

区块链

Web3 视角下的 AIGC 算力进化论

伴随着 ChatGPT 的爆红，AIGC（人工智能生成内容）产业链受到广泛关注。我们发现，同为算力，其与加密资产挖矿产业存在相似之处，均为投入算力和电力，获得经济回报，主要的不同点在产品端：AIGC 产业链的产品是对用户提问而回应的内容，而加密资产挖矿的产品是加密资产。

同为算力“吞金兽”，短期比特币更耗电、长期 AI 算力增速快。同为算力生意，我们做了个有趣的研究，AIGC 与 BTC 挖矿耗电量如何？按当前比特币全网算力 319EH/S，每日耗电量约为 2.2 亿千瓦时；而 OpenAI 在训练、推理端的日均耗电量约 2.6 万千瓦时，约是前者的万分之一，但按 OpenAI 的预测，AI 训练所需的算力每 3-4 个月翻倍，且考虑商业竞争算力增长速度预计将更快，“算力即权力”的时代将来临。中性预期下，ChatGPT 日活稳定在 1 亿人次左右、ChatGPT 每 6 个月模型参数翻倍、国际市场将出现 10 个左右类似于 ChatGPT 的商用大模型，而比特币挖矿耗电量保持当前状态，则大约 4 年后，AIGC 大模型耗电量将超过比特币挖矿耗电量，乐观和悲观预期下，这一数字则分别为 1.6 年和 7.5 年。

算力的竞争性——产业自驱的结果。市场忽略了算力的竞争性，考虑商业因素，算力生意存在分子与分母端。对比特币挖矿而言，关于个体矿工能挖到的可用于变现的比特币数量，分子是个体的算力，分母是全网算力，个体矿工能挖得的比特币数量取决于其自身算力占全网算力的比例。而 AIGC 而言，关于个人内容生产者能获得的可用于变现的用户注意力，分子是个体算力驱动下的内容生产能力，分母是全网的内容膨胀速度，从 UGC 到 AIGC 的升维中，谁的内容生产力提升更快，谁就能获得更多商业利益，因此产业自驱之下，市场将追求更高的算力、更优的模型算法、更高功耗比的网络架构以及更便宜的电力。最终算力的需求规模将由应用定义。

寻找算力进化中的边际变化。当下难以预知未来多模态（图片、视频等）中到底会消耗多少算力，可当我们看到海外已然层出不穷的应用时，算力的增长只是时间问题，更重要的边际变化在于算力进化中在芯片、网络、连接等诸多领域有哪些创新方案？我们认为，在 AIGC 的推动下，更有利于新技术、新架构、新材料的落地应用，例如光连接中的 CPO（光电共封装）、MPO（多纤连接器）；芯片层面的 Chipllet；网络架构层面的边缘计算等。

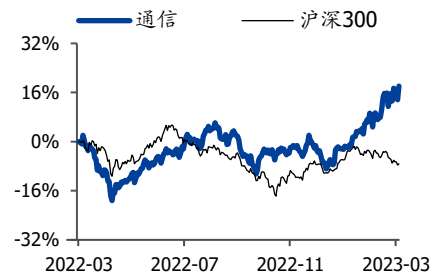
AIGC 驱动算力产业版图生变。我们研究了过往加密资产领域的算力变迁，云服务与分布式早已应用（BTC 云挖矿与 Filecoin 分布式存储），而今在 AIGC 的算力版图亦同步发生。1）英伟达募资百亿进入 AI 云算力，其在 GPU 供给、算力复用度上占优，且将改变其商业模式；2）高通在 AIoT 领域布局终端算力，推理与内容生产端算力与存储的梯度分布（云—雾—端）。这些变化将驱动终端 IT 服务和硬件的升级。

投资建议：建议关注：1）算力侧：英伟达、微软、寒武纪、天孚、太辰光、锐捷网络、中兴、紫光、美格智能、新易盛、中际旭创、Chipllet 产业链等；2）应用及 IP：科大讯飞、汤姆猫、万兴科技、中文在线、昆仑万维、视觉中国、值得买等。

风险提示：AIGC 技术发展不及预期，AIGC 监管趋严。

增持（维持）

行业走势



作者

分析师 宋嘉吉

执业证书编号：S0680519010002

邮箱：songjiaji@gszq.com

分析师 孙爽

执业证书编号：S0680521050001

邮箱：sunshuang@gszq.com

相关研究

- 1、《通信：新版 Office 发布，AI 应用侧边际变化何在》2023-03-17
- 2、《通信：GPT-4 发布：多模态、强智能、大算力》2023-03-15
- 3、《区块链：数字人——GPT4 下的杀手级应用》2023-03-13

内容目录

1 如何理解 AIGC 的算力生意？	3
1.1 算力经济模型：投入硬件和电力，获得经济回报	3
1.2 对用户的内容生成，是 AIGC 产业链的产品	4
2 算力“吞金兽”：当前 BTC 耗电更大，而 AIGC 增速快	4
2.1 AIGC 和比特币挖矿耗电量比较：当前比特币挖矿更耗电	4
2.2 未来，AIGC 耗电量或将超越比特币挖矿	7
3 算力的竞争性——不容忽视的产业自驱结果	7
3.1 比特币挖矿的内化竞争	7
3.1.1 “产量”恒定，全网算力越多，自身获益越少	7
3.1.2 内部过度竞争的出现有前提	8
3.2 AIGC 算力的外化竞争	9
3.2.1 需求端：刚被点燃，放量在即	9
3.2.2 供给端：算力和数据快速进化，模型获开源赋能	10
4 AIGC 算力进化的边际变化	14
4.1 提升超算通信效率，高密度光连接成首选	14
4.2 基于功耗考虑，CPO 方案渗透率有望逐步提升	14
5 AIGC 驱动算力产业版图生变	15
5.1 云化：算力服务商与云厂商走向历史性合作	15
5.2 推理边缘化：边缘算力有望成为推理主体	16
投资建议	17
风险提示	18

图表目录

图表 1: AIGC 与加密资产挖矿产业链	3
图表 2: ChatGPT 每日电费估计	5
图表 3: Stable Diffusion 算力成本	6
图表 4: ChatGPT 类 AIGC 模型耗电量或将超越比特币挖矿	7
图表 5: AIGC 与比特币挖矿各自内部过度竞争的可能性	8
图表 6: 比特币全网算力与单位算力的挖矿利润（截至 2023 年 2 月 23 日）	8
图表 7: 比特币价格与单位算力的挖矿利润（截至 2023 年 2 月 23 日）	9
图表 8: 英伟达部分显卡的算力增长	10
图表 9: 2012-2018 年 AI 训练算力增长速度	10
图表 10: 合成数据发展预测	11
图表 11: 不同类型数据的消耗趋势和耗尽日期	12
图表 12: GPT-3 训练的自然语言数据	12
图表 13: AIGC 驱动全网内容大爆炸，个人唯有能生成更多内容，方能获得更多用户注意力	13
图表 14: 大语言模型参数数量的增长趋势	13
图表 15: 国外主要 AIGC 预训练模型一览表	13
图表 16: 嘉楠科技边缘计算芯片 K230	16
图表 17: AIGC 投资标的	17

1 如何理解 AIGC 的算力生意？

伴随着 ChatGPT 的爆红，AIGC（人工智能生成内容）产业链受到市场广泛关注。算力，作为 AIGC 的基础，更被称为基石，近日英伟达在财报中的表述也佐证了这一点。市场一直在问：未来的算力需求到底有多大？我们发现现有的论述虽就 GhatGPT 的需求进行了推演，但未来 AIGC 必然走向多模态，其需求量和需求逻辑也将有所差异，算力的增长将大幅超预期，而理解这些的一切都建立在如何认识 AI 算力生意上。

