

AI 算力的三条价格曲线：涨价从内存开始，在合同里结束

AI 算力的三条价格曲线 · 涨价从内存开始，在合同里结束

这轮算力涨价的源头不是需求溢价，是内存成本。NVIDIA 已把成本以超过 15% 的系统提价转出去，GPU 云以约 25% 的 SKU 提价与合同价从每兆瓦 2,000 万美元向 2,500 万美元的上移接住并转出，而房东档用十到二十年、年递增约 3% 的固定租约把自己锁在每兆瓦 170 万至 240 万美元——在一场成本推动的涨价里，唯一不能重定价的那一层就是最终的税基。至于「租金在跌」，那是把现货、合约与整柜三条不同曲线混读的结果。

涌现资本产业研究部 · 2026-08-31

30秒结论卡

项	内容
一句话判断	市场同时在传两个方向相反的价格读数——租户侧的 GPU 小时租金在跌、厂商侧的整柜与整 GW 单价在涨。把两个读数各自还原到自己的曲线上之后，只剩一个方向： 从内存到整柜到合同，全链在涨 。跌的那条线来自单一现货平台在五至六月的一段窗口，与同期的标准化指数、一年期合约价指数方向相反。
机会状态	🟡 观察 —— 涨价是真的，但它是 成本推动 不是需求溢价。可交易的变量不是价格方向，而是 这笔成本停在谁的账上 。
市场错在哪	把「价格」当成一个数。它至少是三个数：现货小时价（波动剧烈、无定性效力）、一年期合约价（2025-10 至 2026-03 上涨约 40%）、整柜系统价（VR200 NVL72 单柜约 780 万美元，为上一代的约两倍，其中内存 200 万）。三条曲线用同一个词描述，于是任何立场都能找到支持它的数字。
重点标的	🟢 NVDA（\$217.55·已把内存成本以逾 15% 的系统提价转出，三季度毛利率指引仅下调 1pp）· 🟢 IREN（\$35.45·全行业唯一在法定披露文件里写下「合同定价正在上涨」并给出每兆瓦数字的公司）· 🟡 NBIS（\$209.18·拍卖清算价创新高，但已通电容量与已签约容量差五倍）· 🟡 CRWV（\$84.23·七月对各 SKU 提价约 25%，98% 收入来自承诺型合同，压力在利息不在售价）· 🔴 WULF（\$15.35） / APLD（\$25.34） / CIFR（\$15.17）（房东档把十至二十年的价格锁在每兆瓦 170 万至 240 万美元）
最大反证	单柜价格翻倍不等于单位算力变贵。若 Vera Rubin 相对 GB300 的同代性能提升超过约 95%，则每 FLOP 成本仍在下降，「涨价」只是账单变大而不是变贵。本部本轮未取得官方性能倍数，这是全篇唯一未收口的一环，也是本篇结论会被推翻的地方。

一、投资要点

框架一句话：在一场由内存成本推动的涨价里，价格上行本身不产生阿尔法，能否把成本转出去才产生阿尔法；而合同期限，就是转出能力的倒数。

#	核心观点	方向
1	涨价的源头是内存：NVIDIA 供应承诺单季由 1,190 亿跳增至 2,790 亿美元，公司口径说明主要与内存采购有关	成本
2	NVIDIA 已把成本转出：含其芯片的服务器系统提价超过 15%，2027 年初出货生效，三季度毛利率指引仅由 75.0% 降至 74.0%	
3	GPU 云接住并再转出：CoreWeave 七月对各 SKU 提价约 25%，近期合同贡献利润率较此前高 5 至 10 个百分点	
4	合同价的一手证据一致向上：IREN 三年期合同已超每兆瓦 2,000 万美元、在谈约 2,500 万，Nebius 中期合同同区间、拍卖清算价再高 15%	
5	房东档转不出去：四家独立样本落在每兆瓦 170 万至 240 万美元、期限十至二十年、年递增约 3%，在涨价周期里卖掉了上行	
6	「租金在跌」不成立于任何标准化口径：同期 B200 指数七日仍在上行、按需价十二个月上涨 4.8%、一年期合约价上涨约 40%	
7	现货价格波动不传导到 GPU 云的收入侧：CoreWeave 98% 收入来自承诺型合同，真正的压力在利息与折旧	

