

AI 基础设施产业研究

CASPA CEO Summit 参会手册

从数据中心系统到部件 工艺 材料与设备

研究截止 2026 年 10 月 15 日

覆盖能源与散热 机架与系统 计算与存储 网络与光互联 晶圆制造 封装与测试

23 个模块 313 条来源记录 6 张产业图

目录

导读 AI 基础设施产业全景与 CASPA 参会研究

01 从一块企业级 SSD 看懂整条产业链

总图 全景图与系统连接

01A NAND 的工艺、材料和设备怎样衔接

02 从电网到芯片：供电架构与材料

02B 热量怎样离开机柜：液冷、建设与运维

03 数据怎样移动：机内、机架、集群与跨中心网络

03B 把一条光链路拆到材料、晶圆、装配与测试

03C 网络与光互联企业地图、瓶颈与可证伪投资命题

04 从 SRAM 到磁带：AI 的内存与存储全景

04A SSD、网络存储、HDD 与磁带怎样服务 AI

04B 存储公司版图与可验证的投资问题

05 晶圆制造：把电路图变成三维器件

05B 设备公司地图：按工艺能力定位

06 材料地图：同样叫“绝缘”，实际承担不同任务

06B 材料公司分布与投资验证

07 先进封装与电测试：系统性能的另一半

07B 光子封装测试与罗博特科案例

08 计算架构与芯片设计

08B 计算与设计公司地图

09 CASPA 会议与公司对应

10 技术壁垒与投资验证

11 社区异议与深入阅读

来源附录按编号列出原始资料 链接 证据范围及对应章节

导读 AI 基础设施产业全景与 CASPA 参会研究

研究截止 2026 年 10 月 15 日。读者任务是理解 AI 基础设施怎样工作、怎样制造、哪些环节可能形成经济壁垒，并据此在 CASPA CEO Summit 提问。正文把系统运行关系与制造供应关系分开，避免将设备商、芯片设计商、晶圆厂、封装测试厂、系统集成商视为同一类公司。

核心判断

本报告的分析结论：AI 基础设施是相互制约的产能系统。 可交付的计算能力取决于芯片、内存、封装、网络、供电、散热、软件与运维是否同时到位；单一部件的峰值指标不能代表可售服务。NVIDIA 将 Rubin 描述为多种芯片与机架共同组成的平台，AMD 将 Helios 定义为供伙伴构建系统的参考设计，这两种组织方式都要求跨层协同。[R004][R005]

本报告用五个问题替代公司名称堆砌：工作负载要搬运多少数据；哪个物理约束先到；解决它需要哪种部件或工艺；供应商能否稳定批量交付；价值能否留在供应商的现金流里。前四项支撑技术判断，第五项决定技术优势是否可能转化成投资回报。

建议阅读顺序

1. 先读 SSD 实例，建立“终端产品—芯片—工艺—设备—材料”的共同语言。
2. 看全景图与系统连接，辨认电力、热、数据及制造供应链四类关系。
3. 选择存储、计算、光互联、能源或制造模块向下钻取；每个模块用能力与验收指标解释公司位置。
4. 最后读投资验证框架、论坛异议与参会问题。来源附录可以反查所有编号；同一事实只在主章节展开。

证据与状态约定

官方事实表示官方规范、产品资料或披露文件确实这样记载；其中性能、领先性和未来需求仍可能是厂商自述。**研究推论**表示本报告从工程约束推导的判断；**社区观点**用于发现问题，不作为客户关系、份额或财务事实；**待核实**表示公开证据不足。

产品阶段按“研究/仿真—流片—回片验证—送样—宣称量产—客户部署—可复现基准”记录。它们不是一条自动完成的证明链。网页出现产品名称，不能证明已出货；采购合同、订单、收入、现金回款分别核查。

地域表按企业主要总部或产业归属归类，台湾地区单列，实际生产地点可能跨国。表中企业是有公开产品线的代表性供给方，**不是经过份额排序的全球完整工商名录**；未公开的供应关系与客户认证不作填补。

研究方法边界

CASPA 嘉宾 Dylan Patel 所在的产业研究机构为 **SemiAnalysis**。其公开方法说明将物理容量、网络连接比例与 TCO 联系起来；本报告的分析基于公开资料，不引用付费底层模型或未公开客户份额数据库。

Seeking Alpha 的公开访谈用于提炼毛利率、存货和订单兑现的分析问题，相关目标价及荐股观点不作为本报告事实。[R002][R020][R021]

本研究覆盖 AI 基础设施的架构层级、关键工艺及代表性企业，以公开一手资料为基础，辅以论文、标准、公司技术分享和社区讨论。来源附录区分全文资料与摘要；视频和播客作为延伸阅读，技术结论以可核对的材料为依据。保密配方、未公开供应合同及缺乏可比口径的市场份额不在证据覆盖范围内。

01 从一块企业级 SSD 看懂整条产业链

先固定研究对象

以一块装进 AI 存储服务器的 NVMe SSD 为例：它把主机发来的逻辑块，转换成 NAND 闪存中的实际位置。SSD 是整机部件；NAND 是保存数据的存储介质；控制器是执行协议、调度和纠错的逻辑芯片；固件决定这些硬件如何配合。把“闪存厂”“主控厂”“SSD 品牌”“存储系统厂”都归为一个赛道，会同时错判竞争关系和利润来源。[\[S001\]](#)[\[S017\]](#)

一个可手算的容量例子。假设使用 1 Tbit 裸片，并按芯片密度的二进制口径计为 2^{40} bit，则单颗为 128 GiB，8 颗合计 1 TiB 原始容量。这里还没有扣除 SSD 预留空间、坏块替换和系统元数据，也没有规定封装数量。SSD 对外标称 TB 常采用十进制，不能混算。1 Tbit TLC 是实际产品规格；8 颗组合是教学假设。Kioxia 密度脚注明确 $1\text{ Gb} = 2^{30}\text{ bit}$ 。[\[S013\]](#)

运行时 一笔写入经过什么

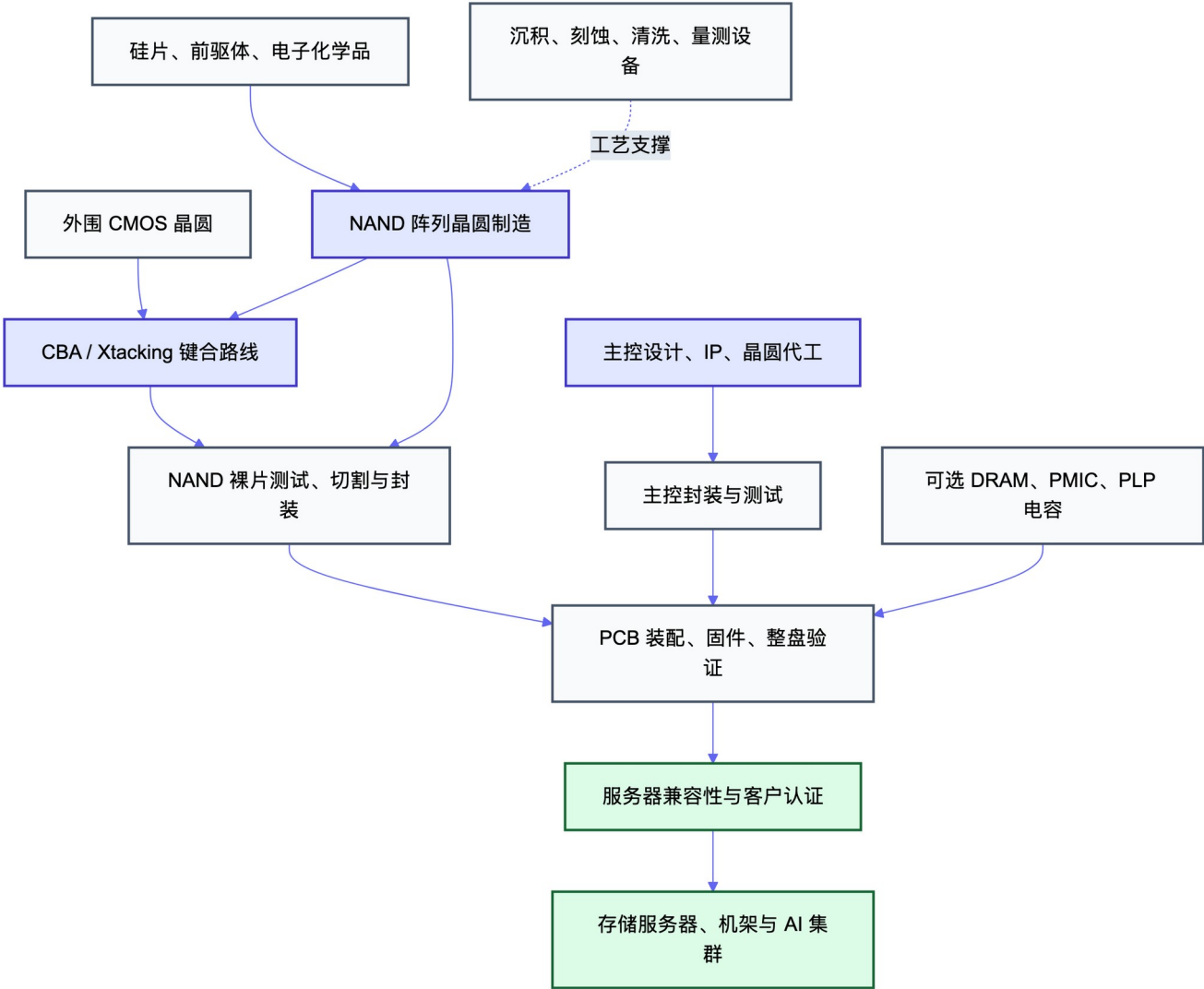
顺序	发生什么	实际需要的部件或能力
1	CPU 提交 NVMe 命令，由 PCIe 搬运数据	CPU PCIe 根端口、主板走线、连接器，必要时经过 PCIe 交换机；NVMe 是协议，PCIe 是一种传输路径。 [S020]
2	控制器找到可写位置，更新逻辑到物理映射	控制器处理器、片上 SRAM、可选外置 DRAM、FTL 固件；映射表不是用户文件系统。 [S001] [S025]
3	编码并并行写入多个 NAND 通道	ECC/LDPC、数据路径校验、NAND PHY、通道调度；并行度、纠错开销和功耗要共同设计。 [S017]
4	后台搬迁有效页，回收整块空间	垃圾回收、磨损均衡、备用块；NAND 按页写、按块擦除，因此主机写入量可能小于内部写入量。 [S018]
5	突然掉电时保存写入及关键映射状态	电压检测、保持供电电容、PMIC、固件恢复机制；保护范围必须按具体产品确认。 [S016]

