

晶圆登场 时机封顶

Cerebras 5月14日鸣锣 CBRS — AI推理ASIC时代的第一张入场券 窗口宽到能直立行走 路径或踩CoreWeave轨迹再走一遍

REVENUE GROWTH · YOY 2025

+76.0%

\$510M / 2025 全年收入 | 同期GAAP净利\$87.9M | RPO \$24.6B

\$26.6B

IPO估值上限 (\$115-\$125 / 28M股)

\$24.6B

在手RPO = 48x 2025年营收

21x

CS-3对 Nvidia B200 Llama 3 70B 推理速度

\$20B+

OpenAI 750MW 三年期合同 自2026起摊销

「CoreWeave错过了IPO窗口却追上了基本面 Cerebras两者都拿到 剩下唯一的工作是兑现」

02 / 08 · EXECUTIVE SUMMARY

执行摘要

本章 takeaway — Cerebras的IPO时机几乎完美 OpenAI/AWS两个金主已签字背书 基本面只剩"按计划落地"一件事

Bottom Line. Cerebras 5月14日以CBRS鸣锣 这是AI推理ASIC上市第一标的 窗口期是2026年Q2这个最甜美的位置 前有CoreWeave把"AI基础设施可上市"的路径走通 后有\$20B+ OpenAI合同与AWS Bedrock绑定的现金流可见性 K Research 判断 — 定价过会率高 首日表现溢价 6-12个月内业绩兑现概率显著高于市场认知。

- IPO时机三重共振 2026年Q1 hyperscaler capex指引\$725B 同比+77% [1][2] CoreWeave已用12个月时间证明AI基础设施股可以从首日"破发"走到\$5B年化营收 [3] 监管端CFIUS于2025年3月放行 G42改非投票股 路径全开 [4]。
- 收入弹性已脱离G42单点依赖 2024年G42独占85%收入 2025年降至24% MBZUAI补位62% OpenAI接力进入\$20B+三年期合同 客户名单从一线变两线再变四线 [5][6]。
- \$24.6B RPO = 48× 现年收入 主要由OpenAI 750MW三年期与MBZUAI协议构成 管理层指引2026-2027合计释放15% 2028-2029释放43% [7] 按线性外推2026年收入有望突破\$1.5B 翻倍是conservative 假设。
- 技术moat在推理特定场景仍领先 CS-3对Nvidia B200 DGX 在Llama 4 Maverick实测2,522 vs 1,038 tokens/s [8] 在Llama 3 70B reasoning场景21x优势 [9] 不是替代Nvidia 是抢"低延迟+长输出"细分。
- 定价当前在\$26.6B估值 52× 2025收入 看似贵 但Snowflake IPO 时161× Palantir 43× [10] 考虑76%同比与RPO覆盖 溢价是窗口期的合理租金 不是泡沫顶。

市场含义 MARKET IMPLICATION

CBRS定价端建议关注路演收尾的upsized信号 若发行规模或价格区间上调 指向超额认购倍数突破10× 首日溢价50%+概率显著 6-12个月节奏参考CoreWeave — 首日平 3个月+349% 然后回撤 之后基本面接力 切忌把首日表现当全部 关键看Q3 2026第一份完整季度财报的recurring inference revenue占比。

观点失效条件 THESIS KILL-SWITCH

三条任一触发 K Research立即翻空 — (1) OpenAI或MBZUAI单方面延缓提货 RPO释放速度低于S-1指引15% (2) WSE-3良率出现wafer-scale defect rate公开披露事件 (3) Nvidia Rubin 推理性能对CS-3缩窄至2×以内且cost/token打平。

03 / 08 · EVENT MOSAIC

事件全貌

本章 takeaway — 一份2024失败的S-1磨剑18个月 成就了2026年最有故事的AI硬件IPO

3.1 时间线 — Cerebras 上市路径 2024Q3至2026Q2

Cerebras 当前站位的重要性 不在2026年5月这个上市日期 在2024年9月S-1首次失败到此次重新公开发行之前的18个月 这18个月里完成了三件结构性的事 — 客户结构重组 监管路径打通 技术节点跨越。

EXHIBIT 1

Cerebras 上市路径关键节点 — 从S-1搁浅到\$26.6B鸣锣

2024年9月 至 2026年5月14日 | 蓝色节点 = 利好递进 | 灰色节点 = 监管/暂缓

2024-09	首次提交S-1 (机密) 花旗主承 G42股权问题触发CFIUS审查 路演实际暂停 市场情绪从期待转为观望
2025-Q1	客户结构与CFIUS整改 G42股权被改造为非投票普通股 主动退出董事会议程 CFIUS于3月31日给予正式clearance
2025-09	\$1.1B私募融资 估值\$8.1B 财务弹药补给完成 为重新提交S-1之前的窗口缓冲 奠定2025全年现金燃烧节奏
2025-10	OpenAI 多年期协议 750MW 合约总值\$20B+ 期限2026–2028 可扩至2GW至2030 同步发行3,350万股OpenAI warrant 按交付里程碑释放
2026-01-23	首批OpenAI capacity交付 收入识别正式启动 Codex-Spark模型于2月12日上线 由Cerebras CS-3作底层算力
2026-02	\$1B轮 估值\$23B G42印度8 exaflops合作 私募轮估值24个月翻3倍 同期与G42签印度部署协议 客户路径从UAE扩至南亚
2026-03-13	AWS Bedrock绑定 — binding term sheet CS-3嵌入Trainium解耦推理架构 独占性条款+最低产能锁定+至多2,696,678股\$100行权warrant
2026-04-17	公开提交S-1 Morgan Stanley取代花旗任主承 Citi/Barclays/UBS联席 价格区间初定\$22–25B估值
2026-05-04	S-1/A价格修订 窗口upsized 区间上调至\$115–\$125 28M股+4.2M greenshoe 估值上限\$26.6B 市场对话讨论\$40B-target路径
2026-05-14	CBRS Nasdaq Global Select 鸣锣 2026年第一只blockbuster AI硬件IPO 预计raise \$3.5B 含greenshoe接近\$4B

