
单台英伟达DGX H100服务器最高功耗（Wh）	10200			10200		
单张H100空间成本（$/h）	0.14			0.21		
单张H100“硬件+能耗+空间”成本（$/h）	**0.87**	**0.88**	**0.90**	**1.13**	**1.15**	**1.16**

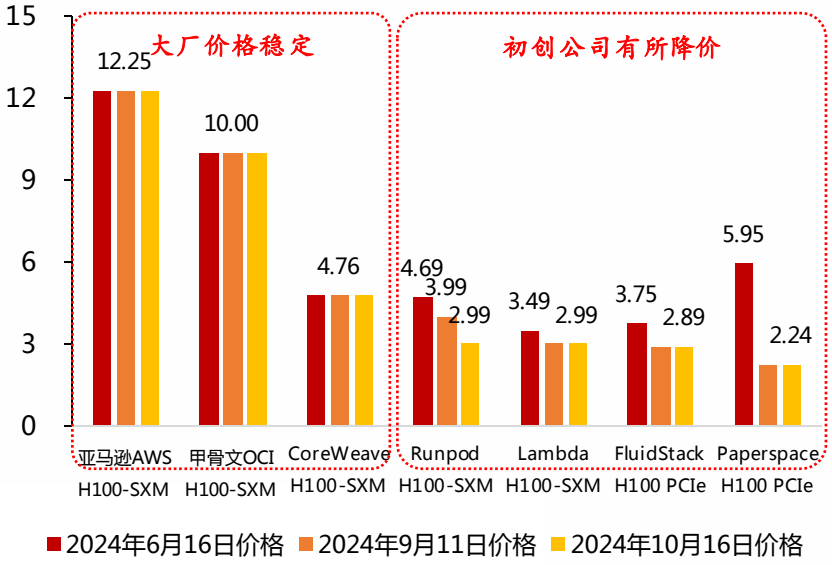
资料来源：semianalysis，substack，西南证券整理

2.3.2 收益测算

◆ **收入影响因素——租赁定价及折扣**

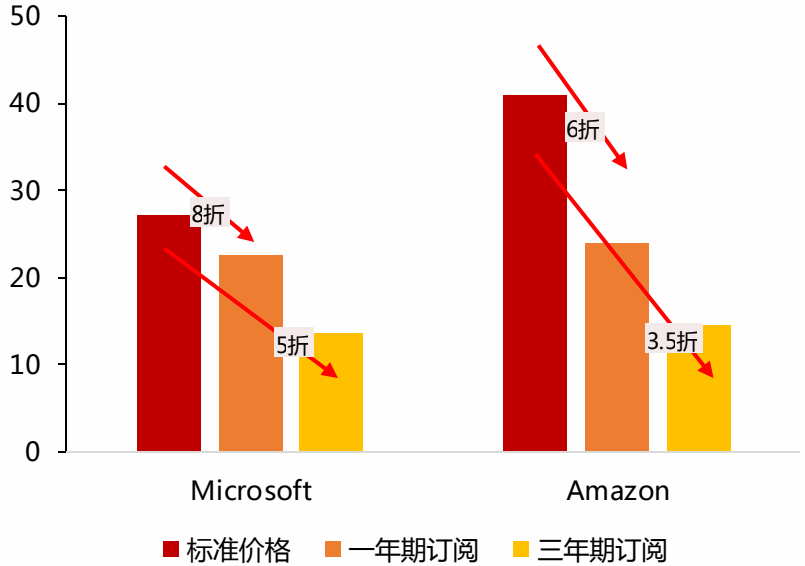
❑ **GPU更新周期促使过往代际产品折价，云厂商推出多年期订阅折扣。**按照英伟达迭代节奏，基本为两年一代际。未来随着GPU的更新升级，云厂商也将推出基于新款GPU的云计算服务，过去的虚拟机通常会面临一定的折价，新一代的云服务与上一代产品形成价格梯队，从而促进产品或服务间的差异化销售。从微软Azure和亚马逊AWS对上一代A100服务器的定价策略来看，对多年期订阅均给予一定优惠，微软基于其标准价格分别对一年期和三年期推出8折和5折优惠，亚马逊则分别推出6折和3.5这优惠。因此，在收益测算中，我们也将基于云厂商当前对H100的租赁价格，对不同折价情况下的收益和利润测算进行敏感性分析。