读取的产业逻辑反向展开：控制器先查映射，NAND 感测电路读取电压状态，纠错引擎恢复数据，再经主机接口返回。顺序吞吐率、随机访问延迟、满盘稳态写入和掉电一致性是不同能力，单个“GB/s”数字无法替代整盘评价。[\[S001\]](#)[\[S019\]](#)

制造时 谁把“材料”变成“可交付 SSD”

下图中的实线是制造或装配关系，虚线是设备/材料支撑关系；它不是同一公司内部的必经组织架构。
NAND 与主控通常采用不同工艺，不能把一块 SSD 的所有芯片都叫作“某个纳米制程”。例如慧荣 SM2508 的逻辑主控采用台积电 6nm，NAND 自身另按存储结构和世代演进。[\[S017\]](#)

SSD 制造、装配与设备材料支撑关系



图中并列保留了阵列直接进入后续工序及先与 CMOS 键合的路线：Kioxia CBA、长江存储 Xtacking 都把外围电路和阵列分别加工再连接；其他集成方式不必复制完全相同的流程。图是基于公开结构形成的教学抽象，不是厂商保密配方。[\[S003\]](#)[\[S007\]](#)

三种“堆叠”必须分开

名称	实际增加什么	主要约束
NAND 单 die 的垂直 cell/word-line 层数	一颗裸片内部可存储的垂直位置	孔深、孔形、字线电阻、膜应力、层间干扰。 [S002] [S008]

名称	实际增加什么	主要约束
NAND 阵列与 CMOS 的晶圆键合	独立制造两种功能，再建立大量连接	表面平整度、洁净度、对准和键合缺陷；并不自动增加存储 cell 层数。 [S003] [S007]
一个封装内堆多颗 NAND die，或 HBM 内堆 DRAM die	封装内的裸片数量	裸片厚度、连接、热、封装测试；HBM 还涉及 TSV 和宽接口。 [S026] [S068]

读路线图的规则：比较良品 bit 成本，而非比较层数标题。 Kioxia 2025 年公布的第九代 512 Gb TLC 采用基于第五代技术的 120 层工艺和新 CMOS/CBA，目标是成本与性能组合；2026 年第十代 1 Tb TLC 则在 7 月宣布送样，采用更先进堆叠。不同世代名称不保证层数严格单调，更不能把送样当作收入量产。[\[S057\]](#)
[\[S013\]](#)

同一块 SSD 上的四种经济模型

- **NAND 制造商：**为晶圆厂、工艺平台、良率和每 bit 成本投入资本；层数提高可能同时增加工序、设备占用和缺陷风险，最终看每片晶圆产出多少可售良品 bit。[\[S003\]](#)[\[S008\]](#)
- **独立控制器公司：**通过逻辑芯片和固件适配不同 NAND，在纠错、调度、安全和低功耗上竞争。主控支持某 NAND，不代表客户所有固件版本与容量组合都已认证。[\[S017\]](#)
- **SSD 方案及品牌公司：**组合介质、主控、固件、PCB 与可靠性保障。企业客户购买的是可持续交付的完整盘及其工作负载表现；需要预处理和一致条件下的整盘测试。[\[S019\]](#)[\[S023\]](#)
- **设备与材料公司：**销售刻蚀/沉积/清洗等工艺能力以及持续消耗的化学品。投资验证要追到具体步骤的客户导入、扩产需求与单片消耗，而不是从 NAND 需求直接跳到某供应商收入。[\[S004\]](#)[\[S006\]](#)[\[S009\]](#)

下一节把制造过程打开到孔、薄膜和金属；第 04 章再把 SSD 放回 SRAM、DRAM、HBM、HDD 与磁带共同构成的系统。

总图 全景图与系统连接

从一个 AI 请求看整座设施

典型在线推理可拆为请求进入前端网络、CPU 执行鉴权和调度、数据或模型从存储进入主存、加速器完成计算、结果返回。模型权重、KV cache 和中间激活的驻留与搬运由运行时决定；PagedAttention、批处理和分阶段服务会改变真实的容量及带宽需求。因此，硬件分析必须带着软件的访问模式进行。[\[R013\]](#)[\[R014\]](#)

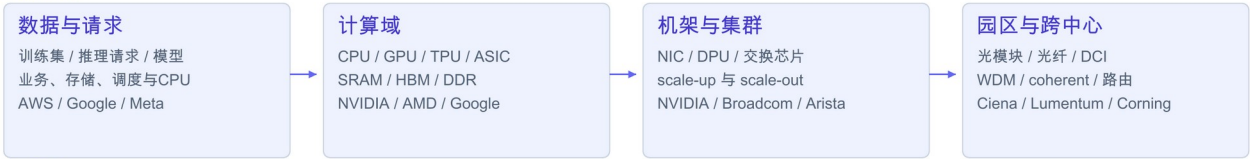
在本文框架中，**scale-up** 是一个紧耦合计算域内的加速器互联，**scale-out** 连接多个计算域；它们与“同机架/跨机架”并非永远一一对应。前端业务、后端训练/推理通信、存储网络、带外管理网络应分别画出，避免将所有交换机和光模块算成一张网。[\[N021\]](#)[\[N022\]](#)[\[N024\]](#)

图一 全产业链总图

01 AI 基础设施全产业链

按四条链阅读：数据流、电力流、热流，以及制造工序与供应关系。图中举例不表示指定客户采购关系。

数据与计算



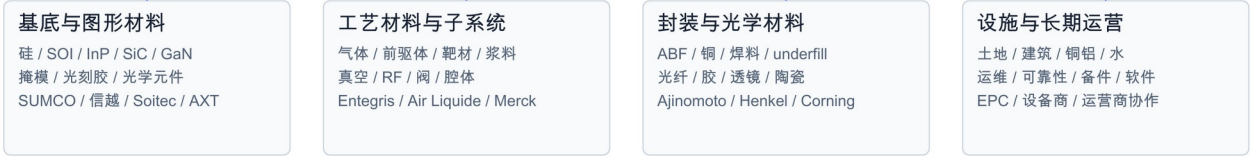
电力与热



设计与制造



材料与底座



制造链由设计驱动并依赖材料；运行链同时受可用电力、可排热量、内存容量和网络通信约束。

AI 基础设施产业研究 | 截止 2026-10-15 | 本报告原创示意图 | 产品阶段与企业证据见正文及来源附录

AI 基础设施四条链总图

图注：这是本报告对各章证据的综合架构，不是某厂商的采购 BOM。实线表示系统中的电、热或数据流，制造链箭头表示供应和工序依赖。公司举例及来源在各模块给出。纵向读法为“电网到 token”；横向读法为“设计到制造再到系统交付”。

图二 数据中心与机架内部

02 数据中心内部与机架连接

功能连接示意，省略冗余路由；CPU不必位于所有GPU数据传输的必经路径。光模块可位于不同网络层级。



AI 基础设施产业研究 | 截止 2026-10-15 | 本报告原创示意图 | 产品阶段与企业证据见正文及来源附录

数据中心及机架连接

图注：GPU 到 HBM 是封装内的短距宽接口；CPU 到内存一般是 DDR 通道；CPU/加速器/NIC 的本地连接可含 PCIe，CXL 在特定 CPU 与设备组合上增加一致性/内存语义。GPU scale-up 可以经过专用交换结构；scale-out 通过 NIC 和交换网络。此图不表示全部 GPU 内部通信都先经过 CPU。[\[R004\]](#)[\[R006\]](#)[\[N025\]](#)[\[N026\]](#)[\[N021\]](#)[\[N024\]](#)

七个层级的进入点

层级	要回答的问题	对应硬件与供应者类型	应观察的约束
园区与设施	何时可有电力和机房	公用事业、变压器、开关设备、EPC、UPS、发电机	已送电 MW、施工与许可、可靠性
机架	电与冷却怎样送到托盘	机柜、母排、PDU、power shelf、CDU、歧管、冷板	每架功率、维护窗口、漏液与保护
系统板	算力与内存如何连接	CPU、GPU、NIC/DPU、DDR、SSD、PCB、连接器	供电完整性、信号完整性、布局 and 散热

层级	要回答的问题	对应硬件与供应者类型	应观察的约束
封装	多种裸片如何组合	HBM、interposer、RDL、TSV、基板、键合与测试	带宽密度、翘曲、良率、热阻
裸片	如何实现计算或存储	逻辑、SRAM、DRAM、NAND、光子和功率器件	面积、能耗、器件与互连限制
晶圆制造	怎样重复形成目标结构	晶圆、掩模、光刻、沉积、刻蚀、CMP、检测	工艺窗口、缺陷、设备利用率
材料与子系统	工艺为何能稳定重复	高纯气体、前驱体、光刻胶、靶材、真空/RF/阀门	纯度、均匀性、颗粒、寿命与认证