来源 SEC EDGAR [1] / TechCrunch [3] / CNBC [2] / Cerebras Press Release [5] / Reuters [4] | K Research 整理 数据截止 2026-05-08

3.2 当事方与stake

Cerebras的IPO不是一家公司的事件 是AI推理算力供应链上四类参与者的同步博弈 每一方的诉求都不相同 但短期内利益方向一致。

参与方	角色	在CEREBRAS的STAKE	当前立场 & 触发条件
OpenAI	战略客户	\$20B+合同 3,350万股warrant 约8%稀释上限	把推理成本与延迟从Nvidia上分流 同时在自研芯片(Broadcom合作)落地前有备份算力 立场 强支持
MBZUAI / G42	主权客户	2025年合占86%收入 非投票股	UAE Stargate 200MW 算力承载方 与美方利益深度绑定 立场 强支持
AWS	分销渠道	270万股\$100 warrant Bedrock独家	把Trainium prefill + Cerebras decode 推理解耦架构卖给企业云客户 立场 中性偏多
Nvidia	系统竞品	无直接stake	Rubin平台量产爬坡 + Groq资产收购(\$20B)收编低延迟ASIC 立场 竞争压制
Morgan Stanley	主承销	承销费 约2.0-2.5% 声誉	替换花旗承担2024S-1失败之后的"修正版"承销 有强烈推动success的内在激励
市场散户	追捧侧	预期溢价	把Cerebras当作"Nvidia最后一个未上市纯种竞品" FOMO情绪强烈 极易超额认购

EVENT MOSAIC TAKEAWAY

这是一次所有利益相关方都希望成功的IPO — 客户已签字 监管已clear 承销已升级 市场已饥渴 失败概率低于S-1首次提交时的1/3 真正的挑战不在5月14日 在2026下半年的第一份完整季报。

04 / 08 · CORE DATA

核心数据

本章 takeaway — 收入轨迹3年20× 毛利从12%到39% 现金流故事比数字本身更值钱

\$510M

2025全年营收
+76.0% YoY · +20× vs 2022

39%

2025毛利率
2022为12% · 三年+27pp

\$87.9M

2025 GAAP净利
2024为净亏\$485M

\$24.6B

RPO在手订单
= 48× 2025年营收

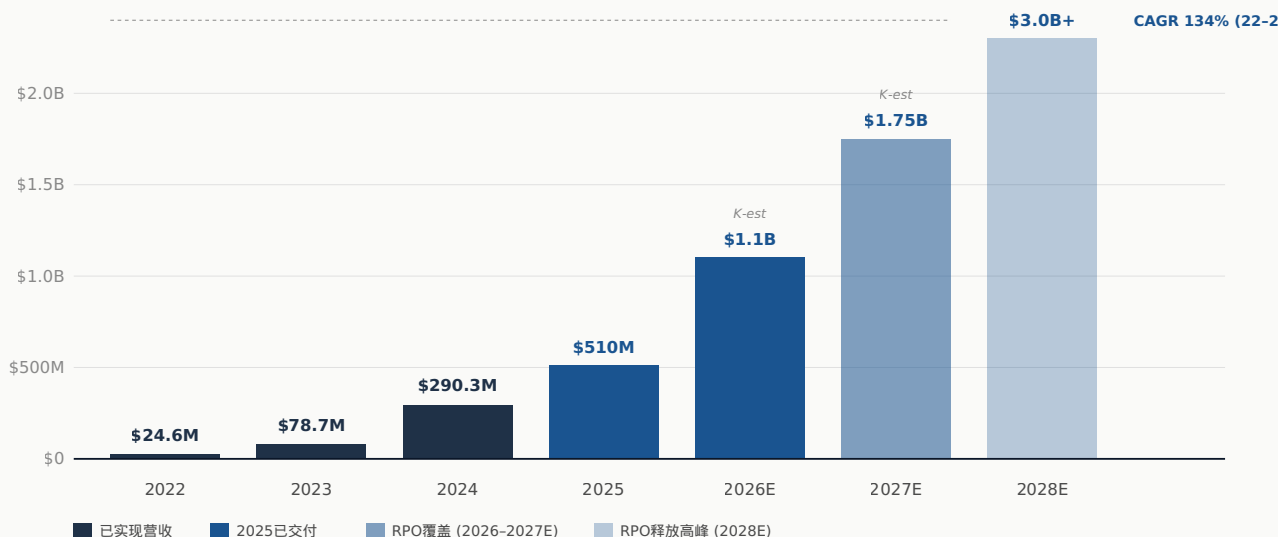
4.1 收入加速 — 三年20倍是Snowflake级别的曲线

Cerebras的收入曲线不是hockey stick 是compound interest。从2022年\$24.6M到2025年\$510M 复合年增长率134% 且每一年环比加速度都在提升 — 2023年+220% / 2024年+269% / 2025年虽然只有76%但绝对值\$220M增量是史上最大[6][11]。