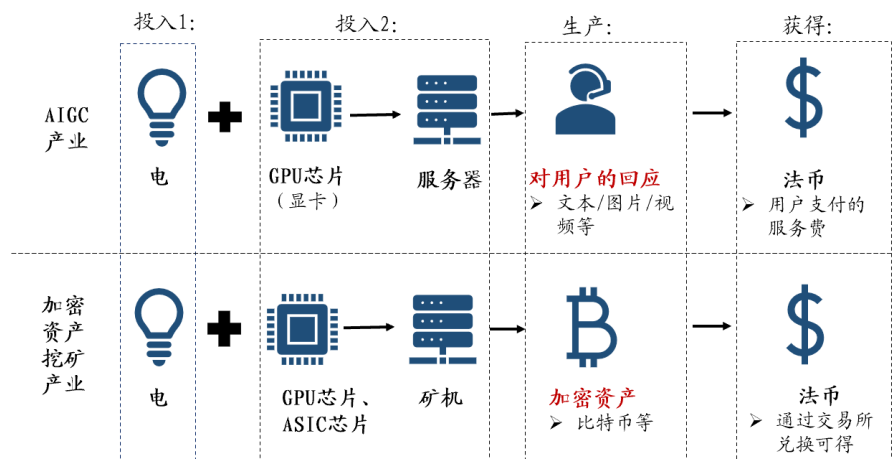
过去我们讨论算力较多着眼于 IDC，其商业逻辑着重于“数字地产”，即购买或租赁土地，投资建设机房并按机柜（或功耗）租赁给下游政企客户，从用途上兼顾“存”与“算”，但大部分一线城市以外的机柜都以“存”为主。AI 算力则聚焦于“算”，在我们和国内外团队交流中，构建 AI 超算中心的资本开支核心在于 GPU，这种高密度的计算未来将着重受限于两个指标：**算力、功耗比**。这与过往的“挖矿”市场在经济模型上有一定相似之处，研究这一特征有助于预测 AIGC 算力生意的未来走向。

1.1 算力经济模型：投入硬件和电力，获得经济回报

过去我们看到很多超算中心提供气象、交通、工业等领域的服务，但更多侧重于社会价值，而对 AIGC 或“挖矿”而言，都直接以经济回报为目的，投入产出要求更高。产业链方面，可以看出，AIGC 产业链与加密资产挖矿存在若干相似之处。

- 成本主要是电费和算力投资，并且这两项费用主要取决于芯片本身的性能。其中，电费取决于芯片的功耗，而算力费用取决于芯片的算力水平和云计算费用。
- 目的都是为了获得经济回报。从 OpenAI 最近的合作及收费方式可以清晰看到其商业飞轮模型，正因为有了这样的反馈才能推动其快速发展。

图表 1: AIGC 与加密资产挖矿产业链



资料来源：国盛证券研究所整理

1.2 对用户的内容生成，是 AIGC 产业链的产品

AIGC 产业链与加密资产挖矿的不同之处在产品端，AIGC 产业链的产品是对用户提问而回应的内容，而加密资产挖矿的产品是加密资产。具体看来，两大产业的不同包括：

- 成本端，
 - **AIGC**：使用的芯片主要是 GPU（图形处理器/显卡）；**硬件之外，相较于加密资产挖矿，AIGC 投入了更大规模、更昂贵的人力资源。**例如，澎湃新闻称，为 ChatGPT 项目做出贡献的人员不足百人（共 87 人）。根据 O'Reilly 2021 年 6 月的调查，OpenAI 所在的加州数据和人工智能岗位的平均年薪为 17.6 万美元，假设 OpenAI 员工的平均年薪也与此相当，则 OpenAI 一年需付出 1531 万美元（约合 1 亿人民币）。
 - **加密资产挖矿**：使用的是 ASIC（专用集成电路）和 GPU（当前 GPU 因以太坊共识机制改变已经退出）。相较于 GPU，ASIC 芯片的针对性更强，只适用于各加密资产各自的算法。加密资产挖矿行业所需的专业人才主要是芯片设计人才和矿机运维人才，根据我们的了解，相较于前者，后者工资较低，而前者规模相较于重人力资本投入的 AIGC 产业明显更少，原因主要是**加密资产挖矿（例如比特币矿机）芯片设计不需要类似于 AIGC 模型一般持续、大规模的人工调试和算法优化。**
- 产品端，
 - **AIGC**：产品是对用户提问的回答，形式包括文本、图片和视频等多模态输出，当前主要是前两种。
 - **加密资产挖矿**：产品是各加密资产系统为奖励矿工维护系统账本安全性，而发放的加密资产奖励，以及账本内交易中给矿工的手续费，例如比特币。
- 收入端，
 - **AIGC**：会预先从客户处收取使用费，例如，ChatGPT 高级版（Plus）的订阅费为每月 20 美元；同时对下游 B 端客户开放的 API 接口收费。
 - **加密资产挖矿**：需要将生产所得的加密资产，通过海外加密资产交易所等渠道兑现。

2 算力“吞金兽”：当前 BTC 耗电更大，而 AIGC 增速快

2.1 AIGC 和比特币挖矿耗电量比较：当前比特币挖矿更耗电

由于 AIGC 算法和加密资产挖矿算法的不同，难以直接比较其参数，另外由于 AIGC 尚处于发展早期，年收入尚难估算，但我们找到了直接对比 AIGC 和比特币挖矿的一个直观指标：耗电量。

根据我们的计算，ChatGPT 每日训练和推理的耗电量为 25677 千瓦时，是比特币挖矿日耗电量的万分之一（比特币挖矿为 2.2 亿千瓦时）。

此处我们使用的数据是：

- 比特币挖矿方面，
 - 已知项：2023 年 2 月 28 日，比特币全网算力为 319（Ehash/s）。
 - 假设项：全网矿机一半为蚂蚁矿机 S9（算力为 14T hash/s，功耗为 1400W），一半为蚂蚁矿机 S19（算力为 141T hash/s，功耗为 3031.5W），其平均算力为 78T hash/s，平均功耗为 2216W。可以算出，按照前述全网算力，全网大约有 412 万台矿机。

- AIGC 方面:
- 已知项: 1) 英伟达单个 DGX A100 服务器内含 8 个 A100 GPU; 2) 单个英伟达 A100 GPU 的峰值浮点运算次数是 624 TFLOPS (稀疏化后, 指每秒可做 10^{12} 次浮点运算); 3) 单个英伟达 DGX A100 服务器的最大功率是 6.5 千瓦; 4) ChatGPT 前身——GPT3 的模型参数量是 1750 亿个、训练的标识符 (token) 数是 3000 亿个。
 - 假设项: 1) ChatGPT 模型参数是 3000 亿个, 推理时可以通过蒸馏等技术仅使用 300 亿个参数; 2) 单个模型参数、单个标识符训练时所需的浮点运算次数为 6N, 推理时为 2N, 其中 N 为模型参数, 单位为 FLOPs; 3) 单个服务器训练时的峰值利用率为 46.2%, 推理时为 21.3%, 与 GPT3 保持一致; 4) 训练用标识符数量为 5000 亿个; 5) 关于推理时用的标识符数量, 假设根据 Similarweb 统计, 截至 2023 年 2 月 21 日的 28 日 chat.openai.com 的每日访问人数均值为 28696429, 此处简化为 3000 万人, 平均每人提出 10 个问题, 每个提问含 50 个单词, ChatGPT 为单个提问生成 5 个回答, 每个回答包含 100 个单词, 每 1000 个单词对应 750 个标识符; 6) 假设 ChatGPT 每个月训练十次, 每个月有三十天。可以算出, ChatGPT 每次训练需要 1506 个英伟达 DGX A100 服务器, 每日推理需要 144 个, 相当于 13194 个 A100 GPU。