二、钱流：这笔涨价从哪里发动，往哪里走

顺着账单往上游看，起点非常明确。

第一站，内存。 服务器 DRAM 合约价三季度环比上涨 13% 至 18%，已签长约的美系云厂不适用；2027 年服务器 DRAM 仍处短缺。这个价格变化在整机层被放大：Vera Rubin VR200 NVL72 单柜物料合计约 780 万美元，其中内存一项就是 200 万美元，占约四分之一，是上一代 Grace Blackwell 机柜内存账单的两倍以上。

第二站，NVIDIA 的资产负债表。 二季度供应承诺由上一季的 1,190 亿美元跳增至 2,790 亿美元，单季增加 1,600 亿，公司说明主要与内存采购有关。同期存货由 258 亿升至 316 亿。这是本轮全链最大的一笔预先锁量动作，也是「成本推动」这个判断最硬的一手依据。

第三站，价格转出。 含 NVIDIA AI 芯片的服务器系统整体提价超过 15%，2027 年初出货生效，覆盖 Vera Rubin 与 Grace Blackwell 两代，幅度随代次与内存配置浮动。转出的效果在报表上可以直接读到：二季度 GAAP 与非 GAAP 毛利率均为 75.0%，环比持平；三季度指引 74.0%，只下调 1 个百分点。**成本涨了一倍量级，毛利率只让出 1 个百分点，这就是定价权的样子。**

第四站，GPU 云。 CoreWeave 七月对各 SKU 提价约 25%，理由是需求环境与零部件成本上升；管理层称 Blackwell 与 Vera Rubin 世代的定价与利润率在创新高，而上一代 SKU 的定价处于或高于数年前水平。近期新签合同的贡献利润率比前几个季度高 5 至 10 个百分点。

第五站，戛然而止。 到了只出电力与外壳的房东档，链条断了。这一层的合同在过去十二个月里陆续签成十到二十年期、年递增约 3% 的固定租约。他们不采购内存，因此不承担这次成本；但他们也不参与这次

提价。

层级	本轮价格动作	方向
内存	服务器 DRAM 三季度合约价环比 +13~18%	● 对上游
整机系统	含 NVIDIA 芯片服务器系统提价 >15%，2027 年初生效	●
GPU 云（自持算力）	SKU 提价约 25%；合同价每兆瓦 2,000 万 → 在谈 2,500 万美元	●
托管房东（出电出壳）	每兆瓦 170 万至 240 万美元，锁十至二十年，年递增约 3%	●
现货小时市场	双向剧烈波动，无单一方向	●

三、价值链五层图：三条曲线不能放进同一张图

需求源（模型商 / 云厂 / 量化）
↓ 付「每 GPU 小时」—现货 + 承诺用量两个市场
GPU 云（自持 GPU） ← CRWV / NBIS / IREN
↓ 付「每关键 IT 兆瓦租金」—十至二十年固定 + 约 3% 递增
托管房东（出电出壳） ← WULF / APLD / CIFR / CORZ / GLXY
↓ 付「整柜系统价」—本轮提价 >15% 的那一层
整机与芯片 ← NVDA / AVGO
↓ 付「内存合约价」—涨价真正的发源地
HBM 与服务器 DRAM

这条链上的价格有三条互不通约的曲线，日常讨论里被同一个词覆盖：

- 1. 现货小时价：**Silicon Data 指数口径下，8 月 31 日 H100 每卡每小时 2.53 美元、H200 3.26、B200 5.45 至 5.58，B200 七日仍在上行 1.6%。第三方按需价格指数（47,048 条周观测、76 家供应商）给出的读数是过去十二个月上涨 4.8%、最近四周下跌 1.3%。而单一现货平台在 2026 年五月底至六月的一段窗口里录得 B200 由每小时 6.11 美元跌至 4.22 美元。**同一张纸上，三个方向。**
- 2. 一年期合约价：**H100 由 2025 年 10 月的每小时 1.70 美元升至 2026 年 3 月的 2.35 美元，上涨约 40%。
- 3. 整柜系统价：**GB300 NVL72 约 370 万至 400 万美元，VR200 NVL72 约 780 万美元。