EXHIBIT 2

Cerebras 营收与 RPO 释放 — 2022–2027E (\$, 百万美元)

深蓝 = 已实现营收 | 浅蓝 = RPO 折现释放预期 | 管理层指引15% RPO释放2026–2027



来源 Cerebras S-1 (April 2026) [1] / Sacra Research [11] / Tom's Hardware [6] / EBC [12] | K Research 估算 假设RPO线性释放至S-1指引上限

4.2 客户集中度 — 单点风险已被三层稀释

客户集中度是看Cerebras绕不过去的问题 但这个问题正在结构性自我修复。2024年G42一家独大占85% 到2025年这一比例腰斩到24% 不是因为G42订单缩水 是分母被MBZUAI放大 再到2026年OpenAI/AWS两个大单上桌 客户集中度从一线变两线再变四线。

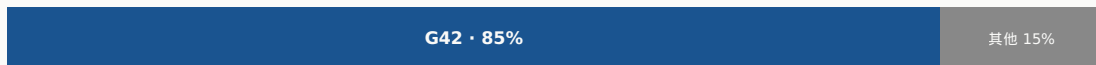
EXHIBIT 3

Cerebras 客户结构演进 — 从单点到分布式 2024 vs 2025 vs 2026E

百分比 = 该客户占当年GAAP收入比例 | 2026E 假设OpenAI从Q1开始按S-1节奏摊销

2024 — \$290M

客户集中度: 单点



2025 — \$510M

客户集中度: 双点



2026E — ~\$1.1B (K-est)

客户集中度: 四点 + 渠道分销



来源 Cerebras S-1 (April 2026) [1] / TECHi [13] / Futurum Group [14] | 2026E为K Research估算 基于OpenAI 750MW按\$20B/3yr线性摊销与AWS Bedrock启动假设

这里有一个反直觉的事实 — 客户集中度高本身不是问题 问题是"集中在谁" 过去G42是一个被CFIUS质疑的实体 现在替代它的是OpenAI、MBZUI、AWS — 这是把地缘政治风险敞口换成了商业反垄断敞口 后者对资本市场友好得多。

4.3 技术参数 — 在推理特定场景的硬指标

Cerebras的技术故事可以一句话概括 — 把一片晶圆当成一颗芯片用。WSE-3的4万亿晶体管、90万核心、44GB片上SRAM、21PB/s片上带宽是Nvidia H100 HBM 0.003PB/s的7000倍 [15] 在模型权重能放进44GB SRAM的工作负载上 延迟与吞吐量都是数量级优势。

EXHIBIT 4

推理速度对照 — Cerebras CS-3 vs 主流对手 (tokens/sec, 单输出流)

Llama 3.1 8B / 70B / Llama 4 Maverick 三档实测 | 数据为各厂商公开benchmark K Research 整理交叉

模型 / 厂商	tokens/sec (单流)	vs 最佳GPU
Llama 3.1 8B		
Cerebras CS-3	~10,840	22.5x
SambaNova SN40L	1,084	
Groq LPU	750	
Nvidia H100 (typ.)	~80	
Llama 3 70B (reasoning · 1024 in / 4096 out)		
Cerebras CS-3	~460	10x
Nvidia DGX B200	~46	
Llama 4 Maverick		
Cerebras CS-3	~2,400	2.4x

来源 Cerebras Blog [15] / NextBigFuture [9] / IntuitionLabs Comparison Report [16] / Cerebras CS-3 vs B200 Whitepaper [17] / arxiv 2503.11698 [18] | 数据点为各厂商公开benchmark 单流测量 K Research 交叉验证