图表 2: ChatGPT 每日电费估计

$$\begin{aligned}
 & \text{ChatGPT耗电量} = \left[\text{训练所需服务器数} + \text{推理所需服务器数} \right] \times \text{单个服务器的最大功率 (6.5kw)} \\
 & \text{其中, 训练/推理所需服务器数} \\
 & = \frac{\text{训练/推理所需的浮点运算次数 (FLOPs)}}{\text{单台服务器的可用浮点运算次数 (FLOPS)}} \\
 & = \left[\frac{\text{单个模型参数、单个Token需要的FLOP次数} \times \text{模型参数} \times \text{每秒需训练/推理的token数量}}{\text{单个服务器包含的GPU数 (8)} \times \text{单个GPU的峰值FLOP次数 (624)} \times \text{FLOPS峰值利用率}} \right] \\
 & \text{每秒需推理的token数量} = \frac{\text{每日提问的人数} \times \text{平均每人提问数} \times \text{每个问题与其对应回答的总单词数} \times \text{生成单个单词包含的token数量 (1000/750)}}{\text{每日包含的秒数 (24*60*60)}}
 \end{aligned}$$

资料来源: Substack 作者“Sunyan”, 国盛证券研究所

注: 浮点 (floating-point) 指的是带有小数的数值, 浮点运算即是小数的四则运算; FLOPS (floating point operations per second), 指每秒浮点运算次数, 指计算速度, 用于衡量硬件性能; FLOPs (floating point operations), 指浮点运算数, 指计算量, 用于衡量算法/模型的复杂度。

ChatGPT 之外, 我们也注意到其他 AIGC 文本服务和多模态领域对算力和电量的海量需求。

大语言模型训练成本高昂。参考 Substack 作者“Sunyan”的测算, 模型参数为 2800 亿个 Gopher (Google DeepMind) 的训练成本为 200 万美元; 模型参数为 5300 亿的 MT-NLG (Microsoft/Nvidia) 的训练成本为 400 万美元; 模型参数为 5400 亿的 PaLM (Google Research) 的训练成本为 1100 万美元。根据《财富》杂志, 在尚未收费的 2022 年, OpenAI 该年度净亏损即高达 5.445 亿美元 (不含员工股票期权), 我们预计其中算力成

本占据了重要位置。

■ 图片

生成图片所需的算力取决于多个因素，例如生成图片的分辨率、所使用的算法、训练数据集的大小和质量，以及使用的硬件等。

➤ Stable Diffusion

- 根据 Stability AI 的创始人兼首席执行官 Emad Mostaque，Stable Diffusion 使用了 256 张 Nvidia A100 训练，所有显卡总计耗时 15 万小时，成本约为 60 万美元。
- 根据 Business Insider 2022 年 10 月的报道，Stable AI 的运营和云计算成本超过 5000 万美元。当月，Stability AI 完成了 1 亿美元融资，投后估值 10 亿美元，为 AI 训练助力。

图表 3: Stable Diffusion 算力成本



资料来源: Stable Diffusion CEO Emad Mostaque, 国盛证券研究所

➤ Dall-E

- 根据《Zero-Shot Text-to-Image Generation》，OpenAI 在 Dall-E 的整个训练中，共使用了 1024 块 16GB 的 NVIDIA V100 GPU。团队从网上收集了一个包含 2.5 亿个图像文本对的数据集，在这一数据集上训练一个包含 120 亿个参数的自回归 Transformer。
- 另外，论文介绍道，用于图像重建部分的 dVAE 模型的训练，共用了 64 块 16GB 的 NVIDIA V100 GPU，判别模型 CLIP 则共使用 256 块 GPU 训练了 14 天。

■ 视频

生成视频所需的算力和存储空间比生成单张图片更高，因为生成视频需要在时间维度上连续地生成，而视频中的每一帧都需要根据前一帧生成。

因此，生成视频所需的算力取决于多个因素，包括视频分辨率、帧率、视频长度、所使用的算法和训练数据集的大小和质量。使用 GPU 可以显著加速训练，但是在生成高分辨率视频时，可能需要使用多个 GPU 和分布式计算来加速训练。

近期我们尝试定制了 2D 仿真数字人，在每次内容生产时就是将语音和文本内容添加到

已经构建好的人物模型中（数字人模型），在后台通过 GPU 生成短视频，其生成速度与 GPU 规模密切相关。对于创业企业而言，未来的定价就是基于算力成本，用户也是购买算力时长后自由使用。

2.2 未来，AIGC 耗电量或将超越比特币挖矿

这个问题在近期的交流中常被问及，海外有一系列的论文对此算力进行了测算，但我们认为更重要的各类应用涌现、商业竞争加剧带来的算力“军备竞赛”是非线性的，算力与应用之间会形成正反馈。

考虑到 ChatGPT 的日活还在不断增长、国际 AIGC 服务的竞争还在不断加深，我们预计，不考虑模型效率提升对算力需求的下降，也不考虑模型训练的标识符（token）的增长，**中性预期下，ChatGPT 日活稳定在 1 亿人次左右、ChatGPT 每 3 个月模型参数翻倍、同时假设全球市场将出现 10 个左右类似于 ChatGPT 的商用大模型，而比特币挖矿耗电量保持当前状态，则大约 4 年后，AIGC 大模型耗电量将超过比特币挖矿耗电量。乐观和悲观预期下，这一数字则分别为 1.6 年和 7.5 年。**注意，此时我们只是考虑了类似于 ChatGPT 的商用大模型，而实际可能出现可能消耗更多算力的多模态商用大模型，如果考虑到这一点，AIGC 耗电量的增长将会更快。

图表 4: ChatGPT 类 AIGC 模型耗电量或将超越比特币挖矿

		悲观预期	中性预期	乐观预期
假设项	ChatGPT 训练模型参数翻倍需要的月份数 (个)	9	6	3
	ChatGPT 日活 (亿人)	0.5	1	2
	ChatGPT 类大模型数量 (个)	5	10	15
结果项	ChatGPT 类大模型每日训练与推理耗电量超越当前比特币挖矿每日耗电量所需的时间 (月份数)	90	48	19.5
	ChatGPT 类大模型每日训练与推理耗电量超越当前比特币挖矿每日耗电量所需的时间 (年数)	7.5	4	1.6

资料来源: OpenAI, 国盛证券研究所 注: 此处比特币挖矿耗电量指 2023 年 2 月 28 日耗电量, 计算假设见“2.1”部分内容。

3 算力的竞争性——不容忽视的产业自驱结果

很多时候，算力的投入并不是独立的，而需要和全网总量竞争。我们认为，AIGC 产业和加密资产挖矿中的比特币挖矿，都存在内部过度竞争的可能性。对于比特币而言，算力竞赛的目的是要占据更多的全网算力比重，从而提高获得奖励的可能性；对于 AIGC 而言，在内容创作领域，需要保证自身的内容生产速度快于行业平均，尤其是在短视频、游戏等相对重度的领域中。