跌幅最大的读数来自档位混读。 同一时点、同一卡型，超大厂档与市场型竞价档的价格可以相差 3 倍以上；把峰值取自前者、谷值取自后者，就能凭空造出一条七成以上的跌幅曲线。跨源比价前必须先归一到同一档位、同一供应商类型、同一计价单位——某家 GPU 云页面上每小时 49.24 美元的报价除以 8，恰好等于同页另一列的单卡价，那是八卡节点价而非单卡价。

四、卡点评分：稀缺在哪一层，定价权就在哪一层

层级	短缺强度	替代难度	集中度	定价权	可投性	象限
HBM 与服务器 DRAM	高	高	三家	强	可投	真卡点可投
整机系统（NVDA）	高	高	单极	强	可投	真卡点可投
GPU 云	中高	中	分散	中（合约期内）	可投	真卡点可投
已通电电力与场地	高	高	分散	弱（已被合同锁定）	可投	真卡点买不到
现货算力小时	波动	低	极分散	无	期货即将挂牌	非卡点

这张表里最反直觉的一格是房东档：它的物理稀缺度最高，定价权却最弱。原因不在市场结构，在合同结构。稀缺是它的，价格却是两年前签下的。四家独立样本的每关键 IT 兆瓦年收入分别是——Anthropic 与 TeraWulf 的二十年约反算约 237 万美元、Galaxy 交付给 CoreWeave 的一期约 240 万美元、Applied Digital 十五年约反算约 170 万美元、Cipher 按毛功率口径折算后约 174 万美元。四家彼此独立、口径各异，却收敛在 170 万至 240 万美元的窄带。这就是这一层的真实单价，且它已经被锁定到 2040 年代。

同一时点，自持 GPU 的那一层：IREN 近期三年期合同超过每兆瓦 2,000 万美元，在谈约 2,500 万；Nebius 中期合同同为 2,000 万至 2,500 万，六个月以内的短期合同 4,000 万至 5,000 万。两层相差约十倍，差额基本是 GPU 本身的钱。

期限即定价权的倒数：短期合同每兆瓦 4,000 万至 5,000 万，中期 2,000 万至 2,500 万，十五至二十年的房东档 170 万至 240 万。期限每拉长一档，单价掉一个台阶。在成本上行周期里，签得越长的人让渡得越多。

五、受益链：这笔内存税最后停在谁的账上

- **需求源**：模型商与云厂，承担全部终端账单，通过 token 单价与自研硅两条路径回避。
- **直接受益**：内存三巨头与 NVIDIA。前者是价格的发起方，后者是唯一验证了完整转出能力的中间层。
- **二阶受益**：GPU 云。提价 25% 且新签合同利润率上行，说明转出成功；但它同时承担了另一笔账——设备融资利息。
- **隐形受益**：定制硅通道。某 XPU 平台首期 350 亿美元对应 1 GW，与 NVIDIA 侧每 GW 约 350 亿美元的内容量同一量级——两条路线的每 GW 单价并没有拉开差距，这削弱了「自研省钱」的直觉。
- **不受益但被计入主题**：托管房东。它出售的是电力与时间，不含内存，因而既不承担成本也不分享涨价。
- **可投代理**：即将挂牌的算力期货。CME 与 Silicon Data 定于 2026-10-05 上市 H100 与 B200 两只小时租赁指数期货，每张合约代表一个月租金，NYMEX 规则下现金结算。挂牌后的首个结算价，是这条曲线第一次拥有可交易、可复核的锚。