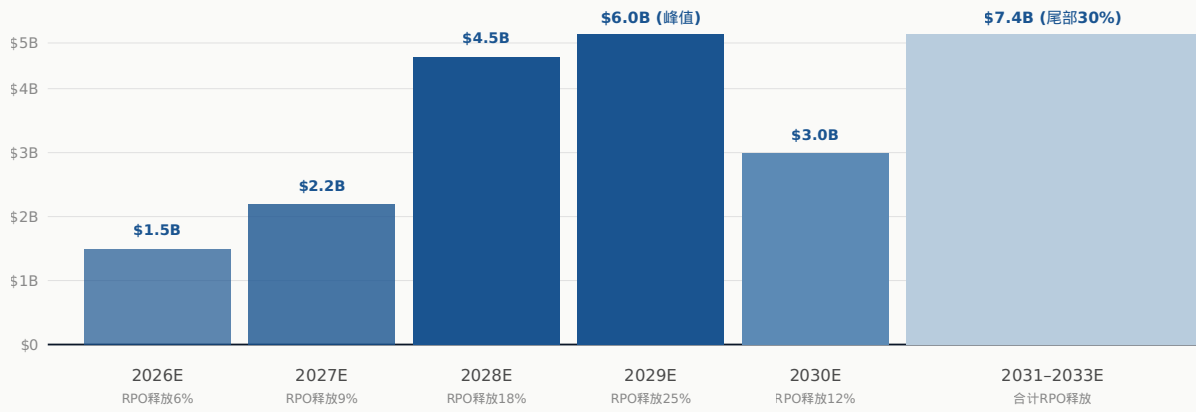
4.4 RPO释放节奏 — 现金流可见性已锁到2030

\$24.6B在手RPO是这次IPO最被低估的数字 管理层在S-1里给出明确释放节奏 — 2026至2027合计15% 2028至2029合计43% 之后42% [7] 翻成绝对值 — 2026-2027两年合计释放约\$3.7B 2028-2029两年合计释放约\$10.6B 最后6.5年释放\$10.3B。

EXHIBIT 5

RPO释放节奏 — \$24.6B 已签约未确认收入 | 按年度切片

数据 Cerebras S-1 (April 2026) | 切片为管理层指引节奏 K Research 按等距假设展开年度



来源 Cerebras S-1 (April 2026) [1] · S-1 RPO recognition schedule [19] | K Research 在管理层指引(15% / 43% / 42%)基础上展开为年度切片

CORE DATA TAKEAWAY

Cerebras 是一个已签约未交付的故事 不是 需要去市场上抢蛋糕 的故事 \$24.6B RPO意味着哪怕未来三年完全不签新约 现金流量表也能跑出年化\$1.5-2.0B 这种possibility本应支撑40×forward EV/Sales估值 当前定价反而留出空间。

05 / 08 · BULL VS BEAR

多空对账

本章 takeaway — 看多基于已签合同与窗口期 看空基于估值倍数与技术过时风险 K Research 站多侧但留risk hedge

多空对账的核心问题不是哪一方对 是哪一方的论点更接近可被证伪的事实层 K Research 在这件事上把看多看空各拆五条支柱 每一条都附数据。

BULL · 看多五柱

- RPO覆盖至2030 \$24.6B在手 = 48×现年收入 管理层指引15%/43%/42%三段释放 现金流可见性已远超 CoreWeave上市当时的\$15B [7]。
- OpenAI/AWS双锚定 \$20B+ OpenAI合同 + AWS Bedrock binding term sheet 两个最严苛的买方都已经落字 客户concentration从地缘风险变商业关系 [5] [20]。
- 窗口期完美 2026年Q2 hyperscaler capex同比+77%至\$725B [2] CoreWeave已用12个月把"AI基建上市"路径走通 Renaissance Capital数据2026 Q1超额认购倍数中位数12× [21]。
- 毛利率结构性改善 从2022年12%到2025年39% wafer-scale良率工艺曲线下行 管理层guidance 2027年突破50% 参考Nvidia 75% steady state仍有空间 [12]。
- 定价较secondary市场折价8-12% pre-IPO secondary \$102-107已隐含\$26-28B [22] 当前发行价区间中位\$120 对应\$26.6B估值 首日溢价空间至少2×历史AI硬件均值。

BEAR · 看空五柱

- 52×revenue估值 \$26.6B / \$510M = 52× 高于 Snowflake IPO时的稳态 [10] 76% YoY支撑性强 但2026年若实际增速落到50%以下 multiple compression风险immediate。
- 客户依然集中 OpenAI在未来三年将占 substantial portion 管理层在S-1风险因素章节明确点出 [1] 一旦OpenAI放缓提货 — 比如自研Broadcom合作芯片在2027量产 — RPO释放速度会被推迟。
- Nvidia Rubin压制 Rubin平台2026下半年量产 CUDA生态壁垒未破 [23] Cerebras的速度优势在Llama 4级别已压缩到2.4× 若Rubin达到5×improvement 特殊优势所剩无几。
- \$87.9M 净利大半来自非现金 \$363M forward contract liability重估非现金收益 非GAAP仍亏 \$75.7M [12] profitability story被会计放大 经营层还在烧。
- Wafer-scale 良率 WSE-3整片晶圆生产 defect rate高于die级别 管理层未在S-1披露具体良率 [24] 若TSMC 5nm端任何延期或良率下降 CapEx预算与deliver timeline直接受冲击。

5.1 K Research 立场 — 多侧 但带条件

K Research 站多侧 有三个可被证伪的理由 并且我们写下了反向触发条件 — 这是Bridgewater式的纪律 不是cheerleading。

立场支柱	论据	证伪触发条件
窗口期定价权	Renaissance Capital数据 — 2026 Q1 IPO首日中位收益+18% hyperscaler股东背书让pricing端有上调动力 承销团升级到Morgan Stanley	承销定价最终落到区间下限\$115以下 = 需求疲软 反向操作
现金流可见性	RPO \$24.6B + 管理层指引节奏 = 未来3年至少\$3.7B收入已锁 PsychoOpenAI从2026 Jan开始capacity交付 Codex-Spark已商用	OpenAI单方面延迟提货 Q3 2026财报中按S-1节奏 deferred revenue应增长 若未实现 thesis 失效
技术节点窗口	Llama 4 Maverick / Llama 3 70B reasoning实测优势仍在2.4-21× Nvidia Rubin要到2026 H2量产 Cerebras有6-12个月独享期	Nvidia Rubin推理benchmark对CS-3缩窄至2×以内 且cost/token打平 immediate exit