3.1 比特币挖矿的内化竞争

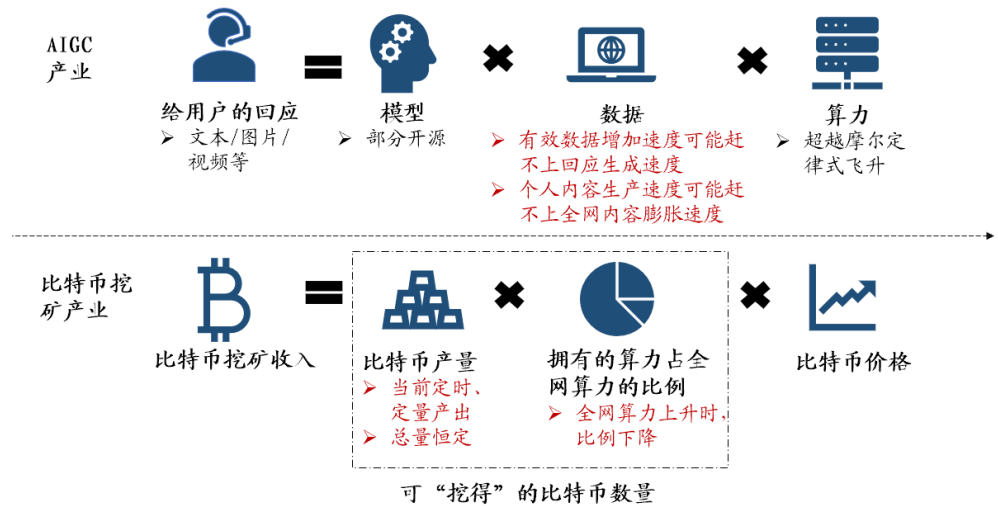
3.1.1 “产量”恒定，全网算力越多，自身获益越少

比特币全网每 10 分钟出一个“区块”，成功“挖掘”出该“区块”的矿工将获得系统发

行的比特币奖励和该区块内所有交易的手续费（前者占主要部分），当前的出块奖励是6.25个比特币（起初是50个，奖励数每四年减半，即“50→25→12.5→6.25→...”）。

根据这些限制条件，比特币的“产量”总量恒定（约为2100万），在出块奖励未改变的每四年内也保持了相对恒定（如当前每天几乎是900个），而矿工能挖得的比特币数量取决于其自身算力占全网算力的比例。当其他矿工认为当前挖矿性价比比较高而将算力投入挖矿，则单个矿工自身算力占全网算力的比例下降，此时如果比特币价格没有出现明显上升，则该矿工的比特币挖矿收入和利润出现下降，这就是比特币挖矿内部过度竞争的逻辑。

图表5: AIGC与比特币挖矿各自内部过度竞争的可能性

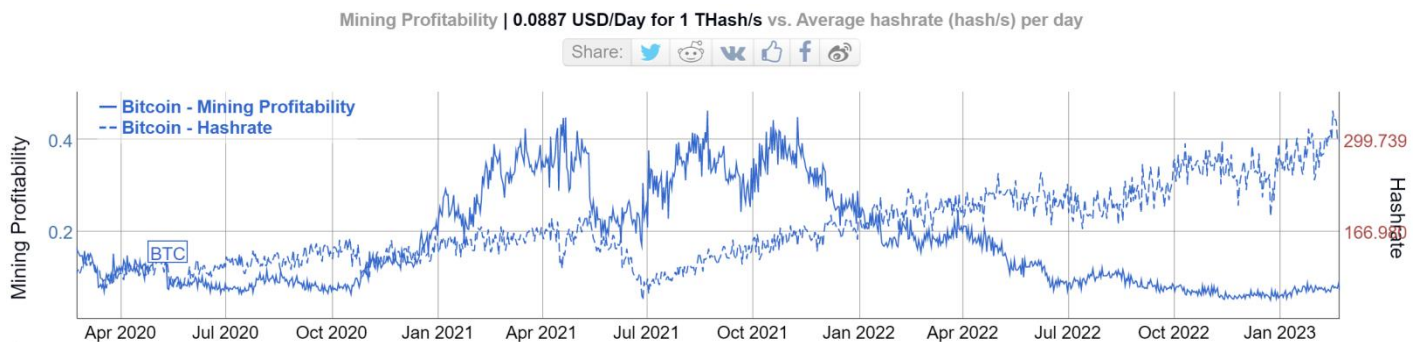


资料来源: 国盛证券研究所整理

3.1.2 内部过度竞争的出现有前提

需要注意到，比特币挖矿出现内部过度竞争存在若干前提，实际情况中，内部过度竞争并非一定会出现。根据 Bitinfocharts 统计，截至 2023 年 2 月 23 日的三年以来，比特币全网算力呈现平稳上升趋势，而与此同时，单位算力的挖矿利润呈现平稳下降趋势，这呈现出一定的“内部过度竞争”特征。一个重要原因是，比特币价格在此期间呈“横盘震荡”状态。

图表6: 比特币全网算力与单位算力的挖矿利润（截至 2023 年 2 月 23 日）



资料来源: Bitinfocharts, 国盛证券研究所

从趋势上看，比特币挖矿单位算力挖矿利润与比特币价格更为相关，当比特币价格上升，前者即能出现较为明显的上升态势。这在一定程度上意味着，打破矿工之间在算力和可挖得比特币“数量”上的内部过度竞争，获得更多比特币挖矿收入和利润，需要比特币“价格”因素发力。

图表 7: 比特币价格与单位算力的挖矿利润 (截至 2023 年 2 月 23 日)



资料来源: Bitinfocharts, 国盛证券研究所

3.2 AIGC 算力的外化竞争

3.2.1 需求端：刚被点燃，放量在即

创新应用往往是供给引导需求。当前，AIGC 正走向多模态，各种服务已经逐步实现商业化，用户需求刚被供给端点燃，从最新的增长情况看，放量在即。

- AIGC 文字服务 ChatGPT
 - 高级版 (Plus) 订阅服务每月费用为 20 美元/账户。
- AIGC 图片服务 Stable Diffusion
 - Stability AI 推出了基于 Stable Diffusion 的图片 AIGC 平台和 API DreamStudio。DreamStudio 会根据用户对图片像素和操作步骤的需求收费，每张图片消耗 0.2-8.2 个积分。新注册用户可以免费获得 200 个积分，此后每 100 积分消耗 1 英镑，默认设置生产 4 张 512*512 的图消耗 0.8 个积分。
 - 此外，公司还对企业客户提供定制化服务。
 - 截至 2022 年 10 月，DreamStudio 拥有超过 150 万用户，创建了超过 1.7 亿张图像，而 Stable Diffusion 全渠道的用户数量每天超过 1000 万。
- AIGC 图片服务 Dall-E
 - 定价是 15 美元生成 115 次。每次生成请求会返回 3~4 张 1024*1204 尺寸的结果图片，每次耗费 0.13 美元。
 - 2022 年 7 月，OpenAI 正式宣布 DALL-E 准备向 100 万个用户开放测试版，第一个月，用户有 50 个免费积分，以后每个月有 15 个免费积分。
 - 根据 OpenAI，截至 2022 年 11 月 3 日，已经有超过 300 万人使用 DALL·E2，每天创建的图像数量达到 400 万张。
- AIGC 图片服务 Midjourney
 - 按月订阅收费，个人用户有每月 10 美元和 30 美元两档，分别对应 200 和 900 次标准生成或放大请求（1 次标准生成请求对应 1GPU 分钟，在 fast 模式下。

每次生成会获得四张 256*256 的初始结果，用户可以选择放大，每次放大也耗费 1GPU 分钟)。

- 超过限额后，每 60GPU 分钟售价 4 美元。但使用这两档订阅服务，用户的提示词和结果图片都将暴露在公共空间，此时再付出 20 美元可享有私密服务通道。
- B 端每年费用为 600 美元。
- 截至 2022 年 11 月，Midjourney 在 Discord 服务器中已累计有超 300 万用户。