六、公司卡

NVDA \$217.55 · 市值 52,532 亿美元 · TTM PE 27.5 · ●

位置：整机与芯片层，涨价链条的转出枢纽。为什么重要：唯一同时握有成本锁量与终端定价两端的公司——供应承诺 2,790 亿美元锁住内存，系统提价逾 15% 转出成本。证据：二季度营收 962.21 亿美元（同比 +106%），数据中心 890.23 亿（同比 +117%），毛利率 75.0%，三季度指引 1,080 亿、毛利率 74.0%。估值状态：TTM 27.5 倍，与 106% 的营收增速相比不构成拥挤。催化：三季度 Vera Rubin 放量，提价 2027 年初生效。反证：应收账款周转天数由 45 天升至 60 天，公司说明为对投资级客户放宽账期——这是需求侧融资条件的第一处松动。

IREN \$35.45 · 市值 127 亿美元 · ●

位置：自持 GPU 云 + 自有电力，全栈档。为什么重要：全行业唯一在法定披露文件中直接写出「合同定价正在上涨、带动每兆瓦价值提升」，并给出可复核数字的公司。证据：三年期合同超每兆瓦 2,000 万美元（关键 IT 口径）、在谈约 2,500 万，约两年回本；客户预付款覆盖 GPU 资本开支的 45% 至 55%；2026 年 0.3 GW、2027 年 0.8 GW 交付节奏；已具名客户包括 Microsoft、Cohere、Perplexity、Figure AI。估值状态：AI 云业务全年仅 1.288 亿美元收入，市值定价的是 2027 年。催化：2027 年产能销售落地。反证：40 亿美元合同 ARR 与 1.288 亿美元已确认收入之间相差三十倍，公司自己声明确认收入可能显著更低。

NBIS \$209.18 · 市值 569 亿美元 · ●

位置：自持 GPU 云。为什么重要：拍卖清算价高于历史最高成交价 15%，是价格上行最锋利的一手读数。证据：二季度营收 5.823 亿美元（同比 +454%），ARR 30 亿；中期合同每兆瓦 2,000 万至 2,500 万美元、预付款覆盖资本开支 50% 至 60%；七成成交带预付款。估值状态：全年资本开支指引 200 亿至 250 亿美元，远大于收入。催化：年底已通电容量。反证：年底已签约电力目标 5 GW，而已连接电力仍为 800 MW 至 1 GW，相差五倍——单价再高，也要先通电。

CRWV \$84.23 · 市值 465 亿美元 · ●

位置：最大的自持 GPU 云。为什么重要：验证了「现货波动不传导到收入」这条链路——98% 收入来自承诺型合同。证据：二季度营收 25.75 亿美元（同比 +112%），已通电电力单季增近 500 MW 至 1.5 GW；七月各 SKU 提价约 25%。估值状态：经营亏损 0.49 亿、净亏 6.26 亿。催化：合同交付转收入。反证：二季度利息支出 6.40 亿美元，三季度指引 8.60 亿至 9.40 亿——售价上行的同时，资金成本上行更快。这批股票真正的风险从来不在租金曲线上，在利息与折旧两行。

WULF \$15.35 · APLD \$25.34 · CIFR \$15.17 · ● （房东档合并卡）

位置：出电力与外壳。为什么重要：本轮涨价唯一完全的旁观者。证据：TeraWulf 与 Anthropic 二十年约 401 MW 关键 IT、约 190 亿美元；Applied Digital 合同 IT 负载 1,410 MW、十五年基础期约 360 亿美元；Cipher 700 MW 毛功率、平均年化 NOI 约 7.93 亿美元，逐年递增约 3%。估值状态：三家均为负 PE，市值定价的是尚未交付的容量。催化：交付节点。反证：它们的物理资产在涨价，合同现金流不在。十至二十年、约 3% 年递增的租约，等于在通胀刚开始时把整条价格曲线卖成了债券。

七、情景与标的表

情景	核心假设	结果	动作
熊（约 25%）	同代性能提升超过单柜涨幅，每 FLOP 成本继续下降，涨价被性能吃掉；同时新供给上线压低承诺用量价	提价无法转成利润，GPU 云利润率回吐	● 回避高杠杆全栈档
基准（约 55%）	内存紧缺延续至 2027，提价按期生效，合同价维持在每兆瓦 2,000 万至 2,500 万美元	上游与整机赚定价，全栈档赚量，房东档赚资本成本	● 上游与整机 · ● 全栈档
牛（约 20%）	2027 年合同价站上每兆瓦 2,500 万美元并转成确认收入，通电节奏兑现	全栈档同时享受量与价	● IREN / NBIS