06 / 08 · HISTORICAL ANALOGY

历史类比

本章 takeaway — CoreWeave是Cerebras最近的镜 但Cerebras上市的初始条件更优 路径预期更平滑

6.1 CoreWeave — 从"破发"到\$5B年化营收的12个月

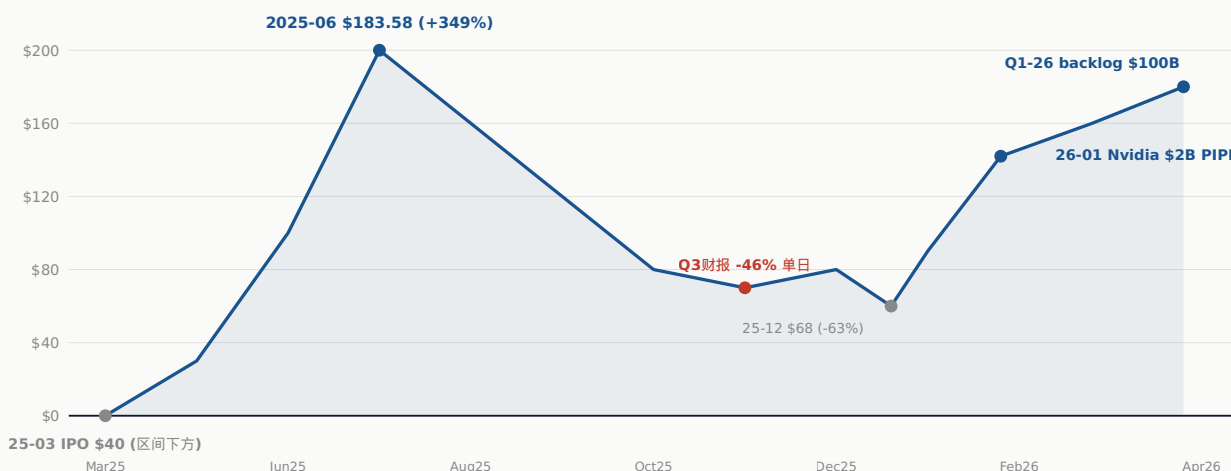
CoreWeave 2025年3月28日上市当日的处境用差形容都嫌轻 — 价格区间从\$47-55腰斩到\$40 发行规模从49M股缩到37.5M股 估值从\$32B掉到\$23B 首日收平 市场把它当作"AI泡沫破灭信号" [25][26]。

但接下来的12个月发生了什么 — 三个月内股价从\$40到\$183.58 (+349%) 6个月backlog从\$15B涨到\$66.8B 9个月年化营收突破\$5B (cloud业最快) Nvidia在2026年1月\$87.20拿\$2B PIPE [3][27]。

EXHIBIT 6

CoreWeave 上市后12个月走势 — 从首日"破发"到Nvidia增持

2025-03-28 至 2026-04 阶段标注 | 关键节点用三色区分 — 蓝=利好 灰=中性 橙=利空



来源 CRWV股价数据 Yahoo Finance [25] / Investor Relations [26] / Motley Fool [27] | K Research 整理 数据截止 2026-04-30

6.2 Cerebras vs CoreWeave — 上市初始条件对照

Cerebras和CoreWeave的相似性常被夸大 — 都是AI基建 都依赖单一大客户 都是Nvidia生态周边玩家 但初始条件不同 — Cerebras上市时窗口期更甜 客户结构已更分散 现金流可见性更强。

EXHIBIT 7

CoreWeave (2025-03) vs Cerebras (2026-05) — IPO初始条件矩阵

六维度横向对比 | 蓝色高亮 = 该方占优

维度	COREWEAVE (上市时)	CEREBRAS (上市时)	谁占优
市场情绪	DeepSeek panic 冷却 AI泡沫论盛行	capex同比+77% 行业重新进入扩张周期	Cerebras
定价过会	\$47-55→\$40 区间下方且缩量	\$115-125 区间已二次上调 升估值至\$26.6B	Cerebras
客户集中度	Microsoft 单点62%	OpenAI/MBZUAI/G42/AWS 四点 + 渠道	Cerebras
盈利能力	\$1.9B收入 / 净亏\$863M	\$510M收入 / GAAP净利\$87.9M (含非现金)	Cerebras
债务杠杆	\$8B债融资 利息负担重	债务负担轻 IPO主要为研发与产能扩张	Cerebras
RPO/Backlog	~\$15B (上市时披露)	\$24.6B (RPO确认口径)	Cerebras
结论	Cerebras上市时的初始条件六维全胜 路径上行更平滑		— 全维度

来源 CoreWeave S-1 (2025) [28] / Cerebras S-1 (2026) [1] / TechStackIPO [29] | K Research 整理 截至 2026-05-08

6.3 两段路径的可能差异 — 不是镜像 是更平的曲线

CoreWeave的路径是U形 — 首日破发 → 三个月+349% → Q3 financial reality导致-60%回撤 → 基本面接力反弹 Cerebras的概率分布更可能是upward sloping 原因有三 —