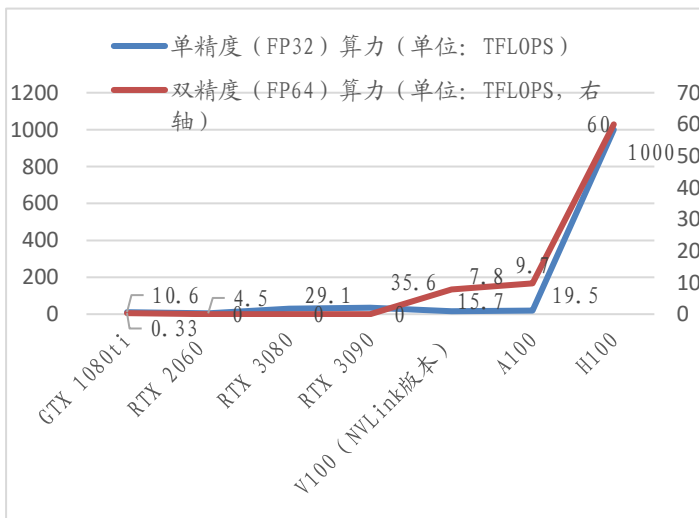
3.2.2 供给端：算力和数据快速进化，模型获开源赋能

我们认为，如果“需求横盘震荡或下滑”前提不存在，则 AIGC 并不会陷入内部过度竞争状态。目前看来，业内公认的 **AIGC 大语言模型 (LLM)** 的三个关键因素都在快速飞升，但一旦 AIGC 内部过度竞争出现，“数据”和“算力”可能最先，“模型”则可能次之，部分模型的开源能为其带来长久活力。

■ 算力：六年来增长了 182 倍

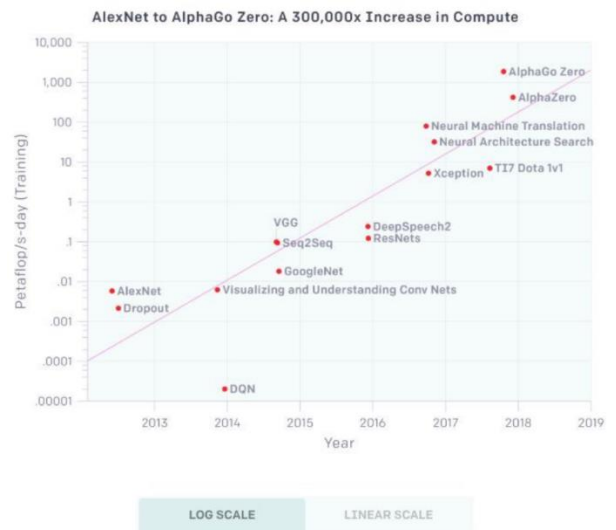
- 摩尔定律中，集成电路上可容纳的元器件的数目，约每隔 18-24 个月便会增加一倍，性能也将提升一倍。
- 继摩尔定律后，英伟达 CEO 黄仁勋提出黄氏定律：每 12 个月 GPU 性能翻一倍，且不受物理制程约束。
- 根据 OpenAI 测算，自 2012 年至 2018 年，用于训练 AI 所需要的算力大约每隔 3-4 个月翻倍，总共增长了 30 万倍（而摩尔定律在相同时间只有 7 倍的增长），每年头部训练模型所需算力增长幅度高达 10 倍，整体呈现指数级上涨。
- 我们统计了部分英伟达有代表性的显卡，发现 2017 年以来，2022 年 4 月面世的 H100 较 2017 年 3 月推出的 GTX 1080ti，在单精度 (FP32) 上，是后者的 94 倍 (1000/10.6)，在 AIGC 常用的双精度 (FP64) 算力上，是后者的 182 倍 (60/0.33)。

图表 8: 英伟达部分显卡的算力增长



资料来源: CSDN, 国盛证券研究所

图表 9: 2012-2018 年 AI 训练算力增长速度



资料来源: OpenAI, 新智元, 国盛证券研究所

■ 数据：当前互联网数据将耗尽，合成数据将成主流

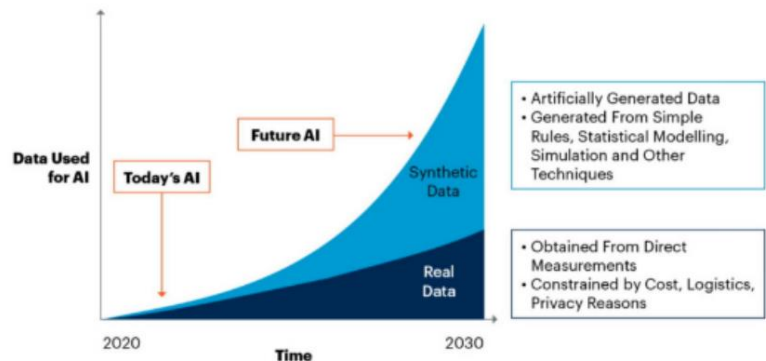
➤ 输入端

合成数据，不基于任何现实现象或事件的数据组成，而通过计算机程序生成。AIGC生成的数据是合成数据生产的重要方式。Gartner 预测：

- ✧ 到 2024 年用于训练 AI 的数据中有 60% 将是合成数据。
 - ✧ 到 2025 年，AIGC 生成的数据将占有所有合成数据的 10%，而目前这一比例还不到 1%。
 - ✧ 到 2030 年 AI 模型使用的绝大部分数据将是人工智能合成的。
- 这一方面体现出市场对合成数据的看好，另外一方面，我们认为，这反应出市场当前互联网公开数据可能存在被 AI 模型“耗尽”的担忧。

图表 10: 合成数据发展预测

By 2030, Synthetic Data Will Completely Overshadow Real Data in AI Models



Source: Gartner
750175_C

Gartner

Synthetic data will become the main form of data used in AI. Source: Gartner, "Maverick Research: Forget About Your Real Data – Synthetic Data Is the Future of AI," Leinar Ramos, Jitendra Subramanyam, 24 June 2021.

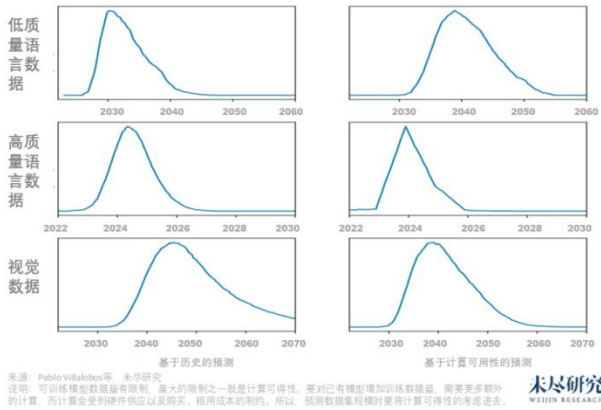
资料来源: Gartner, 国盛证券研究所

根据未尽研究的梳理，阿伯丁大学、麻省理工大学、图宾根大学的 Pablo Villalobos 等 6 位计算机科学家在论文《我们会用完数据吗？机器学习中数据集缩放的局限性分析》中预计：

- ✧ 对于语言模型来说，耗尽当前互联网数据的情况将在 2030 年到 2040 年之间发生；
- ✧ AIGC 文本训练使用的是高质量数据集，它通常包括 50% 的用户生成内容，15%~20% 的书籍，10%~20% 的科学论文，近 10% 的代码和近 10% 的新闻。
- ✧ 假设数字化书籍、公共 GitHub 和科学论文中可用文本的全部数量占高质量数据集的 30% 到 50% 之间，当前高质量语言数据的总存量为 9 万亿（即 9e12，上下限大概为 4.6 万亿到 17 万亿）个单词，每年增长率为 4% 到 5%。以高质量语言数据库作为上限来预测语言数据集的增长，放缓发生得更早，在 2026 年之前。