本表是对正文的索引与分析分类；具体产品能力分别见能源、网络、存储、制造、材料和封测章节，不将类别共存推断为直接供货。

一套可核算的示例

以下全部是**教学假设，不是项目报价或市场预测**：一座数据中心有 10 MW IT 负载，以每机架 100 kW 折算为 100 个功率等效机架；实际计算机架数量还取决于网络、存储、CPU 与功率预留。若全年满功率运行、PUE 为 1.20，设施输入为 12 MW，年用电 105.12 GWh。电价每上涨 0.01 美元/kWh，年度电费增加约 105.12 万美元。这里 PUE 的分母是 IT 能耗，不是 GPU 单独能耗。

再假设每个计算机架 72 个加速器、每加速器 8 个 HBM 堆栈，则每架需要 576 个 HBM 堆栈；若每堆栈 12 个 DRAM die，则为 6,912 个 DRAM die，尚未计底座 die、冗余和良率损耗。封装 die 需求不能直接换算成晶圆需求，还需 die 面积、每片可切 die、KGD 良率与组装良率。这个算法展示“连接比例”怎样将系统规模传导到封装材料和设备，数字不是特定产品的已核实配置。

可售 token 的教学公式为：年度有效产出 = 实测每秒合格输出 × 可运行秒数 × 真实利用率。年度单位成本 = 年化资本成本、电费、运维、网络、软件与融资成本之和 / 有效产出。合格输出必须满足模型质量、延迟和可用性要求；不能把电源铭牌、理论 FLOPS 或标称端口速率直接代入产出。[\[R012\]](#)[\[R020\]](#)

01A NAND 的工艺、材料和设备怎样衔接

先理解 cell 电荷改变可读取的状态

现代电荷捕获型 3D NAND 的一条垂直 string，沿沟道排列多个受字线控制的 cell。读操作辨别阈值状态，编程/擦除改变存储电荷。TLC 是每 cell 3 bit，QLC 是 4 bit；对应更多可区分状态，不是把同一层简单切成更多独立小芯片。电荷保持、擦写退化与 cell 间干扰因此成为材料和算法共同承担的问题。[\[S011\]](#)[\[S012\]](#)

公开器件研究把 tunnel/trap/blocking/channel 层分别作为可靠性优化对象：薄膜缺陷会影响漏电、俘获和释放；局部几何与电场会改变编程表现。必须区分论文器件改善、厂商工艺导入和批量可靠性认证，不能从一篇新材料论文直接推导量产替代。[\[S012\]](#)[\[S058\]](#)

典型 replacement-gate 路线的工序图谱

本表概括一种有代表性的路线，各家材料顺序、腔室与分段策略不同；每一行是设备厂可以争取的工艺位置，不是某个客户已公开的完整 BOM。

工序	要解决的问题	设备与材料能力	壁垒及应检查的量
交替堆积绝缘/牺牲层	形成高而均匀的多层结构	CVD/PECVD；SiO ₂ 与 SiN 等膜层；Lam 公布有 Ox/SiN 层控与应力管理方案。 [S005]	厚度与成分跨片/跨层一致性、沉积速率、晶圆弯曲、颗粒。 [S008]
光刻、硬掩模与通孔刻蚀	在高堆叠中打出细而直的 channel hole	高选择比硬掩模、等离子体介质刻蚀、低温化学与脉冲控制；Lam、TEL 是有公开技术披露的代表。 [S004] [S005]	底部能否贯通、侧壁形貌、孔径与扭曲、掩模损耗、处理时间。
形成沟道和存储介质	让孔顶到底的 cell 尽可能一致	共形薄膜沉积；隧穿、俘获、阻挡层及 poly-Si channel；更深结构要求前驱体充分到达底部。 [S002] [S008]	台阶覆盖、界面缺陷、污染、阈值分布、保持与耐久。
阶梯和 slit 加工	将不同高度的字线分别接出来并提供工艺通道	阶梯图形化、轮廓控制、填充；Applied 公布 Sym3 阶梯刻蚀与 PE-HARP 介质填充。 [S002]	阶梯占地、接触对准、跨片 CD 一致性、空洞与应力。
移除牺牲 SiN	在保持氧化层和脆弱结构的同时留下字线空间	高选择性湿法/化学刻蚀；传统热磷酸体系及配方优化。 [S009]	深处去除充分性、SiO ₂ 再沉积、残留、槽液寿命和缺陷。
字线金属及连接填充	在狭窄空间获得低阻且连续的导体	W 沉积/填充；Mo 字线相关沉积与选择性刻蚀也在推进；并非所有产品均使用同一种金属。 [S005] [S006]	薄膜电阻、空洞、夹气、晶圆应力、相邻字线短路。

工序	要解决的问题	设备与材料能力	壁垒及应检查的量
阵列与外围集成	让外围驱动、感测和 I/O 不占用过多阵列平面	CuA 或分开制造后键合的 CBA/Xtacking 等；键合前需极平整表面和精密对准。 [S003][S007]	键合缺陷、接点开短路、两类晶圆的良率匹配和热预算。
封装与产品化	把裸片容量变成可以交付的接口与可靠性	多 die 封装、控制器固件、纠错和整盘验证。[S068][S017][S019]	每 bit 成本、封装良率、满盘稳态、寿命、客户资格认证。

高深宽比不是“能挖多深”一个指标。 TEL 在 2023 年公布针对超过 400 层结构的 10 μm 深孔刻蚀技术，披露的实验时间为 33 分钟；这是技术展示的结构及条件，不能当作全厂每片产能或所有厂商量产水平。投资比较至少要同时取得孔形合格率、wafer/hour、腔室匹配、耗材与维护周期。[S004]

低温刻蚀的作用来自反应与轮廓控制的变化。 Lam 的公开路线把 cryogenic etch 与字线 pitch、金属化、背面应力补偿并列，显示继续提高堆叠需要多种步骤共同改善。报告采用其产品能力描述，不把“所有领先客户采用”等营销表述直接折算为已验证市场份额。[S005]

还有一种容易混淆的 stack：分段沟道整合。 Samsung 对 V10 的说明把 3-stack HARC Merge 与阵列/外围 wafer bonding 列为两项独立创新；前者处理高堆叠沟道/工艺整合，后者连接不同功能晶圆。其描述还把孔形、共形薄膜、清洗均匀性和膜应力列为共同约束。这是供应商工艺解释，不等于公开了可复制的完整制程或客户产量。[S072]

“隔离电子的材料”到底是哪一种

位置/材料	功能	最容易产生的误解
相邻字线间的氧化物绝缘层	隔开不同导电层，控制结构和电耦合	它与 cell 内的隧穿氧化层不是同一工艺位置。[S002][S011]
Tunnel dielectric，隧穿介质，常见 SiO ₂ /复合氧化体系	在编程/擦除条件下允许受控电荷转移，平时兼顾保持	“绝缘”不等于所有电场下绝无电荷通过；厚度、缺陷和能带工程是核心。[S012]
Charge-trap SiN，电荷捕获氮化硅	提供俘获电荷的状态，承担存储功能	它不是一层只负责分隔的普通绝缘胶，也不是之后必然挖掉的牺牲层。[S011]
Blocking dielectric，阻挡介质，包括氧化物及高 k 组合	限制朝控制栅方向的非期望电荷交换、调整栅耦合	高 k 材料出现不代表器件就是 DRAM 电容；要看所在位置和目标。[S002][S058]
Sacrificial SiN，牺牲氮化硅	多层模板中后续被选择性去掉，给替换字线留空间	与 charge-trap SiN 化学类别相近，工艺使命相反。[S009]

位置/材料	功能	最容易产生的误解
TiN/W，或推进中的 Mo 金属方案	电极、阻挡/衬层、字线与接触等导电任务	金属不是“层间电子隔离材料”；导电、扩散阻挡和电荷存储需分开。 [S002] [S006]
Poly-Si 沟道及孔芯填充	沟道负责 string 方向载流；孔芯另有绝缘填充	垂直孔不是挖完后保持空心；各同心层有不同功能。 [S002] [S058]

SiN 在此指氮化硅；材料名称必须同时对应化学式、器件位置与工艺步骤。[\[S002\]](#)[\[S011\]](#)

从“有这种化学品”到“能进入量产”的差距

薄膜材料需要可交付的 ALD/CVD 前驱体，前驱体又需要稳定的纯化、储运和供料系统。Merck 的公开薄膜产品资料分别列有有机硅、有机金属、介质和相关供料研发能力；这些目录能够证明供应商有对应材料平台，不能证明某个特定 NAND 客户已经采用其某一配方。[\[S010\]](#)

ALD 的优势是表面自限反应形成的厚度控制和共形性，但“原子层级”不自动保证高吞吐：不同反应物、温度、曝露与清扫条件会影响生产率。ASM 的公开应用包括高 k 栅、MIM 电容、介质及高深宽比结构，因此设备评估要按具体薄膜和步骤比较，不能仅比较“都有 ALD”。[\[S031\]](#)