云厂商H100租赁价格（$/GPU/h）



资料来源：各公司官网，西南证券整理

云厂商A100单台服务器租赁价格（$/h）



资料来源：各公司官网，西南证券整理

2.3.2 收益测算

- **收益测算：假设80%的使用率、五折优惠，各个主流云厂商的H100算力租赁业务均能实现正收益。**
  - **收入端：**根据各个云厂商官网数据，H100的租赁价格从2$/h~13$/h不等，假设数据中心算力使用率为80%、且推出五折优惠，则云厂商每小时对应的实际收益为H100租赁价格的40%。
  - **利润端：**基于前文成本测算，可知GPU每小时的分摊成本主要包括Capex部分（硬件资本成本）和Opex部分（电力能耗成本、托管空间成本），云厂商将实际收益扣除以上成本后，仍具备较高利润率，例如，若租赁价格为10$/h、对应实际收入为4$/h，扣除成本0.88$/h，利润率可达78%。
  - **回本周期：**根据测算，若单张H100的租赁价格在10$/h，在80%的使用率和五折优惠下，回本周期仅需1年。

H100算力租赁收益及回本周期测算

云厂商	AWS	Oracle	CoreWeave	Runpod	Lambda	FluidStack	Paperspace
实际收益							
价格 ($/GPU/h)	12.25	10.00	4.76	3.99	2.99	2.89	2.24
使用率	*80%*	*80%*	*80%*	*80%*	*80%*	*80%*	*80%*
销售折扣（支付原价的百分比）	*50%*	*50%*	*50%*	*50%*	*50%*	*50%*	*50%*
实际收入 ($/GPU/h)	**4.90**	**4.00**	**1.90**	**1.60**	**1.20**	**1.16**	**0.90**
利润及利润率							
Capex+Opex成本 ($/GPU/h)	0.88	0.88	0.88	0.88	0.88	0.88	0.88
利润 ($/GPU/h)	4.02	3.12	1.02	0.71	0.31	0.27	0.01
利润率	**82%**	**78%**	**54%**	**45%**	**26%**	**24%**	**1%**
回本周期							
Capex支出 ($/GPU)	269,010	269,010	269,010	269,010	269,010	269,010	269,010
Opex成本 ($/GPU/h)	0.24	0.24	0.24	0.24	0.24	0.24	0.24
Opex利润	4.66	3.76	1.66	1.35	0.95	0.91	0.65
回本周期（月）	**10**	**12**	**28**	**35**	**49**	**51**	**72**

资料来源：各公司官网（采用2024年9月11日价格数据），西南证券整理

2.3.3 利润率测算

❑ **利润率敏感性分析：**基于80%的使用率，若租赁价格享受5折优惠，则各个云厂商均能实现正收益。

算力租赁利润率敏感性分析

支付原价的百分比		AWS	Oracle	CoreWeave	Runpod	Lambda	FluidStack	Paperspace
使用率					70%			
折扣后的租赁利润率	80%	87%	84%	67%	61%	48%	46%	31%
	60%	83%	79%	56%	48%	31%	28%	7%
	50%	80%	75%	48%	38%	17%	14%	-11%
	40%	75%	69%	35%	22%	-4%	-8%	-39%
	35%	71%	64%	25%	11%	-19%	-23%	-59%
	30%	66%	59%	13%	-4%	-39%	-43%	-85%
使用率					80%			
折扣后的租赁利润率	80%	89%	86%	71%	65%	54%	52%	38%
	60%	85%	82%	61%	54%	38%	36%	18%
	50%	82%	78%	54%	45%	26%	24%	1%
	40%	78%	72%	42%	31%	8%	4%	-23%
	35%	74%	68%	34%	21%	-6%	-9%	-41%
	30%	70%	63%	23%	8%	-23%	-27%	-64%
使用率					90%			
折扣后的租赁利润率	80%	90%	88%	74%	69%	58%	57%	44%
	60%	86%	83%	65%	58%	44%	43%	26%
	50%	84%	80%	58%	50%	33%	31%	11%
	40%	80%	75%	48%	38%	17%	14%	-11%
	35%	77%	72%	40%	29%	5%	2%	-27%
	30%	73%	67%	30%	17%	-11%	-15%	-48%