根据上述论文《我们会用完数据吗？机器学习中数据集缩放的局限性分析》，现在互联网上的存量视觉数据数量在 8.1 万亿到 23 万亿之间，目前的年增长率在 8% 左右，视觉模型数据耗尽的情况将在 2030 年到 2060 年之间发生。相较文本数据，视觉数据对 AI 训练而言，耗尽的时间较为靠后。不过根据 IDEA 研究院计算机视觉与机器人研究中心讲席科学家张磊博士的观点，将视觉数据用于 AI 训练，还存在“大模型学到的全图表征用于细粒度问题会性能递减”“视觉算法方面尚需进一步改进”“视觉大模型目前还没有达到单纯增加数据提高效果的阶段”等问题。

图表 11: 不同类型数据的消耗趋势和耗尽日期



资料来源: Pablo Villalobos, 未尽研究, 国盛证券研究所

图表 12: GPT-3 训练的自然语言数据

数据集	数量: 十亿*	占比	完整跑遍样本次数**
Common Crawl	410	60%	0.44
WebText2	19	22%	2.9
Books1	12	8%	1.9
Books2	55	8%	0.43
Wekipidia	3	3%	3.4

来源: OpenAI, 未尽研究
说明: *指token, 即有意义的标点符号与单词。**指300B大小tokens的训练数据集下需要遍历的样本 (epoch elapsed) 的次数。

未尽研究
WEIJIN RESEARCH

资料来源: OpenAI, 未尽研究, 国盛证券研究所

➤ 输出/产品端

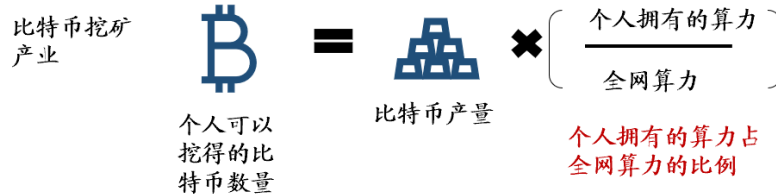
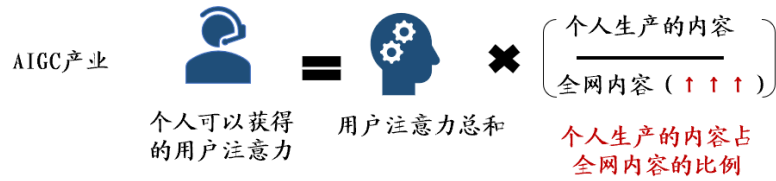
除了“输入”内容/数据的快速进化, 我们还需要考虑的是, 在 AIGC “输出”的内容/数据方面, 当未来 AIGC 的使用越来越普及的时候, 互联网内容竞争的格局会有怎样的变化?

在过去的十年中, 抖音、小红书等平台大幅降低了内容创作、分享的壁垒, UGC 模式的内容丰富程度和规模远超过往。而当前无论是带货直播还是内容创作, 门槛均比之前有所提高, UGC 也“卷”了起来。

而当 AIGC 来临时, 内容竞争又将升维, 例如数字虚拟人将实现包括外形、声纹、对话内容的模拟, 或许看到的很多内容将是 AI 生成, 作为个体需要保证自身的内容生产、进化速度快于行业均值, 背后的算力投入将首当其冲, 否则将直接影响账号盈利能力。

我们认为, 市场忽略了算力的竞争性, 考虑商业因素, 算力生意存在分子与分母端。对比特币挖矿而言, 其分子是个体的算力, 分母是全网算力, 个体矿工能挖得的比特币数量取决于其自身算力占全网算力的比例。而对 AIGC 而言, 其能获得的用户注意力的分子是个体算力驱动下的内容生产能力, 分母是全网的内容膨胀速度。从 UGC 到 AIGC 的升维中, 谁的内容生产力提升更快, 谁就能获得更多商业利益, 因此产业自驱之下, 市场将追求更高的算力、更优的模型算法、更高功耗比的网络架构以及更便宜的电力。

图表 13: AIGC 驱动全网内容大爆炸, 个人唯有能生成更多内容, 方能获得更多用户注意力

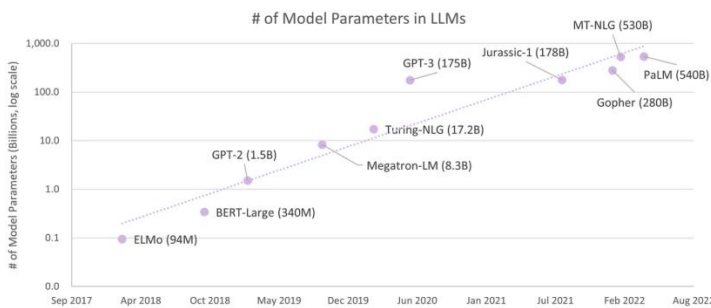


资料来源: 国盛证券研究所整理

■ 模型: 参数迅速堆叠

2017 年以来, AIGC 模型的参数量逐年增加, 从, 2021 年以来, 这一趋势呈现出明显加快的状态, 从 ELMo 的 9400 万个参数一路飙升, 至 2020 年年中, GPT-3 引领行业走入千亿参数时代, 参数达 1750 亿个, 截至 2021 年年中, 英伟达大语言模型 MT-NLG 更是超过了 5300 亿。

图表 14: 大语言模型参数量的增长趋势



资料来源: Substack 作者 “Sunyan”, 国盛证券研究所

图表 15: 国外主要 AIGC 预训练模型一览表

预训练模型	应用	参数量	领域
谷歌	BERT	4810 亿	NLP
	LaMDA		对话系统
	PaLM	5400 亿	NLP
Imagen	语言理解与生成、推理、代码生成	110 亿	多模态
	Parti	200 亿	多模态
微软	Florence	6.4 亿	CV
	Turing-NLG	170 亿	NLP
Facebook	OPT-175B	1750 亿	NLP
	M2M-100	150 亿	NLP
Deep Mind	Gato	12 亿	多模态
	Gopher	2800 亿	NLP
	AlphaCode	414 亿	NLP
Open AI	GPT3	1750 亿	NLP
	CLIP&DALL-E	120 亿	多模态
	Codex	120 亿	NLP
	ChatGPT		NLP
英伟达	Megatron-Turing NLG	5300 亿	NLP
Stability AI	Stable Diffusion		多模态

资料来源: 腾讯研究院, 国盛证券研究所

底层的大模型具有明显的头部效应, 在算力、投资、数据上形成了正反馈, 可以预见, 竞争之下, 大模型会不断集中。而基于大模型开发的垂直应用将层出不穷, 正如我们在上篇报告中所说, 大模型的“通识化”将是未来 AI 快速迭代的基础。

4 AIGC 算力进化的边际变化

当下难以预知未来多模态（图片、视频等）中到底会消耗多少算力，可当我们看到海外已然层出不穷的应用时，算力的增长只是时间问题，更重要的边际变化在于算力进化中在芯片、网络、连接等诸多领域有哪些创新方案？

我们认为，在 AIGC 的推动下，更有利于新技术、新架构、新材料的落地应用，例如光连接中的 CPO（光电共封装）、MPO（多纤连接器）；芯片层面的 Chiplet；网络架构层面的边缘计算等。

4.1 提升超算通信效率，高密度光连接成首选

大规模超算集群，通信效率直接影响训练效率。大模型训练需要超算集群持续运算数月时间，需调用数千片 GPU 运算海量数据，GPU 与 GPU 间、服务器与服务器节点之间存在海量内部数据交互需求，根据木桶效应，只要有一条链路出现负载不均导致网络堵塞，成为木桶短板，则其他链路即使畅通无阻，集合通信时间仍然会大幅度增长，进而直接影响训练效率。集群训练会引入额外的通信开销，从而导致 N 个 GPU 算力达不到单个 GPU 算力的 N 倍。因此，高性能网络互联，是大规模 GPU 集群所必须的。