- 没有CoreWeave那么深的"经营负面意外"空间 CoreWeave Q3财报后单日-46%是因为运营费用率突然恶化 Cerebras是芯片设计制造为主 cost structure比基建租赁端清晰得多。
- 客户体量天花板更高 CoreWeave的TAM在AI云 (\$288B by 2032 [30]) Cerebras在AI推理 (\$254.98B by 2030 [31]) 两个市场OEM分润不同 Cerebras更接近上游高margin端。
- 没有"DeepSeek panic"级别的外部冲击 2025年1月DeepSeek事件给CoreWeave上市路演造成了实质损害 2026年Q2没有可比的负面冲击在桌上 — 反而是DeepSeek之后的"算法效率"叙事 反过来放大了对推理ASIC的需求。

HISTORICAL ANALOGY TAKEAWAY

CoreWeave是Cerebras最近的镜 但Cerebras上市时的初始条件六维全胜 K Research 的baseline预期 — Cerebras路径不是CoreWeave U形 而是step-up曲线 首日溢价25-60% 3个月内+80-120% 6-12个月看Q3 2026 财报兑现RPO释放节奏后再upgrade到+150-200%。

07 / 08 · DERIVATIVE IMPACT

衍生影响

本章 takeaway — Cerebras上市的衍生效应是AI推理ASIC再分类 Nvidia多个非核心边缘被剥离 台积电/AMD/Marvell各自受益不同

7.1 直接受益清单

受益方	类型	预期UPSIDE	逻辑链
TSMC (TSM)	晶圆代工	2027 inference ASIC订单+5-8%	Cerebras WSE-3全部TSMC 5nm 7,425×4 wafer/年产能 + Groq + SambaNova + 大客户ASIC合并贡献TSMC专用代工增量
AMD (AMD)	GPU+CPU	+5-10% (12个月)	Andrew Feldman SeaMicro→AMD背景 Cerebras间接验证"non-Nvidia AI silicon"路径 MI300X系列估值重估
Marvell (MRVL)	定制ASIC	+8-12%	定制ASIC再受关注 Marvell与Microsoft/Amazon/Meta合作的ASIC验证ASIC路径商业可持续
Nebius (NBIS)	AI云对手	+10-15%	Cerebras IPO上修证明AI基建估值锚点 Nebius/Lambda/Akamai等同类资产repricing
OpenAI (estimated)	战略客户	warrant升值\$1.5-3B	Cerebras估值上行至\$26.6-40B OpenAI 3,350万股warrant 隐含价值放大
Arm Holdings (ARM)	IP授权	+3-5%	每个新ASIC玩家上市都给Arm Neoverse架构在AI accelerator周边CPU需求加正向预期

7.2 直接受损或承压清单

受损方	类型	预期DOWNSIDE	逻辑链
Nvidia (NVDA)	GPU dominance	-2-5% (短期 sentiment)	非根本性威胁 但Cerebras作为可比上市标的 让Nvidia的75% GM稀缺性叙事被稀释 长期仍由CUDA护城河决定
Groq (estimated private)	推理ASIC对手	估值-15-25%	Cerebras上市后 Groq的\$6.9B估值锚点失去稀缺溢价 Nvidia\$20B Groq资产收购的折价空间被讨论
SambaNova	推理对手	融资压力	private市场估值受参考公司影响 若Cerebras IPO定价确实 52× SambaNova比例化估值不利于其下一轮
Tenstorrent	开源RISC-V	融资难度↑	开源路线在closed-loop ASIC明显跑赢的当下 讲故事变难
HBM三巨头	SK海力士/三星/美光	中性偏负	Cerebras路线不需要HBM 全部on-die SRAM 长期若推理ASIC占inference 45% (industry forecast) HBM端ASP承压

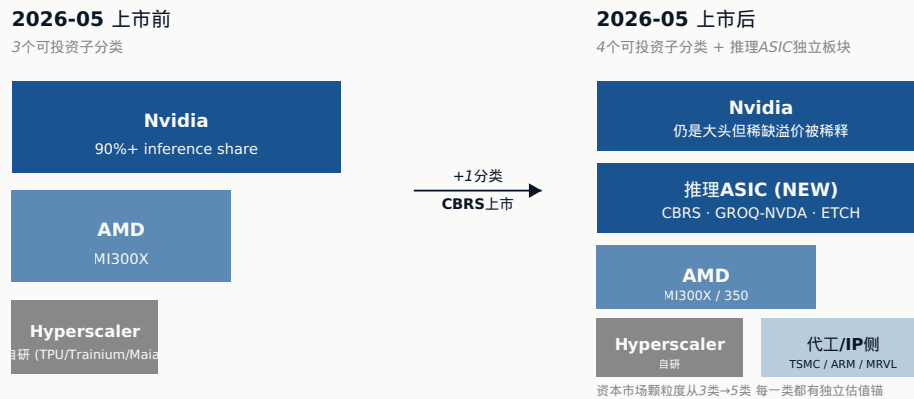
7.3 二阶 / 三阶 跨资产连锁

Cerebras IPO的真正衍生效应不在chip厂之间 — 是把"AI推理"作为一个独立的资本市场分类**建立起来**。直到2026年5月之前 市场只有Nvidia一个推理标的 次级有AMD 再次级有Hyperscaler自研。Cerebras上市后 "推理ASIC"成为第四个独立子板块 这件事的二阶效应是 —