资料来源：各公司官网（采用2024年9月11日价格数据），西南证券整理

2.3.4 回本周期测算

- **回本周期敏感性分析：**基于各公司官方披露价格，数据中心使用率越高，回本周期越短，即使在70%的使用率下，各厂商均可在3年内回本。

算力租赁回本周期的使用率敏感性分析

云厂商	租赁价格($/GPU/h)	使用率	实际收入($/GPU/h)	能耗+托管成本($/GPU/h)	利润($/GPU/h)	经营利润率	硬件资本成本($/台)	回本周期（月）
亚马逊AWS	12.25	70%	8.58	0.23	8.34	97%	269,010	6
甲骨文OCI	10.00		7.00		6.77	97%		7
CoreWeave	4.76		3.33		3.10	93%		15
Runpod	3.99		2.79		2.56	92%		18
Lambda	2.99		2.09		1.86	89%		25
FluidStack	2.89		2.02		1.79	89%		26
Paperspace	2.24		1.57		1.34	85%		35
亚马逊AWS	12.25	80%	9.80	0.24	9.56	98%	269,010	5
甲骨文OCI	10.00		8.00		7.76	97%		6
CoreWeave	4.76		3.81		3.56	94%		13
Runpod	3.99		3.19		2.95	92%		16
Lambda	2.99		2.39		2.15	90%		22
FluidStack	2.89		2.31		2.07	89%		23
Paperspace	2.24		1.79		1.55	86%		30
亚马逊AWS	12.25	90%	11.03	0.26	10.77	98%	269,010	4
甲骨文OCI	10.00		9.00		8.74	97%		5
CoreWeave	4.76		4.28		4.03	94%		12
Runpod	3.99		3.59		3.33	93%		14
Lambda	2.99		2.69		2.43	90%		19
FluidStack	2.89		2.60		2.34	90%		20
Paperspace	2.24		2.02		1.76	87%		27

资料来源：各公司官网（采用2024年9月11日价格数据），西南证券整理

2.3.4 回本周期测算

- **回本周期敏感性分析：**基于80%的使用率，若H100的租赁价格能够给予八折或五折优惠，AWS和OCI均可在1年内回本；三折优惠下，AWS和OCI可在2年内回本。

算力租赁回本周期的价格折扣敏感性分析

云厂商	80%使用率收益($/GPU/h)	折扣（支付原价的百分比）	实际收入($/GPU/h)	能耗+托管成本($/GPU/h)	利润($/GPU/h)	经营利润率	硬件资本成本($/台)	回本周期（月）
亚马逊AWS	9.80		7.84		7.60	97%		6
甲骨文OCI	8.00		6.40		6.16	96%		8
CoreWeave	3.81		3.05		2.80	92%		17
Runpod	3.19	*80%*	2.55	0.24	2.31	90%	269,010	20
Lambda	2.39		1.91		1.67	87%		28
FluidStack	2.31		1.85		1.61	87%		29
Paperspace	1.79		1.43		1.19	83%		39
亚马逊AWS	9.80		4.90		4.66	95%		10
甲骨文OCI	8.00		4.00		3.76	94%		12
CoreWeave	3.81		1.90		1.66	87%		28
Runpod	3.19	*50%*	1.60	0.24	1.35	85%	269,010	35
Lambda	2.39		1.20		0.95	80%		49
FluidStack	2.31		1.16		0.91	79%		51
Paperspace	1.79		0.90		0.65	73%		72
亚马逊AWS	9.80		2.94		2.70	92%		17
甲骨文OCI	8.00		2.40		2.16	90%		22
CoreWeave	3.81		1.14		0.90	79%		52
Runpod	3.19	*30%*	0.96	0.24	0.71	75%	269,010	65
Lambda	2.39		0.72		0.47	66%		99
FluidStack	2.31		0.69		0.45	65%		104
Paperspace	1.79		0.54		0.29	55%		159

资料来源：各公司官网（采用2024年9月11日价格数据），西南证券整理

相关标的：英伟达(NVDA.O)

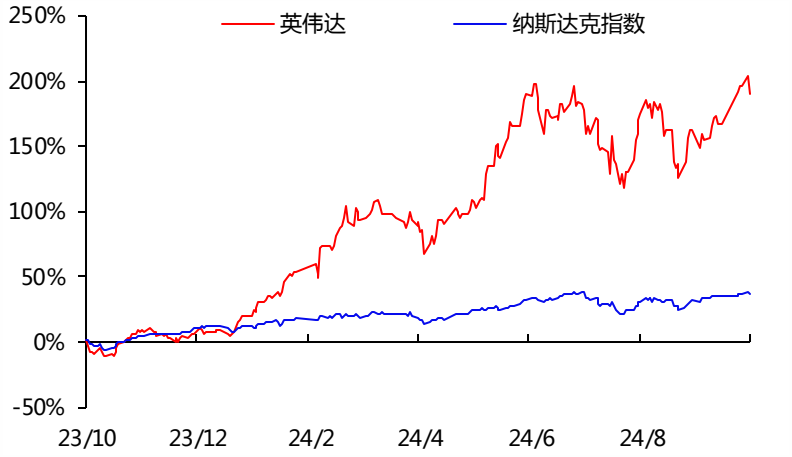
- **投资逻辑：智算引领者，AI芯片生态构筑宽广护城河。**
- ✓ 大模型和生成式AI带动数据中心AI芯片需求持续增长，公司在AI GPU领域具备领先地位，英伟达数据中心GPU市占率超80%；
- ✓ CUDA软件生态上具备高壁垒，公司的超强算力&通信性能+CUDA生态+DGX Cloud AI云服务构筑公司宽广的护城河；
- ✓ 不断推出的新品和技术创新带来的持续竞争力。
- **业绩预测与投资建议：**预计公司未来三年GAAP和Non-GAAP净利润年复合增速分别为62.6%和59.2%，对应PE分别为48倍、32倍和25倍，建议积极关注。
- **风险提示：**AI进展或不及预期；数据中心业务发展或不及预期；竞争加剧等风险。