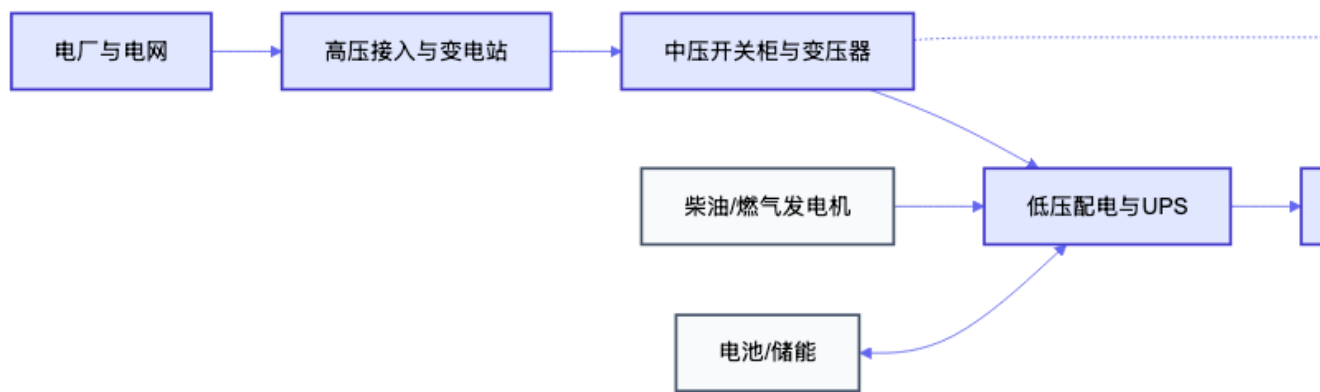
热磷酸去除牺牲 SiN 的案例说明了材料壁垒的真实形态：选择比好仍可能因槽液中的反应产物再沉积而增缺陷；延长槽液寿命又可能恶化这一问题。可投资的差异来自在客户结构上同时控制选择性、缺陷、产能与使用寿命，必须用客户验证数据判定。[\[S009\]](#)

研究判断：3D NAND 扩层的竞争最终是一张约束网络——更高堆叠提高孔深，孔深影响刻蚀/沉积时间；更薄字线降低高度，却放大电阻与金属整合难题；更薄晶圆和更高膜应力又压迫后续键合。突破一个步骤可以释放其他步骤的价值，但也可能把瓶颈转移给相邻工序。[\[S005\]](#)[\[S006\]](#)[\[S008\]](#)

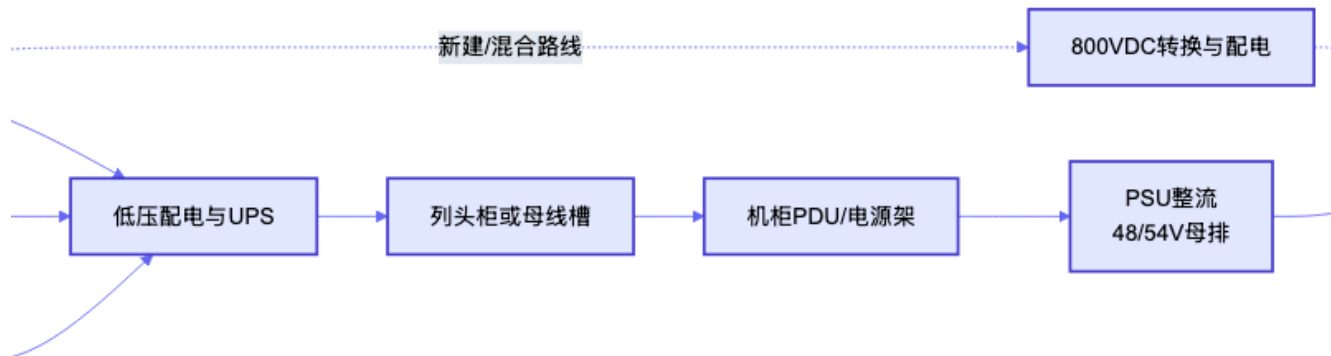
02 从电网到芯片 供电架构与材料

范围：截至 2026-10-15。下图是功能拓扑；项目可以选择不同冗余、交流/直流架构，并非所有方框都串联。机柜内的直流母排不能与园区中压母线混为一谈。[N004][N013][N019]

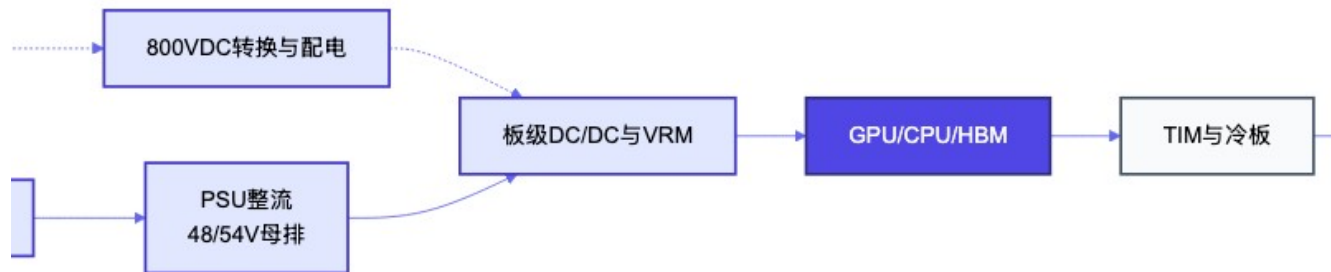
从电网到芯片的功能供电拓扑（从左至右第 1/4 段，相邻区域重叠）



从电网到芯片的功能供电拓扑（从左至右第 2/4 段，相邻区域重叠）



从电网到芯片的功能供电拓扑（从左至右第 3/4 段，相邻区域重叠）



从电网到芯片的功能供电拓扑（从左至右第 4/4 段，相邻区域重叠）



图中电力路径依据 OCP、华为产品手册；直流演进依据 NVIDIA；热路径依据 OCP 冷板要求与 CoolIT。每条箭头表达能量/热量交接，不代表供应商独家关系。[\[N004\]](#)[\[N009\]](#)[\[N010\]](#)[\[N002\]](#)[\[N019\]](#)

首先区分四种“容量”

电源容量、可接入电网容量、IT 可用容量、实际利用率是四个变量。IEA 2026 年更新将全球数据中心用电量由 2025 年约 485 TWh 预测至 2030 年约 950 TWh；这是全体数据中心的用电预测，不是 AI GPU 市场规模。当地输变电、设备、审批和融资约束会改变项目兑现时间，不能用宣布的 GW 直接换算当期收入。[\[N001\]](#)

判断一个园区，先核实并网协议及送电里程碑，再核实变压器/开关柜/冷却交付，最后核实安装的 IT 负载与可售算力。这是由工程依赖推导的研究顺序：已买 GPU 不等于已上电；已上电不等于可持续满载运行。[\[N001\]](#)[\[N012\]](#)[\[N013\]](#)

每一层究竟解决什么

层级	功能与关键部件	最应该问的参数	证据
电网连接	电网公司、输电线路、变电站、保护继电器、计量；确定电力何时抵达	firm capacity、并网条件、可中断负载条款、短路容量	[N001] [N013]
中压变配电	中压开关柜隔离/保护；变压器改变电压；母线与电缆输送功率	额定/短路电流、绝缘等级、温升、选择性保护、交期	[N012] [N013]
备用发电	发动机、发电机、控制器、ATS、燃料和排气系统；电网异常时承担负载	standby 与 prime rating、启动/加载曲线、燃料连续性、排放许可	[N016]
UPS/电池	整流、逆变、旁路、电池、控制；桥接切换和电能质量扰动	满载/部分负载效率、瞬态响应、可维护性、电池寿命	[N019]
BESS/BBU	园区储能与机柜备电分属不同尺度；储能可缓冲，不能凭“有电池”推定可长期离网	能量 kWh/MWh、功率 kW/MW、持续时间、充放电策略	[N004] [N006]

层级	功能与关键部件	最应该问的参数	证据
列头 柜/母线 槽	列头柜集中分路；母线槽加 tap-off box 向机架供电；两者可替代或组合	分路数量、插接能力、计量、故障分区、铜/铝截面	[N014] [N060]
rack PDU	把输入电力分给设备插座；智能 PDU 增加遥测/控制；不是交换机	插座/回路额定值、相平衡、冗余输入、远程计量	[N015]
power shelf/PS U	power shelf 是承载多个 PSU/整流模块的电源架；OCP 机架经公共直流母排给计算托盘供电	热插拔、N+1、母排温升、连接器、监控故障容错	[N004]
板级供电	中间母线转换器 IBC、负载点 POL/VRM 把输入降为芯片所需电压	输出电流、动态压降、效率曲线、功率密度、散热	[N006] [N057]

“机架”通常强调承重安装结构，“机柜”还包括围护、门、线缆和散热组织；采购中两词会交叉使用。OCP Open Rack、传统 19 英寸机柜、厂商整机架方案的尺寸、电源和接口需要分别核对，不能把一张标准机柜图套在所有 AI 系统上。[N004][N015]

48 54V 到 800VDC 改变在哪里

低压直流的优点是贴近 IT 设备，但同样功率下电压越低，电流越大；由 $P=VI$ 、线损 I^2R 可推得导体、接点温升和布线压力上升。提高分配电压可以减小同功率电流，却带来直流开断、绝缘、故障能量与检修问题。这是架构权衡，不是单个功率晶体管的替换。[N002][N005]

OCP ORv3 的代表架构是机架共享 48V 电源架和母排；某些产品体系称 54V 系统，应按额定工作范围和接口查表，不能把名称差异直接当作不兼容证据。OCP v1.0 电源架文件具体规定多个整流器、N+1 和热插拔，也显示部分方案可由设施 tap-box 直接进入电源架，不必再串一个机柜 PDU。[N004][N006]

当前成熟度边界：NVIDIA 2026-08-11 文件把 MGX 兼容 800VDC power rack 写作“2026 年下半年到来”，属于接入既有 AC 设施的混合路线；集中 row power center 的预期可用时间为 2027 年。园区直接中压 AC→800VDC 与整套固态变压器路线仍需逐项目核验。生态合作、参考设计、样机、可交付产品、规模运营必须分别记录。[N002]