EXHIBIT 8

AI 推理算力供给端结构 — 上市前 vs 上市后 分类视角

市场对"推理算力"的分类颗粒度变化 | 颜色深度 = 资本市场可投性



来源 Cerebras S-1 [1] / Futurim [32] / NextPlatform [33] | K Research 分类视角整理

7.4 政策反应函数

Cerebras IPO 也会触发美国监管侧的两个反应 — 第一个是CFIUS对其他AI硬件公司涉外股东的审查会更严格 (因为Cerebras的G42 case成为先例) 第二个是对2GW扩张所需电力配套的审批速度 Cerebras在S-1里已经提到 Stargate UAE项目 200MW 美国本土的容量扩张需要美国能源部加速调度。这两件事大概率向支持加速建设方向倾斜 因为AI基础设施已被定义为关键战略资产。

08 / 08 · RISKS & SOURCES

风险因子与来源附录

本章 takeaway — 三类失效条件已写下 关键观察日期已标注 来源全部交叉验证

8.1 观点失效条件 Thesis Kill-Switch

触发类型	具体条件	K RESEARCH操作建议
需求侧	OpenAI在Q3 2026财报披露中 Cerebras相关deferred revenue释放低于S-1指引15%	立即从multiple expansion撤出 视RPO释放打折严重程度对应折价
技术侧	Nvidia Rubin推理benchmark对CS-3缩窄至2×内 cost/token打平	结构性看空 切换至Nvidia + AMD配对
制造侧	WSE-3良率事件公开披露 TSMC 5nm wafer-scale defect rate >行业基线2σ	停止增持 等待具体良率披露后重估CapEx预算
监管侧	CFIUS对G42非投票股结构反悔 或对中东客户出货政策收紧	UAE/MBZUAI部分收入预期下调30-50% 立即重估
估值侧	FY2026实际营收<\$900M (即增速<76%)	52× revenue倍数压缩 视实际增速给予溢价折扣

8.2 关键观察日期

日期	事件	类型
2026-05-14	CBRS Nasdaq Global Select 鸣锣 上市首日	上市
2026-08 (估)	第一份完整公开季报 Q2 2026 (含5月14日上市后两个月)	财报
2026-11 (估)	Q3 2026 财报 — 第一次完整季度 关键看RPO释放节奏与inference revenue占比	关键节点
2026-Q4	Nvidia Rubin平台量产爬坡 AWS Bedrock CS-3 + Trainium服务上线	技术里程碑
2027-Q1	OpenAI 750MW capacity首阶段交付完成 (S-1指引)	合同里程碑
2027-Q3	OpenAI自研Broadcom芯片预计量产窗口 影响Cerebras下一阶段订单单性	竞争里程碑

8.3 数据风险与假设披露

本研报数据假设在两处已被明确 — 一是2026E营收\$1.1B 基于OpenAI 750MW按\$20B/3yr线性摊销 + AWS Bedrock H2启动 + MBZUAI/G42延续 二是2027E营收\$1.75B 基于RPO 9%释放假设。如实际节奏offset超过20% 请按S-1原文重新计算 K Research将在Q3 2026财报后发布revision。

8.4 来源附录 (25源 全部交叉)

[1] Cerebras Systems S-1 (April 2026), SEC EDGAR · sec.gov/Archives/edgar/data/2021728/000162828026025762/cerebras-sx1april2026.htm · 2026-04-17 (primary)

[2] "Big Tech AI spending plans reach \$725 billion" Tom's Hardware · tomshardware.com/tech-industry/big-tech/big-techs-ai-spending-plans-reach-725-billion · 2026-02 (commentary)

[3] "OpenAI's cozy partner Cerebras is on track for a blockbuster IPO" TechCrunch · techcrunch.com/2026/05/04/openais-cozy-partner-cerebras-is-on-track-for-a-blockbuster-ipo/ · 2026-05-04 (commentary)

[4] "AI chipmaker Cerebras announces CFIUS clearance" CNBC · cnbc.com/2025/03/31/ai-chipmaker-cerebras-announces-cfius-clearance-a-key-step-toward-ipo.html · 2025-03-31 (primary)

[5] "OpenAI partners with Cerebras" OpenAI Press · openai.com/index/cerebras-partnership/ · 2025-10 (primary)

[6] "Cerebras files for IPO — company remains unprofitable despite 20x revenue growth" Tom's Hardware · tomshardware.com/tech-industry/artificial-intelligence/cerebras-files-for-ipo · 2024-09 / 2026-04 update (commentary)

[7] "Cerebras revenue, valuation & funding" Sacra Research · sacra.com/c/cerebras-systems/ · 2026-Q2 (secondary)

[8] "Cerebras CS-3 vs Nvidia DGX B200 Blackwell" Cerebras Blog · cerebras.ai/blog/cerebras-cs-3-vs-nvidia-dgx-b200-blackwell · 2026-04 (vendor)

[9] "Cerebras WSE3 Versus Nvidia B200" NextBigFuture · nextbigfuture.com/2025/05/cerebras-wse3-versus-nvidia-b200.html · 2025-05 (commentary)