提升超算通信效率，高密度光连接成为首选。超算除了需要高密度光传输端口外，端口和端口间主要以光纤连接器来实现光互联，多采用 MTP/MPO 高密度光纤连接器实现连接。其中 MPO 是光纤连接器，主要用途是用以实现光纤的接续，MPO 会直接影响光传输系统的可靠性等各项性能。数据中心的内部光学连接需要借助光模块和光纤连接器来实现。因超算集群内部节点之间对于高效率高速互联的硬性需求，光端口密度的进一步提升，并且综合光纤连接器走线连接的可实施性和可维护性，我们预计超算集群对于高密度光纤连接器的需求量较传统云数据中心有较大幅度的提升，而随着全球科技聚焦大模型 AI 训练，相应超算投入将确定性增长，高密度光连接将确定性受益。

4.2 基于功耗考虑，CPO 方案渗透率有望逐步提升

AI 大背景下，未来基于功耗考虑，CPO 方案渗透率有望逐步提升。CPO 方案通过光电耦合共封装在插槽或 PCB 上，加上液冷板降温控制功耗，有望成为 AI 高算力下高能效比方案。但 CPO 在降低功耗的同时也有一些隐患，主要是光电共封装后，光引擎焊接在同一插槽上不可更换。如果光纤或者光引擎出现损坏，可能会影响整个 CPO 交换 ASIC 基板，对交换机生产的总体成本会产生负面影响。

从 CPO 方案看 MPO 的变化。CPO 交换机内部带来的变化一个是光纤数量增多，一个是交换机内部布线复杂度提升。一个 51.2T 的交换机按单口 100G 来估算将达到 512 通道，对应 1024 根光纤。对应如果是 16 芯的 MPO 需要 64 根，对应 64 个端口。

光纤和 MPO 用量提升后，同时又由于 CPO 方案拉近了光引擎和 ASIC 的距离来降低线上损耗，就导致光纤布线要从原来的机箱外部延伸到内部接到光引擎，等于额外增加了光引擎到交换机机箱前面板的布线。内部光纤路由复杂程度提高。

中间板/板载光互连或成布线解决方案，降低 CPO 出错成本。因为 CPO 下每个光引擎到面板距离不同，导致尾纤长度有差异，且布线复杂易损坏光纤影响整机。考虑在光引

擎和端口面板间增加板中连接器，固定尾纤长度，降低布线复杂度。将 CPO 的试错成本转移到板中连接器和端口的 MPO 上。简而言之，通过增加连接来降低布线复杂度和出错的成本。

海外 AIGC/ChatGPT 持续扩散，带宽密度有望大幅上行。在未来算力/带宽高增的背景下，以及 CPO 的新方案拉动下，MPO 的用量有望大幅提升，在设备内部重要性也逐步提高。

5 AIGC 驱动算力产业版图生变

5.1 云化：算力服务商与云厂商走向历史性合作

当算力竞争加剧，“降本增效”将成为企业的生存关键，通过投资、并购，与上、游合作，将成本“内化”是一条可选路径。我们注意到，当前已有 AIGC 企业选择“投靠”云服务厂商，例如 OpenAI 与微软 Azure 云的合作；也有算力厂商开始了对 AI 模型的研发，例如英伟达大语言模型 NeMo 和 NVIDIA BioNeMo LLM 服务。

其中，OpenAI 与微软 Azure 云的合作具有一定程度的示范效应，它通过深刻利益绑定，我们推测，这种合作能降低 OpenAI 训练 AI 所需的云计算成本。

- 根据《财富》杂志报道，在 OpenAI 的第一批投资者收回初始资本后，微软将有权获得 OpenAI 75% 利润，直到微软收回投资成本（130 亿美元）。
- 当 OpenAI 实现 920 亿美元的利润后，微软的份额将降至 49%。与此同时，其他风险投资者和 OpenAI 的员工，也将有权获得 OpenAI 49% 的利润，直到他们赚取约 1500 亿美元。
- 如果达到这些上限，微软和投资者的股份将归还给 OpenAI 非营利基金会。

除了利润回报，微软微软在整合 ChatGPT 等方面也占尽了先机。除了将搜索引擎 Bing 整合 ChatGPT，2023 年 1 月，微软宣布推出 Azure OpenAI 服务，Azure 全球版企业客户可以在云平台上直接调用 OpenAI 模型，包括 GPT-3.5、Codex 和 DALL·E 模型。

英伟达向 SEC 提交百亿增发申请，全球算力龙头启动战备。2 月 28 日，英伟达向美国证券交易委员会（SEC）提交近亿美元的股票增发申请，此次增发申请通过储驾发现模式进行，可在三年内自主决定发行证券的具体时机。英伟达作为全球算力龙头，在本轮 chatgpt 带动的大模型发展，科技巨头跑步进场，AI 算力底层基础设施迎确定性爆发，英伟达在此前财报会议上表示将与头部云服务商合作，提供 AI 即服务，帮助企业访问英伟达世界领先的 AI 平台，本次融资或代表全球算力龙头英伟达拉开了算力竞备序幕。

根据福布斯中国的梳理，客户使用自己的浏览器，就可以通过 NVIDIA DGX Cloud 来使用 NVIDIA DGX AI 超级计算机，该服务已经在 Oracle Cloud Infrastructure 上可用，预计不久后也将在微软 Azure、谷歌云和其他平台上线。在 AI 平台软件层，客户将能够访问 NVIDIA AI Enterprise，以训练和部署大型语言模型或其他 AI 工作负载。而在 AI 模型即服务层，英伟达将向希望为其业务建立专有生成式 AI 模型和服务的企业客户提供 NeMo 和 BioNeMo 可定制 AI 模型。

无独有偶，AI 产业算力服务商与云厂商的历史性合作，加密资产挖矿领域早已应用。例如，加密资产矿机厂商比特大陆一度推出云算力服务平台“比特小鹿”，用户无需购买矿机硬件，也可以购入挖矿算力，获得挖矿收入。这一举措，通过“化整为零”的方式，对用户/AIGC 需求厂商而言，能最大化降低入局挖矿门槛，驱动全民挖矿/AIGC 时代降临；对矿机/算力厂商而言，在矿机/芯片进入淡季、矿机/芯片库存趋增时，能通过售卖云算力的方式，平滑收入的波动；对云服务厂商而言，则有助于增加客流。

5.2 推理边缘化：边缘算力有望成为推理主体

高通发布在 **Android 手机上部署 AI 模型的解决方案**。3月2日，高通中国公众号发布了全球首个运行在 Android 手机上的 Stable Diffusion 终端侧演示。高通 AI 部门演示了如何利用高通 AI 软件栈，首次在 Android 智能手机部署 Stable Diffusion 模型。该模型是一个用文字生产图片的 AI 模型，参数超过 10 亿，过去只能在云端计算集群内运行。

边缘算力有望成为推理主体。在本次实验中，高通 AI 团队从 Hugging Face 的 FP32 1-5 版本开源模型入手，通过量化、编译和硬件加速进行优化，使其能在搭载第二代骁龙 8 移动平台的手机上运行。我们认为，未来 AI 的推理过程，通过一定的针对性优化后，完全有能力通过边缘算力实施。同时，边缘算力具有低时延、安全、隐私等优势，符合未来 AIGC 时代，对于 AI 创作所有权和隐私权的要求。手机，智能模组等算力相对于云端访问，对于普通用户来说更加便捷和易于学习，对于 AIGC 应用的推广也更加有利。