Infineon 的系统图把 SiC 高压转换、GaN 高变比 DC/DC、Si/GaN 低压转换及最终 VRM 分开；因此“800V 受益”不等于每一颗开关都改用 SiC，也不等于 SiC 晶圆需求与机柜功率线性同步。具体 BOM 由电压、开关频率、拓扑、可靠性与成本共同决定。[N005][N006]

再向材料和工艺下钻

部件	材料/加工能力	壁垒落点	证据
功率半导体	Si、SiC、GaN 对应不同功率转换位置；器件设计、晶圆工艺、封装和驱动协同	高温/高压可靠性、开关损耗、良率、客户认证	[N005]
电感/变压器/滤波	磁芯、绕组、功率电感、脉冲变压器、共模扼流圈；不是只有晶体管才能形成技术含量	饱和电流、磁损、温升、寄生参数、体积和供应稳定	[N058]
铜/铝导体和连接器	电缆、母线、表面处理、接点和绝缘构成连续导电路径	接触电阻、温升、机械插拔、装配误差、安规	[N014] [N060]
电源模块封装	功率级、控制 IC、电感靠近负载；包含板顶、板底或垂直供电布局	电流路径缩短、寄生电感、可返修性、热机械耦合	[N057]
电气绝缘/阻燃	塑料、绝缘片、薄膜、陶瓷等与功率部件共同定义耐压和阻燃	材料本体性能与装配后的爬电距离、热老化同时达标	[N004] [N011]

上述“壁垒落点”是本文的尽调指标，不是对任何公司的已验证独家能力判定。尤其磁性材料、连接器、绝缘与控制软件会成为高密度供电的共同限制；只统计 SiC/GaN 器件市场会漏掉系统 BOM。[N004][N006]
[N058]

企业位置与投资检查

供电不能只按公司国籍分赛道：Eaton 的中压开关柜扩产公告是设备制造和交付证据；华为/科华是电源系统集成；Delta 同时覆盖电源与冷却；Infineon/MPS 提供器件与模块；TDK 是磁性/无源元件。这些公司的订单、利润、认证周期和收入确认方式不同。[N012][N018][N019][N020][N005][N057][N058]

推论：优先寻找“客户必须验收、失效代价高、生产扩容需要时间”的位置，再看市场已经给了多少估值。Eaton 2026-04-08 公告计划新厂在 2027 H1 投产，能证明扩产时间差，不能单独证明当前缺货价格或未来毛利率。现场应追问：积压订单中多少已锁定送电项目、取消条款、交付验收、预收款、扩产后价格和营运资本。[N012]

02B 热量怎样离开机柜 液冷、建设与运维

三条环路不能混

热路径依次是：芯片/封装→热界面材料 TIM→冷板→技术冷却环路 TCS→CDU 换热器→设施水环路 FWS→最终环境。液到液 CDU 通常隔离服务器侧与设施侧水质/压力，泵、过滤器、阀、传感器与控制形成完整产品；液到空气 CDU 则把热交给机房空气，不能凭“用了液冷”推定建筑端已经不需空调。[\[N009\]](#)[\[N011\]](#)
[\[N017\]](#)

冷板直冷和浸没冷却是不同路线：前者将液体封闭在冷板/管路内；后者把适配的器件浸于介电液体。两条路线的材料兼容性、维护、连接器、热设计和保修约束不同。采购决策需要验证整机配置，不应由单一介电液、泵或冷板的性能推导整套系统可用。[\[N010\]](#)[\[N017\]](#)

高温供水能增加自然冷却机会，但不能简单宣称“温水液冷不耗水”。ASHRAE 的 W17/W27/W32/W40/W45/W+ 是供水温度能力分类；干冷器、冷却塔、冷机的选择仍受气候、湿球/干球、温差、设备允许温度和可用空间影响。[\[N007\]](#)

液冷的逐层 BOM 和验收问题

下表部件功能有资料支撑；验收栏是本文按失效机制建立的调查模板，不代表已经获得任一供应商的验收结果。

位置	部件、材料与制造能力	必须核实的验收内容	证据
芯片接触面	TIM 填补微小接触空隙；导热片、膏/相变类材料须按实际配置选择	接触热阻、压力、老化、泵出、返修流程；不能只看材料导热率	[N011]
冷板	铜/铝等导热结构、流道、密封、连接口；设计与加工共同决定热阻和压降	指定流量/温度下热阻、均流、压降、污染容忍度、泄漏和耐压	[N010] [N055]
机架管路	供回水歧管 manifold、软管、快接 QD、盲插、阀和接头	插拔寿命、残液/泄漏、液体兼容性、误插防护、热插拔维护	[N010] [N017]
CDU	泵、板式换热器、过滤器、传感器、控制、冗余电源及接口	冗余切换、流量/压差控制、传感器故障、过滤与维护、告警联动	[N008] [N009]
二次液体	水/乙二醇等指定配方，添加剂与所有金属/聚合物接液材料协同	pH、电导率、颗粒、腐蚀、微生物；DI 水不能当作永远绝缘	[N008] [N010]
建筑端	水泵、水管、阀、蓄冷、冷机、冷却塔、干冷器；与 TCS 边界清楚	最坏气象、部分负载、泵/冷机故障、可维修与剩余风冷负载	[N007] [N013]

位置	部件、材料与制造能力	必须核实的验收内容	证据
监测与 控制	漏液、露点、温度、流量、压差、设备/BMS/ DCIM 接口	报警、隔离、关阀、降功率、停机动作的 先后和演练	[N008] [N013]

本文推论：冷板的微细流道越依赖高传热，颗粒、压降、材料腐蚀和维护策略就越不能分开评估。设备商声称的 MW 通常要附设施供水温度、技术侧进出口温差、流量和冗余条件；不同条件的铭牌容量不能直接排名。[N008][N010][N054]

CoolIT 2026 年 CHx2000 产品说明覆盖泵、过滤、传感器和控制；Vertiv XDU1350 规划文件列出液体质量监测选项、过滤和通讯接口。两者都说明 CDU 的价值不只是“一个换热器”。厂商的最高密度、最低功耗等相对宣传仍须在同工况下验证。[N008][N054]

用一个小算例看工程约束

假设 1 MW 热量全部进入水环路，水比热取约 4.18 kJ/(kg·K)，供回水温差假定 10K。按 $Q=\dot{m}c_p\Delta T$ ，所需流量约 23.9 kg/s，即约 86 m³/h；这只是热平衡示例，未计液体配方、旁通、残余风冷与冗余。泵选型还要结合整条管路压降，所以“1 MW CDU”与“1 MW GPU”不是自动一对一。[N009][N010]（本文示例计算）

进一步假设一台机架 100 kW、热捕获率 80%，则还有约 20 kW 留给空气路径。这个算例用于检查容量边界，不代表任何具体 NVIDIA 产品。重要问题是每一类发热件有没有被冷却、停泵多久触发降频、UPS 是否也保障泵与控制，而非只问冷板覆盖了多少芯片。[N010][N013]（本文示例计算）

建设依赖图 土地不是机房交付

阶段	交付物	主要供应/服务	易漏掉的约束	证据
选址/许可	电力接入、土地、用水、光纤路由、施工/环境条件	开发商、电网、地方政府、工程顾问	电源与光缆“路由不同”是否实际共沟；社区/许可时间	[N001][N013]
土建/结构	建筑、楼板/基础、抗震、屋面、装卸、吊装	总包、结构/岩土工程、钢材/混凝土	满液机柜重量、冷却设备荷载、设备更换通道	[N092]
机电 MEP	变电、配电、制冷、管线、消防、监控	电气/HVAC/消防承包商、设备厂	并行系统之间的接口和联锁	[N013][N019]
布线/网络	光纤、光配架、连接器、铜缆、桥架、测试	光缆/连接器厂、网络集成商	弯曲半径、清洁、极性、插损预算、标识和备件	[N045][N050] [N060]
综合调试	假负载、全链路切换、网络与冷热联动验证	commissioning 团队、设备厂、运维方	单机测试通过是否覆盖全厂失电/泵失效/恢复顺序	[N013][N018]
持续运营	容量、能效、可用性、维护与变更	运营商、DCIM/BMS/EMS、现场服务	负载改变后的参数重整、维护窗口和报警质量	[N008][N013]

预制电力模块、模块化机房把部分装配与调试前移到工厂，有利于减少现场接口；它仍需要实际接电、消防及本地审批。科华 2026-08 文件列出变压器/UPS/配电整合 Power Pod 和项目案例；这些可证明系统范围，不代表其宣传的节省比例可直接迁移到其他国家项目。[N018]

PUE WUE 与“算力收入”的关系

PUE 的分母是 IT 能耗，分子是设施总能耗；WUE 关注用水。PUE 下降不自动意味着总电量、总用水或碳排放下降：IT 规模可以更快上升。比较项目需要注明计量边界、年度/瞬时值、气候和负载率，并把硬件利用率、任务吞吐和可用性一起看。[N007][N001]

推论：一个较低 PUE 但频繁因温度/网络降频的站点，可能有更高单位有效 token 成本；一个电价低却没有准时送电的项目，也没有可售的 GPU 小时。因此投资模型至少分列 IT MW、PUE、电价、可售利用率、设备可用率、资本成本和收入确认，不能仅用“每 MW 建设成本”。[N001][N013]