业绩预测和估值指标

指标	FY2024A	FY2025E	FY2026E	FY2027E
营业收入（百万美元）	60922.00	132735.09	193612.34	246865.19
营业收入增长率	125.85%	117.88%	45.86%	27.50%
GAAP净利润（百万美元）	29760.00	67113.23	99539.61	127931.43
GAAP净利润增长率	581.32%	125.51%	48.32%	28.52%
EPS（美元）	12.10	27.28	40.47	52.01
P/E（GAAP）	81.2	48.1	32.4	25.2

数据来源：公司公告、西南证券（PE截至2024年10月15日）

相对指数表现



数据来源：Wind、西南证券整理

相关标的：微软(MSFT.O)

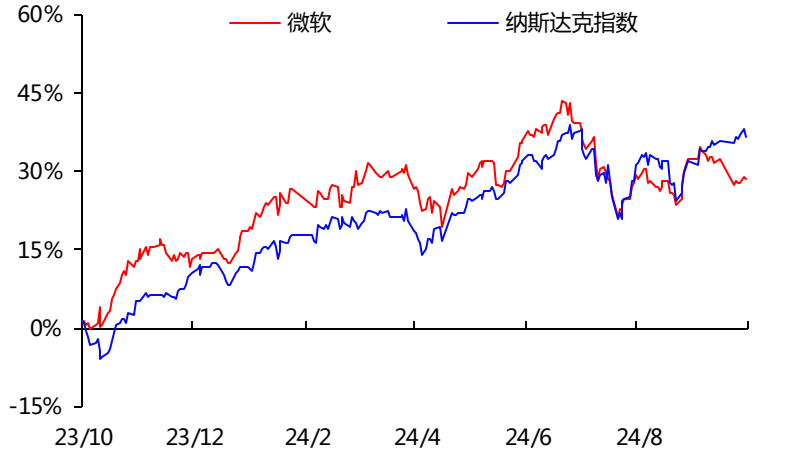
- ❑ **投资逻辑**：云服务领先提供商，**产品线AI化有力拓宽公司护城河。**
- ✓ 受益于企业上云需求增长、以及公司在云业务方面的AI服务供应的增加，云计算业务中的AI贡献有望持续提升，通过多样化AI云服务以更好地满足客户需求。
- ✓ 公司致力于将AI整合至Microsoft 365/Dynamics 365等生产力与流程业务、Bing等搜索业务、以及Windows个人计算业务等，产品线AI化将有力拓宽公司的护城河。
- ❑ **业绩预测与投资建议：**预计公司2025-2027财年归母净利润分别为1027.5亿美元、1198.2亿美元、1380.4亿美元，未来三年复合增速为16.1%。公司为云服务领先提供商，各项业务持续健康发展，前瞻布局AI领域并持续投入，AI商业化加速，订单量有望持续增长，建议积极关注。
- ❑ **风险提示：**AI商业化不及预期、云计算市场竞争加剧、PC市场回暖不及预期等风险。

业绩预测和估值指标

指标	FY2024A	FY2025E	FY2026E	FY2027E
营业收入（百万美元）	245122.00	283481.10	325657.86	370475.13
营业收入增长率	15.67%	15.65%	14.88%	13.76%
GAAP净利润（百万美元）	88136.00	102746.31	119819.61	138038.22
GAAP净利润增长率	21.80%	16.58%	16.62%	15.21%
EPS（美元）	11.86	13.82	16.12	18.57
P/E（GAAP）	35.3	30.3	26.0	22.5