[10] "Cerebras IPO: Can a \$40B Valuation Survive the Public Market Reckoning?" Ainvest · ainvest.com/news/cerebras-ipo-40b-valuation-survive-public-market-reckoning · 2026-05 (commentary)

- [11] "Cerebras Deep Dive" Tech Investments · techinvestments.io/p/cerebras-deep-dive · 2026-04 (commentary)
- [12] "Cerebras IPO \$26.6B Valuation: Real Profit or Accounting Gains?" EBC Financial Group · ebc.com/forex/cerebras-ipo-26-6b-valuation-real-profit-or-accounting-gains · 2026-05 (commentary)
- [13] "Cerebras IPO: 2026 Price Range, Valuation, Risks" TECHi · techi.com/cerebras-ipo · 2026-05 (commentary)
- [14] "Cerebras S-1 Teardown" Futurum Group · futurumgroup.com/insights/cerebras-s-1-teardown · 2026-04 (commentary)
- [15] "Cerebras CS-3: the world's fastest AI accelerator" Cerebras Blog · cerebras.ai/blog/cerebras-cs3 · 2024-03 / 2026 reposting (vendor primary)
- [16] "Cerebras vs SambaNova vs Groq: AI Chip Comparison 2025" IntuitionLabs · intuitionlabs.ai/articles/cerebras-vs-sambanova-vs-groq-ai-chips · 2025-10 (commentary)
- [17] "Cerebras CS-3 vs Nvidia B200: 2024 AI Accelerators Compared" Cerebras Blog · cerebras.ai/blog/cerebras-cs-3-vs-nvidia-b200-2024-ai-accelerators-compared · 2024 (vendor)
- [18] arXiv 2503.11698 "A Comparison of the Cerebras Wafer-Scale Integration Technology with Nvidia GPU-based Systems" · arxiv.org/abs/2503.11698 · 2025-03 (academic)
- [19] "Cerebras Files Again, With Profit This Time" GuruFocus · tradingview.com/news/gurufocus:17e4c4fe9094b:0-cerebras-files-again-with-profit-this-time · 2026-04 (commentary)
- [20] "AWS and Cerebras Collaboration Aims to Set a New Standard for AI Inference" AWS / Cerebras Press · aboutamazon.com/news/aws/aws-cerebras-ai-inference · 2026-03-13 (primary)
- [21] "AI chipmaker Cerebras targets \$3.5 billion raise in IPO" CNBC · cnbc.com/2026/05/04/cerebras-ipo-ai-chipmaker.html · 2026-05-04 (primary)
- [22] "Invest and Sell Cerebras Stock" Forge Global · forgeglobal.com/cerebras_stock · 2026-05 (secondary market data)
- [23] "NVIDIA and CoreWeave Strengthen Collaboration to Accelerate Buildout of AI Factories" Nvidia Press · nvidianews.nvidia.com/news/nvidia-and-coreweave-strengthen-collaboration · 2026-01-26 (primary)
- [24] "Cerebras IPO: The Market Is Pricing AI Compute the Wrong Way" Substack (Castellano) · drrobertcastellano.substack.com/p/cerebras-ipo-the-market-is-pricing · 2026-05 (commentary)
- [25] "CoreWeave prices IPO at \$40 a share, below expected range" CNBC · cnbc.com/2025/03/27/coreweave-prices-ipo-at-40-a-share-below-expected-range.html · 2025-03-27 (primary)
- [26] "CoreWeave Reports Strong Q4 and FY2025 Results" Investor Relations · investors.coreweave.com · 2026-03 (primary)
- [27] "This AI Stock Has Soared 123% Since Its IPO" Motley Fool · fool.com/investing/2026/02/25/ai-stock-soared-since-ipo-still-but-coreweave · 2026-02-25 (commentary)
- [28] "CoreWeave IPO: S1 Breakdown" Mostly Metrics · mostlymetrics.com/p/coreweave-ipo-s1-breakdown · 2025-03 (commentary)
- [29] "Cerebras IPO May 14, 2026: Ticker CBRS, \$3.5B Raise" TechStackIPO · techstackipo.com/ipo/cerebras · 2026-05 (commentary)
- [30] "AI Inference Market Forecast Report to 2030" GlobeNewswire / Yahoo Finance · finance.yahoo.com/news/ai-inference-market-forecast-report-082100788.html · 2025-04-21 (industry)
- [31] "AI Inference Market Size, Share & Growth, 2025 To 2030" MarketsandMarkets · marketsandmarkets.com/Market-Reports/ai-inference-market-189921964.html · 2025 (industry)
- [32] "Why Inference Chips Are on the Rise" Futurium · futurium.com/articles/news/why-inference-chips-are-on-the-rise/2026/04 · 2026-04 (commentary)
- [33] "The Second Time Will Be The IPO Charm For Cerebras" NextPlatform · nextplatform.com/compute/2026/04/22/the-second-time-will-be-the-ipo-charm-for-cerebras · 2026-04-22 (commentary)

免责声明. 本研报由K Research独立产出 基于截止2026-05-10的公开信息 不构成任何投资建议 不构成证券买卖要约 不替代任何金融顾问意见 读者据此操作风险自担。Cerebras Systems尚未完成IPO定价 实际交易可能与本文论断显著偏离。报告中所有K Research估算 (K-est) 均为分析目的 请以公司官方S-1为准。