地区地图与可检验壁垒

地区/企业	已公开可核实位置	应进一步索取的能力证明	证据
美国：Eaton；Vertiv；Boyd	配电/UPS；CDU；冷板/TIM；角色不同	指定架构认证、产能、交付、系统效率与服务覆盖	[N012][N008] [N011][N055]
欧洲：Schneider Electric；Siemens；Infineon	母线/配电；设施自动化/电气/消防；功率器件	本地安规、保护配合、800V 产品状态、维修与备件	[N014][N013] [N005]
中国大陆：华为数字能源；科华；英维克	PowerPOD/UPS；预制电力模块；液冷全链路产品	接液材料清单、部署案例、整机厂验证、海外认证	[N019][N018] [N017]
中国台湾：台达 Delta	电源/磁性组件与液冷模块、CDU	量产范围、主要客户集中度、功率/冷却边界	[N020]
日本：TDK；古河电工	电感/磁性/保护元件；光纤/冷却相关产品	材料规格、工厂产能与客户 qualification	[N058][N051]
加拿大：CoolIT	数据中心液冷 CDU 与系统	相同温差/流量下容量、现场服务、故障恢复	[N009][N054]

韩国：LS ELECTRIC 已公开数据中心 Remote Power Panel 配电资料；Hyosung 展示变压器、开关设备、SST/HVDC/STATCOM 等电网方案。展品、产品目录与已取得客户订单是不同证据，上述企业的 AI 数据中心收入占比及韩国电池企业的具体项目供货仍缺乏充分公开证据。[N071][N072]

Reddit r/datacenter 的改造讨论中，用户反复提到预留管线空间、保留残余房间制冷、按服务器厂商核实二次液体。这些属于有工程细节、可与手册交叉检查的经验线索；作者身份、项目规模和长期结果均未验证，不能用于证明市场份额或投资回报。[N052]

03 数据怎样移动 机内、机架、集群与跨中心网络

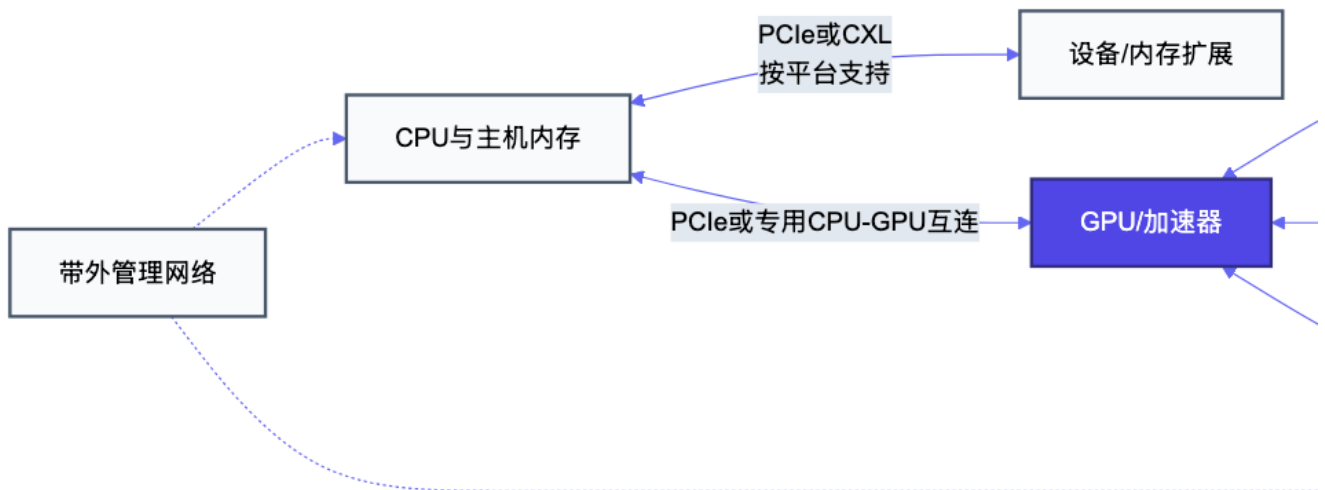
先纠正三个容易混的名词

CPU 执行通用计算与控制；**DPU** (Data Processing Unit) 承担网络、存储、安全等基础设施卸载；**DSP** (Digital Signal Processor) 在高速电/光链路处理中承担均衡、恢复等信号处理。DPU 与光模块中的 DSP 承担不同功能，不能混同；DPU 本身也不是光模块。[\[N027\]](#)[\[N065\]](#)[\[N090\]](#)

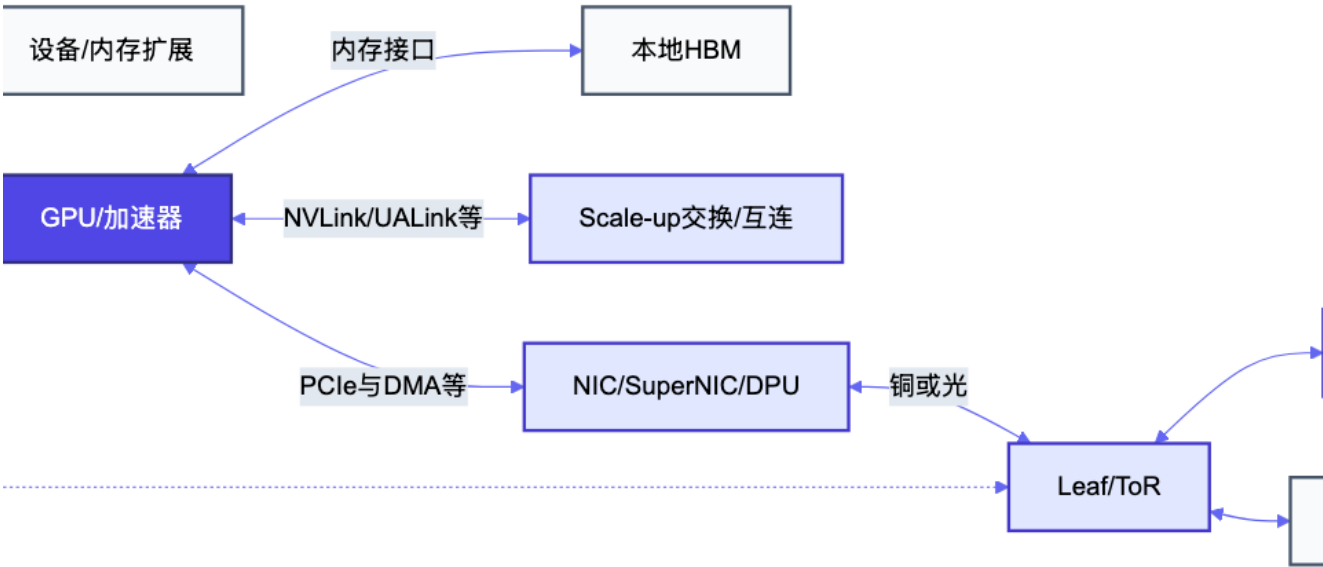
“光互联”描述传输介质与物理实现，“以太网/InfiniBand”描述网络体系，“NVLink/UALink”描述加速器 scale-up 互连，“PCIe/CXL”描述主机/设备/内存连接；这些处于不同抽象层。不能把 CPO 与 Ethernet、把光纤与 RDMA 放在互斥选项中。[\[N081\]](#)[\[N084\]](#)[\[N024\]](#)[\[N025\]](#)[\[N026\]](#)[\[N047\]](#)

四层连接蓝图

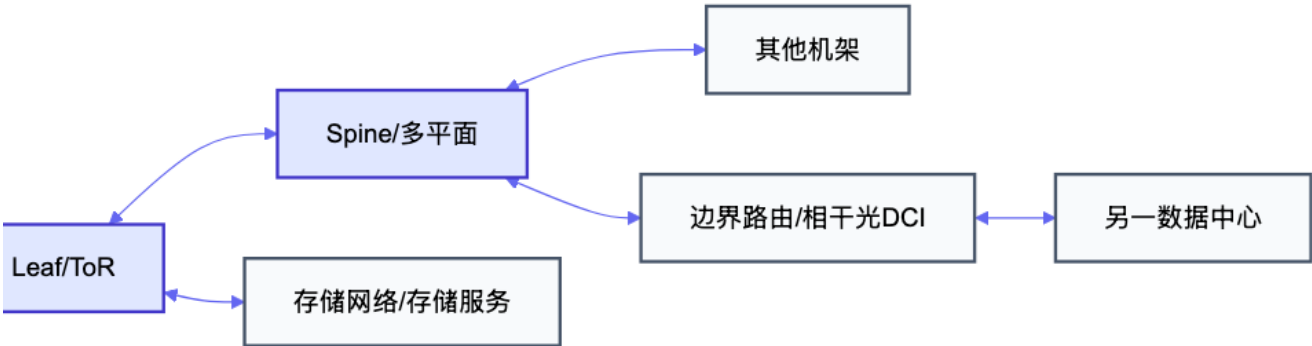
主机、加速器、机架、集群与跨中心互联（从左至右第 1/3 段，相邻区域重叠）



主机、加速器、机架、集群与跨中心互联（从左至右第 2/3 段，相邻区域重叠）



主机、加速器、机架、集群与跨中心互联（从左至右第 3/3 段，相邻区域重叠）



图是本文按功能绘制的通用抽象；GPU 到 NIC 的实际数据路径取决于平台拓扑与 DMA 能力，未画成“必经 CPU 内存”。HBM 仍是封装附近的本地存储接口，不能因为 GPU 使用网络便将 HBM 带宽与网络带宽相加。[N021][N022][N026][N027]

层	任务	典型组件	边界	证据
板/托盘/机内	CPU、GPU、NIC、SSD 等互连；内存扩展与一致性	PCIe root complex、switch、retimer；CXL controller	PCIe 兼容插槽不保证 CXL 功能，需主机/固件/OS 支持	[N025] [N026] [N028]
scale-up	紧密协同的加速器之间高带宽访问与通信	NVLink/NVSwitch、UALink、专用 ASIC/连接	通信语义、拓扑和软件决定能力；不局限于一台箱体	[N081] [N024]