标的	现价	位置	定价权	状态
NVDA	\$217.55	整机与芯片	强，已验证转出	●
IREN	\$35.45	全栈，自有电力	强，合同价上行有一手披露	●
NBIS	\$209.18	全栈	强，通电是约束	●
CRWV	\$84.23	全栈	中，压力在利息	●
AVGO	\$368.79	定制硅通道	中，每 GW 单价与通用路线同量级	●
WULF	\$15.35	房东	弱，已锁二十年	●
APLD	\$25.34	房东	弱，已锁十五年	●
CIFR	\$15.17	房东	弱，年递增约 3%	●

八、数据冲突与口径

冲突一：B200 现货价的方向。 单一现货平台录得 2026 年五月底至六月三周内由每小时 6.11 美元跌至 4.22 美元；同期标准化日频指数在八月底仍显示七日上行 1.6%，而同一指数在三月单月上行 23.6%。本部取舍：现货曲线在 2026 年是双向剧烈波动，取样窗口决定结论，因而单独的现货读数不具定性效力。有定性效力的是合约价与整柜价，两者同向上行。

冲突二：表外承诺的总量口径。 九家大型科技公司表外承诺合计约 3 万亿美元，约为其最近十二个月合计资本开支的五倍；此前针对五家公司的口径为 1.65 万亿美元。家数、是否含未启动租赁、是否含能源采购三项差异即可造成接近翻倍的差距，跨报道直接比较无效。本部在正文只使用可回到单一发行人法定披露的数字：NVIDIA 自身的未来承诺为 560 亿美元（AI 云协议 360 亿 + 未开始的第三方数据中心租赁 200 亿），另有担保最大敞口 1,085 亿美元。

冲突三：容量口径。毛功率与关键 IT 负载的实测折算率为 66.5%（Galaxy 交付 200 MW 毛功率对应 133 MW 关键 IT）。Cipher 使用毛功率口径，其余样本使用关键 IT 口径，直接横比会低估 Cipher 单价约三分之一。本篇所有每兆瓦单价已统一到关键 IT 口径。

冲突四：担保标的。NVIDIA 披露的担保标的是土地、电力与外壳租约，最大敞口 1,085 亿美元，**不是 GPU 残值**。任何以「芯片残值被兜底」为前提的推论不适用于已披露事实。

九、风险与反证条件

#	风险	什么条件出现会推翻本篇结论	影响标的
1	涨价被性能吃掉	Vera Rubin 相对 GB300 的官方同代性能提升超过约 95%，则每 FLOP 成本仍降，「涨价」只是账单变大	全链
2	合同价不转成收入	IREN 2027 年确认收入与其 40 亿美元合同 ARR 的差距未收窄至可解释区间	IREN / NBIS
3	利息吞掉售价	CoreWeave 三季度实际利息支出高于 9.40 亿美元指引上沿，且贡献利润率回落	CRWV
4	通电滞后	Nebius 年底已连接电力低于 800 MW 下沿	NBIS
5	内存缓解	2027 年服务器 DRAM 合约价环比转负，则成本推动逻辑消失，提价面临回吐压力	NVDA / 全栈档
6	房东档被重新定价	若出现十年期以内、带通胀挂钩条款的新签托管合同，则「卖掉上行」的判断需要收窄至存量合同	WULF / APLD / CIFR
7	现货与合约脱钩终止	期货挂牌后现货指数与承诺用量价开始联动，则合同隔离这一层保护减弱	CRWV / NBIS

十、行动清单

- **现在关注：**整机与芯片层的转出验证（毛利率是否守住 74% 一线）；IREN 2027 年产能销售的定价披露。
- **只监控不买：**算力期货首个结算价（2026-10-05 挂牌），它将第一次给现货曲线一个可交易的锚；在此之前所有租金读数都缺少统一基准。
- **绝不追：**以「租金暴跌」或「租金暴涨」单一现货读数为论据的任何一侧交易。
- **下次看这五个数：**①三季度服务器 DRAM 合约价环比；②NVIDIA 三季度实际毛利率与供应承诺余额；③CoreWeave 三季度利息支出与贡献利润率；④Nebius 年底已连接电力；⑤IREN 在谈合同是否落到每兆瓦 2,500 万美元。