数据来源：公司公告、西南证券（PE截至2024年10月15日）

相对指数表现



数据来源：Wind、西南证券整理

风险提示

- 市场需求不及预期；
- 行业竞争加剧；
- 投资回报不及预期等风险。

西南证券
SOUTHWEST SECURITIES

分析师：王湘杰
执业证号： S1250521120002
电话：0755-26671517
邮箱：wxj@swsc.com.cn

联系人：尤品柯
邮箱 ypk@swsc.com.cn

西南证券 SOUTHWEST SECURITIES

西南证券研究发展中心

西南证券投资评级说明

报告中投资建议所涉及的评级分为公司评级和行业评级（另有说明的除外）。评级标准为报告发布日后6个月内的相对市场表现，即：以报告发布日后6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。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；香港市场以恒生指数为基准；美国市场以纳斯达克综合指数或标普500指数为基准。

公司评级	买入：未来6个月内，个股相对同期相关证券市场代表性指数涨幅在20%以上
	持有：未来6个月内，个股相对同期相关证券市场代表性指数涨幅介于10%与20%之间
	中性：未来6个月内，个股相对同期相关证券市场代表性指数涨幅介于-10%与10%之间
	回避：未来6个月内，个股相对同期相关证券市场代表性指数涨幅介于-20%与-10%之间
	卖出：未来6个月内，个股相对同期相关证券市场代表性指数涨幅在-20%以下
行业评级	强于大市：未来6个月内，行业整体回报高于同期相关证券市场代表性指数5%以上
	跟随大市：未来6个月内，行业整体回报介于同期相关证券市场代表性指数-5%与5%之间
	弱于大市：未来6个月内，行业整体回报低于同期相关证券市场代表性指数-5%以下

分析师承诺

报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，报告所采用的数据均来自合法合规渠道，分析逻辑基于分析师的职业理解，通过合理判断得出结论，独立、客观地出具本报告。分析师承诺不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接获取任何形式的补偿。

重要声明

西南证券 SOUTHWEST SECURITIES

西南证券研究发展中心

西南证券研究发展中心

上海

地址：上海市浦东新区陆家嘴21世纪大厦10楼

邮编：200120

北京

地址：北京市西城区金融大街35号国际企业大厦A座8楼

邮编：100033

深圳

地址：深圳市福田区益田路6001号太平金融大厦22楼

邮编：518038

重庆

地址：重庆市江北区金沙门路32号西南证券总部大楼21楼

邮编：400025

西南证券机构销售团队

区域	姓名	职务	手机	邮箱	姓名	职务	手机	邮箱
上海	蒋诗烽	总经理助理/销售总监	18621310081	jsf@swsc.com.cn	欧若诗	销售经理	18223769969	ors@swsc.com.cn
	崔露文	销售副总监	15642960315	clw@swsc.com.cn	李嘉隆	销售经理	15800507223	ljlong@swsc.com.cn
	李煜	高级销售经理	18801732511	yfliyu@swsc.com.cn	龚怡芸	销售经理	13524211935	gongyy@swsc.com.cn
	田婧雯	高级销售经理	18817337408	tjw@swsc.com.cn	孙启迪	销售经理	19946297109	sqdi@swsc.com.cn
	张玉梅	销售经理	18957157330	zymyf@swsc.com.cn	蒋宇洁	销售经理	15905851569	jyj@swsc.com.c
	魏晓阳	销售经理	15026480118	wxyang@swsc.com.cn				
北京	李杨	销售总监	18601139362	yfly@swsc.com.cn	王一菲	高级销售经理	18040060359	wyf@swsc.com.cn
	张岚	销售副总监	18601241803	zhanglan@swsc.com.cn	王宇飞	高级销售经理	18500981866	wangyuf@swsc.com
	杨薇	资深销售经理	15652285702	yangwei@swsc.com.cn	路漫天	销售经理	18610741553	lmtyf@swsc.com.cn
	姚航	高级销售经理	15652026677	yhang@swsc.com.cn	马冰竹	销售经理	13126590325	mbz@swsc.com.cn
	张鑫	高级销售经理	15981953220	zhxin@swsc.com.cn				
广深	郑龑	广深销售负责人	18825189744	zhengyan@swsc.com.cn	张文锋	销售经理	13642639789	zwf@swsc.com.cn
	杨新意	广深销售联席负责人	17628609919	yxy@swsc.com.cn	陈紫琳	销售经理	13266723634	chzlyf@swsc.com.cn
	龚之涵	高级销售经理	15808001926	gongzh@swsc.com.cn	陈韵然	销售经理	18208801355	cyryf@swsc.com.cn
	丁凡	销售经理	15559989681	dingfyf@swsc.com.cn	林哲睿	销售经理	15602268757	lzr@swsc.com.cn