层	任务	典型组件	边界	证据
scale-out	多节点/机架组成训练或推理集群	NIC、InfiniBand 或 Ethernet、leaf/spine、铜/光链路	需拥塞控制、路由、collective 与软件一起验证	[N021] [N022] [N023]
scale-across/ DCI	多园区之间数据、复制或分布式计算连接	边界路由、相干 DSP、DWDM、光线路系统	距离时延、故障域和可用带宽 改变应用可行性	[N046]

前端、后端、存储和管理网

前端网连接用户/服务、登录、控制和业务流量；**后端计算网**主要服务加速器 collective、分布式训练等大流；**存储网**输送训练数据、checkpoint、模型和缓存；**带外管理网**承载 BMC、设备配置与故障恢复。它们可以物理隔离，也可以按需求部分共用；“一张高速网”并不自动满足所有流量的隔离、尾延迟和维护要求。
[\[N021\]](#)[\[N022\]](#)[\[N027\]](#)

一次分布式训练迭代可拆成：存储读取→CPU/加速器预处理→本地计算→机内/跨节点通信→同步与 checkpoint。网络升级是否提高 GPU 利用率，取决于通信是否在关键路径；必须拿实际 collective 消息大小、并发作业、故障恢复和端到端训练时间验证。本文将此视为评估方法，不能由交换机标称吞吐推出训练加速比例。
[\[N022\]](#)[\[N023\]](#)

ToR、leaf、spine、rail 的逻辑

ToR (Top of Rack) 首先是接入交换机的部署位置/常用称呼；leaf 是拓扑角色，两者经常重合但不是严格同义。leaf 接端点，spine 连接 leaf，规模再扩大可能增加层次或多平面；高密度 AI 系统也可以采用列端、集中交换等布置，不能要求每个机架顶部必有一台交换机。
[\[N083\]](#)

rail/plane 设计的目的在于安排加速器/NIC 与独立网络平面的对应关系，控制故障域和通信局部性。多 rail 不能仅按“几张网卡”判断：还要核实 rank 放置、collective 拓扑、交换端口与路由是否匹配。以下是本文尽调检查，不把某一种厂商拓扑当作通用标准。
[\[N081\]](#)[\[N082\]](#)

带宽算例 (非产品配置)：48 个 400Gb/s 下联总计 19.2Tb/s，8 个 800Gb/s 上联总计 6.4Tb/s，则下联/上联为 3:1 订阅比；并非每个任务都同时打满，但不能称此拓扑无阻塞。另一个算例：64×800Gb/s=51.2Tb/s，必须核实厂商容量是否按单向端口速率求和、端口组合是否全开，不应再随意乘以 2。
[\[N022\]](#)[\[N063\]](#) (本文计算)

验证顺序：单端口 BER/丢包→多端口满载→incast/all-to-all→真实 collective→并发作业→链路/交换机失效→维护恢复。把平均吞吐、P99 延迟、重传、PFC 暂停、ECN、GPU 等待分别记录，可以区分“电/光链路好”与“集群网络工作好”。
[\[N022\]](#)[\[N084\]](#) (本文研究模板)

Ethernet、RoCE、InfiniBand 与新标准

InfiniBand 和 Ethernet 都有 AI 集群应用。RoCE 在以太网承载 RDMA，需要与交换机、NIC、拥塞策略和配置协同；应直接比较实际 workload 的吞吐、尾延迟、扩展、故障恢复、运维和总成本，不能用“Ethernet 天生不适合 AI”或“换成以太网必然便宜”替代验证。[\[N021\]](#)[\[N022\]](#)

UEC 是围绕 AI/HPC 的以太网通信体系，不只是端口速率。官网版本记录：1.0 首次公开于 2025-06-11，1.0.3 于 2026-07-16，1.1 于 2026-10-12。本文核实版本发布时间，未逐条审阅 1.1 全文，也未验证任意交换机/NIC 已通过该版本完整互操作测试。[\[N023\]](#)

UALink 1.0 在 2025 年 4 月公开，定位开放 scale-up；它与 UEC 的 scale-out 焦点不同。规范发布、芯片流片、开发板演示、量产服务器、多个厂商互通是不同里程碑。不能将联盟成员数量换算为已部署加速器数量。[\[N024\]](#)

PCIe 7.0 于 2025-06-11 发布到会员，速率 128GT/s；CXL 4.0 于 2025-11-18 公布，提供 128GT/s、bundled ports 与内存 RAS 增强。这些是规范状态：具体机器仍按 CPU、NIC、retimer、CXL 设备支持代际核查。[\[N025\]](#)[\[N026\]](#)

铜线并未因光互联而消失

实现	信号路径	优点/限制的来源	典型调查对象	证据
DAC	被动铜导体与连接器	低功耗/低成本；随速率、线径和板级损耗受限	插损、弯曲半径、长度与插接可靠性	[N031]
ACC	铜缆加模拟均衡/放大	降低复杂度；端点信号预算仍需协同	模拟架构、兼容性、链路余量	[N031]
AEC	铜缆加重定时/DSP	可改善链路与运维诊断；增加功耗/成本/时延	SerDes、固件、误码和大规模链路稳定性	[N089]
AOC	两端光电转换与固定光纤整体	安装方便；整根更换与布线维护需衡量	功耗、线缆现场更换、模块配合	[N030] [N048]
光模块+光纤	可插拔模块与独立光缆	分离维护、距离扩展；连接器和清洁进入预算	光功率/插损/反射、DSP、FEC、温度	[N030] [N045]
板级 retimer	电链路中恢复/重新发送信号	延伸 PCIe/CXL 等电信号通道，不是网络交换	协议代际、拓扑、热、遥测与软件支持	[N028]

米级、百米级、公里级只能作场景提示，不能作材料天生边界。距离由速率、波长、调制、FEC、收发器、连接器/熔接损耗和环境共同决定；具体采购必须查对应标准与模块数据表。[\[N031\]](#)[\[N046\]](#)[\[N049\]](#)[\[N063\]](#)

03B 把一条光链路拆到材料、晶圆、装配与测试

05 光互联的信号链与制造链

同一个光链路包含电子、光子、精密机械与控制软件；CPO、LPO、光I/O和DCI处于不同功能边界。



AI 基础设施产业研究 | 截止 2026-10-15 | 本报告原创示意图 | 产品阶段与企业证据见正文及来源附录

光互联信号与制造链

先看信号链 再谈 CPO

一条典型强度调制光链路的发送方向可抽象为：交换 ASIC/NIC 电信号→电通道/SerDes→DSP 或模拟驱动→激光/调制器→耦合器→光纤；接收方向是光纤→光电探测器→TIA→后续信号处理→主芯片。不是每个模块都包含完全相同的 DSP、激光和调制器组合。[\[N090\]](#)[\[N033\]](#)[\[N087\]](#)

光模块里“有硅”不代表激光在硅上直接发光；硅光 PIC 可以承担调制、波导、复用与探测等功能，光源仍可由 InP 平台提供。电子 EIC、光子 PIC、激光、光纤连接和封装是不同生产与测试对象。[\[N033\]](#)[\[N034\]](#)[\[N036\]](#)

方案	光电边界放在哪里	主要系统权衡	截至检索日应怎样表述	证据
retimed	前面板可插拔模块中含信号恢	互操作和维护方便；有模块功	已有多代产品；按具体 400G/800G/1.6T 机	[N029]
pluggable	复/DSP	耗及板级长电通道	型核查	[N048]

方案	光电边界放在哪里	主要系统权衡	截至检索日应怎样表述	证据
LPO	模块保留线性模拟收发链，将更多均衡压力留在主机端	可降低模块功耗；主机 SerDes、线性度、通道协同更关键	产品与演示并存，不等于任意端口可直接替换	[N030] [N049]
NPO	光引擎移近主芯片，仍可与芯片封装分离	缩短电路径；维护与板级布置改变	按厂商产品定义核查，术语边界可能不同	[N036]
CPO	光引擎与交换/计算芯片在共同封装系统中靠近	缩短高损耗电连接；封装、热、光纤接入和返修更复杂	Broadcom 有 Bailly 平台；NVIDIA 有 Photonics 平台发布；客户验收另查	[N032] [N064]
optical I/O	计算芯片封装附近以光 IO chiplet 接外部连接	面向带宽密度/距离/功耗；协议和封装集成是门槛	Ayar TeraPHY 与 SuperNova 是组件体系；平台出货需逐项验证	[N037]

CPO 与 LPO 并不构成覆盖所有距离、成本和维护需求的“最终赢家”。应按每条连接的带宽密度、插损预算、功耗、可更换性和故障代价建立决策。CPO 交换机采用情况不能直接推出 GPU 光 I/O 的采用情况。
[N032][N034][N037] (本文推论)

光器件与材料蓝图

以下“关键参数”为本文建立的尽调模板；来源支持材料/功能对应，参数合格与否必须以供应商数据表和客户验收为准。

部件	材料与功能	关键参数/问题	代表能力位置	证据
EML	InP 平台将 DFB 激光与电吸收调制器 EAM 集成	速率、消光比、啁啾、输出功率、温度与寿命	Lumentum ; Coherent	[N033][N039]
CW/DFB 光源	持续出光，供外部硅光调制器使用；可做 ELSFP	光功率、RIN、效率、波长稳定、失效恢复与激光安全	Lumentum ; 住友电工 ; 古河	[N034][N050] [N051]
VCSEL	GaAs 平台垂直腔面发射激光，服务多模等短距方案	单 lane 速率、温度、阵列一致性与模态条件	Coherent	[N039][N040]
硅光 PIC	波导、调制器、耦合器、探测器、滤波/复用构件集成	PDK、插损、串扰、温漂、工艺偏差和封装接口	GlobalFoundries 等硅光代工平台	[N036]
TIA/驱动/EIC	高速模拟与数字电路，把电子数据接到光子器件	噪声、带宽、线性度、功耗、接点寄生	Semtech ; Credo 相关 DSP/SerDes	[N030][N029]
探测器	将光变为电流，须与 TIA/带宽协同	响应度、暗电流、带宽、饱和、温度	Coherent InP 探测器产品	[N087]
WDM 与滤波	多波长共用光路，复用/解复用；实现依平台	通道间隔、插损、隔离度、温漂和锁定	GF 平台、Ayar 多波长光源体系	[N036][N037]
光纤/带缆	玻璃光波导、涂覆与护套、带状/高密度结构	单/多模、衰减、弯曲、密度、阻燃与施工	Corning ; 住友 ; 古河 ; Prysmian	[N045][N050] [N051][N060]