- **出现即重写结论：**任一家托管房东签出带通胀挂钩条款的新约，或任一家 GPU 云披露合同价环比下行。

附录 · 十一、来源置信度表与数字三道关自评

数字三道关自评：①量级一致性——NVIDIA 二季度数据中心 890.23 亿美元 < 总营收 962.21 亿 < 三季度指引 1,080 亿，逐级自洽；CoreWeave 上半年 46.53 亿 = 一季度 20.78 亿 + 二季度 25.75 亿；IREN 全年 7.070 亿美元 > AI 云 1.288 亿，子项小于总额。②单位显式化——全文单价统一为「每关键 IT 兆瓦·年」与「每 GPU·小时」两种，合同期总额一律已除以年限并标注年限，机柜价统一为「每柜美元」。③双源与公告优先——所有合同与财务数字取自 SEC 8-K、10-Q 或公司 IR 披露；卖方建模值与第三方指数单独标级，不与公司披露混列。

结论 / 数字	来源	口径	置信度
NVIDIA 营收、毛利率、指引、供应承诺 2,790 亿、担保敞口 1,085 亿	SEC 8-K 与 10-Q, 2026-08-26	单季 / 存量	A
服务器系统提价逾 15%	Bloomberg, 2026-08-22	厂商侧主动提价, 2027 年初生效	B
服务器 DRAM 三季度合约价 +13~18%	TrendForce, 2026-07-09	合约价, 含 LTA 例外	A
IREN 合同价每兆瓦逾 2,000 万、在谈 2,500 万、预付款 45~55%	SEC 8-K, 2026-08-27	关键 IT 口径	A
Nebius 中期 / 短期合同单价、拍卖清算价 +15%	二季度电话会, 2026-08-19	管理层表述	B
CoreWeave 提价约 25%、贡献利润率 +5~10pp、98% 承诺型收入	二季度电话会与 10-Q, 2026-08-11 / 08-12	管理层表述 / GAAP	A~B
房东档每兆瓦 170 万至 240 万美元	各公司 8-K 与公告	合同期总额已除年限, 未计递增条款	B (自算)
毛功率→关键 IT 折算 66.5%	Galaxy 一期交付公告	同一公告内双口径并列	A
VR200 NVL72 单柜约 780 万美元、内存 200 万	卖方物料拆解, 2026 年 8 月	建模值, 非官方定价	B
TPU 每 GW 约 270 亿美元	卖方模型, 2026-08-25	分析师建模假设, 非公布价目	B
现货与合约指数读数	Silicon Data / SemiAnalysis / 第三方按需指数	单卡·小时·标准化	A~B
算力期货 2026-10-05 挂牌	交易所官方公告, 2026-08-11	两只指数期货, 现金结算	A
现价 (全部标的)	实时行情接口, 2026-08-31	最新成交	A

未取证，不进数字表：①Vera Rubin 相对 GB300 的官方性能倍数——这是第九节第 1 号反证条件无法当期收口的原因；②各家合同是固定价还是随现货浮动，三家全栈档均未披露定价调整机制；③加权平均合同期限，CoreWeave、Nebius、IREN 均未披露；④GPU 二手成交价数据库，现有读数均为挂牌价；⑤定制硅路线的单机架报价，公开口径只有每 GW 总量。

价格不是一个数。在把任何「算力涨了」或「算力跌了」写进模型之前，先回答三个问题：哪一条曲线、哪一个档位、哪一种期限。三个标签缺任何一个，这个读数就不能用来做决定。

本报告由涌现资本产业研究部独立撰写，所载数据来自公开披露与实时行情接口，力求准确但不保证完整。报告仅供参考，不构成任何证券之要约、招揽或投资建议。市场有风险，决策需独立判断与尽职调查。