边缘算力是未来算力体系的重要一环。当前，无论是需要超强硬件支撑的 AI 迭代与训练，还是运行要求相对较低，需求相对分散的 AI 推理，都放在超算中心内进行。我们认为，在 AI 大爆发周期内，迭代和训练需要的整体算力将会呈指数级增长，增速将会超过单芯片算力增长速度。同时，单个 AI 超算规模将会受到功耗、土地、散热等因素制约。因此，未来的 AI 运算将呈现出训练与迭代在云端，推理与内容生产梯度分布（云侧+雾侧+边缘侧）的格局变化。此外，随着 AIGC 内容愈发丰富，从简单的文字发展到视频、虚拟场景，如果采用云生成然后发送到端的形式，将会产生较多的网络带宽成本和一定程度的时延，进而影响模型的商业化进程与用户使用体验，边缘算力有助于作为补充手段改善这一情形。

我们注意到，比特币矿机厂商嘉楠科技推出了自研的边缘 AI 芯片，其董事长张楠康在一次公开发言中表示，未来会有更多计算发生在边缘侧或者端侧。2022 年 11 月，嘉楠科技宣布即将推出为立体视觉和高性能 AI 而生的端侧 RISC-V AIoT 芯片 K230。值得一提的是，嘉楠科技作为比特币矿机厂商，是首个交付全球 7nm ASIC 芯片的企业，2016 年启动 AI 人工智能芯片研发，2018 年发布了全球首款基于 RISC-V 的 AI 芯片。嘉楠科技的布局，也体现出 AI 算力和加密资产算力在商业模式等层面的一种共通性。

图表 16: 嘉楠科技边缘计算芯片 K230



资料来源：腾讯云，国盛证券研究所

投资建议

综上所述，算力、数据和场景是当下 A 股的主要投资方向。算力层面，我们认为硬件开
销的规模由应用决定，在 AIGC 赋能千行百业的情况下，着重于跟踪应用端的边际变化。
AI 硬件仍以英伟达 GPU 产业链为主；在能耗、电费因素的影响下，东数西算的重要性将
强化，新疆、内蒙古、云南、四川、贵州、西藏等火电、水电、太阳能资源丰富的省份
将成为 AI 大规模训练、应用的优势区域。目前国内的 AIGC 竞赛刚刚开始，预计将首先
出现英伟达 A100 需求的趋紧，随着规模增加，能耗问题也将凸显。

同时，我们看到国内大量拥有用户、IP 的应用型企业也在跃跃欲试，本身对于文字创作、
广告、游戏等赛道而言，“备战 AIGC”将成为当务之急。一方面，其缺乏大模型能力，
必须选择与 OpenAI 或百度等企业的合作；另一方面，其手中有大量的细分场景和付费
群体，也是大模型企业期待落地的合作伙伴。我们认为，如能在场景应用、合规等方面
做好自己的“小模型”，亦有望在 AIGC 时代脱颖而出。

图表 17: AIGC 投资标的

赛道	公司	代码	AIGC
自然语言 处理 NLP	谷歌	GOOG	深度学习模型 Transformer, AIGC 生成应用 Imagen
	微软	MSFT	自然语言处理用 Apache Spark, 投资了 OpenAI
	科大讯飞	002230	语音识别和人工智能
	拓尔思	300229	通过基于大数据和 NLP 的知识图谱的构建, 为虚拟人安装知识大脑
AIGC 生成 算法和数据集	英伟达	NVDA	Omniverse, 生成对抗网络 PoE-GAN, StyleGAN
	Meta	Meta	AIGC 生成应用 Meta AI, 多语言模型 XLS-R
	百度	BIDU	拥有文心大模型, 不仅可以自然语言处理, 还可以生成修复图片
	视觉中国	000681	通过 AIGC 生成艺术照片, 拥有高质量图片数据集
	昆仑万维	300418	旗下海外平台 StarX 完成了首批 5 首完全 AI 作曲的歌曲
	万兴科技	300624	公司旗下首款 AI 绘画软件“万兴 AI 绘画”正式开启公测
	中科金财	002657	通过 AIGC 企业级内容互动创作平台, 实现与元宇宙营销平台“觅际”融合
	汉仪科技	301270	通过 AI 技术有效提高 C 端字体的生产效率及效果
算力	蓝色光标	300058	发布一键生成抽象画平台“康定斯基模型”, 用于 AI 绘画
	寒武纪	688256	人工智能芯片提供商
	天孚通信	300394	全球光模块及数通厂商光学器件供应商
	中兴通讯	000063	5G 设备商龙头
	新易盛	300502	400G 时代领先的光收发器件解决方案提供商
	美格智能	002881	物联网智能模组供应商
	奥飞数据	300738	布局东数西算的 IDC 服务提供商
中际旭创	300308	全球光电子领军企业, FB 供应商	

资料来源: 公开资料, 国盛证券研究所

风险提示

AIGC 技术发展不及预期。虽然 AIGC 产业的技术在飞速迭代中，但可能遇到阻滞。

AIGC 监管趋严。AIGC 产业可能在内容创作、数据使用等方面迎来严格监管。

免责声明

国盛证券有限责任公司（以下简称“本公司”）具有中国证监会许可的证券投资咨询业务资格。本报告仅供本公司的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本公司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。

本报告的信息均来源于本公司认为可信的公开资料，但本公司及其研究人员对该等信息的准确性及完整性不作任何保证。本报告中的资料、意见及预测仅反映本公司于发布本报告当日的判断，可能会随时调整。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。本公司不保证本报告所含信息及资料保持在最新状态，对本报告所含信息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。

本公司力求报告内容客观、公正，但本报告所载的资料、工具、意见、信息及推测只提供给客户作参考之用，不构成任何投资、法律、会计或税务的最终操作建议，本公司不就报告中的内容对最终操作建议做出任何担保。本报告中所指的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。投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。

投资者应注意，在法律许可的情况下，本公司及其本公司的关联机构可能会持有本报告中涉及的公司所发行的证券并进行交易，也可能为这些公司正在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。

本报告版权归“国盛证券有限责任公司”所有。未经事先本公司书面授权，任何机构或个人不得对本报告进行任何形式的发布、复制。任何机构或个人如引用、刊发本报告，需注明出处为“国盛证券研究所”，且不得对本报告进行有悖原意的删节或修改。

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，本报告所表述的任何观点均精准地反映了我们对标的证券和发行人的个人看法，结论不受任何第三方的授意或影响。我们所得报酬的任何部分无论是在过去、现在及将来均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。

投资评级说明

投资建议的评级标准		评级	说明
评级标准为报告发布日后的6个月内公司股价（或行业指数）相对同期基准指数的相对市场表现。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；香港市场以摩根士丹利中国指数为基准，美股市场以标普500指数或纳斯达克综合指数为基准。	股票评级	买入	相对同期基准指数涨幅在15%以上
		增持	相对同期基准指数涨幅在5%~15%之间
		持有	相对同期基准指数涨幅在-5%~+5%之间
		减持	相对同期基准指数跌幅在5%以上
	行业评级	增持	相对同期基准指数涨幅在10%以上
		中性	相对同期基准指数涨幅在-10%~+10%之间
		减持	相对同期基准指数跌幅在10%以上

国盛证券研究所

北京

地址：北京市西城区锦什坊街35号南楼
 邮编：100033
 传真：010-57671718
 邮箱：gsresearch@gszq.com

南昌

地址：南昌市红谷滩新区凤凰中大道1115号北京银行大厦
 邮编：330038
 传真：0791-86281485
 邮箱：gsresearch@gszq.com

上海

地址：上海市浦明路868号保利One56 10层
 邮编：200120
 电话：021-38934111
 邮箱：gsresearch@gszq.com

深圳

地址：深圳市福田区益田路5033号平安金融中心101层
 邮编：518033
 邮箱：gsresearch@gszq.com