部件	材料与功能	关键参数/问题	代表能力位置	证据
插芯/连接器	MT/MPO 等多纤接口或 LC 类接口；精密对中和端面	插损、回损、端面清洁、极性、重复插拔	Corning；古河；住友	[N045][N050] [N051]

隔离与波导的区别：半导体电路的介电绝缘、光波导的折射率约束、激光前的光隔离器是三个问题。“光隔离器”抑制反向光反馈，不是层间电绝缘材料；WDM 的通道隔离度也不是介电击穿强度。采购/研究时应先问要隔离电流、热、光反馈还是相邻波长，才能对应正确组件。[N085][N066]（本文概念拆分）

硅光平台还应继续向下拆材料：Si/SiN 波导负责导光，Ge PIN/APD 可负责探测，金属后端承担电连接；Soitec 的 Photonics-SOI 是工程衬底位置，并非完整光模块。材料名称相同，不等于采用同一工艺、同一厚度或同一光学性能。[N066][N068]

SiN 的低损耗工艺可以涉及 LPCVD，重点在薄膜质量、应力和后续集成；薄膜铌酸锂等异质调制材料则可能通过转印与硅光组合。imec 的相关结果属于平台/研究演示，不能据此断言全部 1.6T 模块将转向某一种材料。[N067][N086]

无源光器件也有独立材料链：例如 Coherent 数据表列出铋掺杂稀土石榴石作为磁光旋转材料，用于隔离器等非互易器件。这是晶体生长、膜层、镀膜、组装与测量的问题，不能以逻辑芯片制程纳米数衡量。[N069]

制造这套器件需要什么

工序	为什么需要	设备/工艺范畴	怎样识别真实壁垒	证据
化合物材料外延	为 InP/GaAs 激光建立受控薄层结构	MOCVD 等外延；AIXTRON 有 GaAs/InP 设备	厚度/组分均匀性、缺陷、重复性、吞吐与维护	[N041][N039]
器件晶圆加工	将外延/硅光平台形成有源/无源器件	光刻、刻蚀、薄膜、掺杂、金属化；具体流程依 PDK	光学性能与电学良率同时达标，不能只比最小线宽	[N036][N033]
晶圆级光电测试	在昂贵封装前剔除问题 die、建立分档	自动探针、激光源、功率/波长/电学测试与校准	接触/耦合重复性、校准、测试覆盖、每 die 时间	[N044]
切割/贴片/互连	PIC、EIC、激光、基板组合	精密贴装、焊接、胶合与封装	热预算、机械应力、寄生、known-good-die 管理	[N036][N042]
光纤耦合与对准	把光高效送入/取出细小波导	多轴主动/被动对准、透镜/阵列、UV 胶或焊接固定	多通道最优点、固化漂移、节拍、长期稳定	[N042][N043]
终测与老化	确认成品在目标条件下工作	LIV、光谱/波束、burn-in、性能验证	温度点、误码、失效分析、追溯与返修	[N042][N044]
系统互通与现场	模块进入交换机、线缆与运维体系	标准/MSA、固件监控、链路测试	同规格不同厂商能否持续工作，故障能否定位	[N023][N047] [N049]

制造“光芯片”与生产“光模块”是不同层级；拥有模块装配能力不自动拥有 InP 外延和晶圆工艺。反向看，优秀激光芯片也需要电子器件、封装、散热、光纤连接、固件与系统认证才能成为可售网络端口。[N033][N036][N048]（本文供应链推论）

测试不是一台仪器 四个对象、四套指标

1. **激光/探测器 die**：LIV、光谱、响应、温度、寿命等；先确定器件本体和工艺窗口。[N042]
2. **PIC/EIC 晶圆**：光电联合探测与校准，关注芯片级参数、一致性与可测性；Keysight NX5402A 属于此类自动测试体系。[N044]
3. **光引擎/模块**：对准、固定、连接、热和固件共同决定最终链路；坐标轴精度只覆盖其中一部分。[N042][N043]
4. **网络系统**：误码/FEC、跨厂互通、拥塞、链路抖动/重训、故障恢复和长时运行；单模块达标不能替代集群验证。[N022][N023][N049]

六自由度通常是在 X/Y/Z 平移之外加三个转动自由度。PI 产品页明确有全电动 6DOF 版本；ficonTEC 说明多轴亚微米主动/被动对准。因此“能做 XYZ 动态对准”能说明能力类别，不能单凭这一点证明罗博特科/ficonTEC 独占市场。[N042][N043]

真正应追问的数据：客户产品/波长/通道数下的耦合良率分布；UPH、换线时间、设备稼动率；胶固化前后与温度循环后的损耗变化；误差校准和追溯；耗材/返修成本；客户量产线重复采购。上述是本文提出的可证伪壁垒清单，尚未拿到各家可比验收原始数据。[N042][N043][N044]

数据中心之间 从短距 IMDD 进入相干传输

园区/城域/长距 DCI 需要把路由/交换业务接口接到相干光模块或传输平台，再经波分复用线路传输。相干接收利用更复杂的光电检测和 DSP 处理，提高频谱利用与链路适应能力；线路端还涉及放大、光路配置、监控和光纤路由。Ciena 产品地图覆盖 coherent pluggables、line systems、backbone 与 submarine，这些不能压缩为一只“光模块”。[N046]

在 DCI 中，DWDM 增加同一光纤承载的波长，ROADM 负责可重构波长路由，EDFA 等补偿光功率损耗；链路预算仍受噪声、色散、非线性与连接损耗限制。800ZR 模块的速率是给定模式/链路条件下的接口能力，不能把展示的距离无条件扩展到所有光纤线路。[N088][N049][N065]

跨海光缆进一步加入海底线路、分支、供电、湿端与登陆站等专门工程。它服务全球数据连接，却不是一座机房内部 AI 后端网络的同类 BOM；本报告仅把它作为 DCI 边界延伸，不将海缆建设金额计入每个 AI 机架的内容价值。[N046][N060]（本文范围界定）

03C 网络与光互联企业地图、瓶颈与可证伪投资命题

本章是按技术位置建立的候选地图，不是市场份额排名，也不声称同一行企业相互供货。集团所在地、研发地、晶圆厂、封装地和销售市场可以不同；出口限制及客户认证须单独尽调。

按部件寻找企业 而非把“光通信”当一类

技术位置	美国/北美	中国大陆/中国台湾	欧洲、日本及其他地区	证据边界
交换/网络平台	NVIDIA ; Arista ; Cisco ; HPE Juniper	华为 ; 新华三 H3C ; 台湾智邦 Accton	各厂在多地区有销售/服务，不按销售地重复计数	平台/手册/ODM 产品；未核独立市场份额。 [N021] [N022] [N061] [N062] [N073] [N074] [N075]
网络 ASIC 与高速 DSP	Broadcom ; Cisco Silicon One ; Marvell ; Credo	盛科：官方披露交换芯片产品；不以整机品牌代替 ASIC 原厂	不因代工厂所在地改变设计商分类	盛科所列 12.8T/25.6T、最大 800G 等能力，不能视作与更高容量平台全面等效。 [N029] [N032] [N065] [N073] [R053]
NIC/ DPU 与板级连接	NVIDIA BlueField/ConnectX ; Astera Labs retimer	中科驭数 K2-Pro ; 云豹智能 DPU ; 分别核实接口速率和软件	UALink/CXL 参与者不等于已出货产品	K2-Pro 目录为 2×100GbE、PCIe Gen3×16；云豹 2021 年公开产品为 2×25G，不能混称 800G 产品。 [N027] [N028] [R054] [R055] [R056]
模拟光电接口	Semtech TIA/laser driver ; 相关 Credo SerDes/DSP	与模块厂共同适配，不能把模块品牌当 TIA 来源	硅光代工平台另有 SiGe 电子支持	retimed 和 LPO 路线有不同信号预算。 [N030] [N029] [N066]
激光/有源光芯片	Lumentum ; Coherent	源杰有 EML/CW DFB 目录；SiFotonics 有硅光产品，两种平台不混称	日本住友电工、古河电工；产品范畴不同	产品目录不能证明全部型号均已批量、良率或份额。 [N033] [N039] [N050] [N051] [N056] [R051]
微光/无源/光源/引擎配套	Corning 等布线与光学生态 须按产品分层	天孚：Mux TOSA、Demux POSA、Lensed FAU；腾景：准直/环形/隔离/VOA/WDM；仕佳：PLC/AWG	分别是组件、平台与装配能力，不是整条光模块业务同质竞争	目录支持产品类别，客户认证与每 SKU 量产规模另查。 [R049] [R050] [R052] [R057]
高速光模块	多家系统/器件商提供模块或光引擎	中际旭创旗下 InnoLight；新易盛 Eoptolink；光迅 Accelink	不按模块最终装配地推定上游芯片全本地化	InnoLight 组合、Eoptolink OIF 展示、Accelink 年报。 [N048] [N049] [N076]

技术位置	美国/北美	中国大陆/中国台湾	欧洲、日本及其他地区	证据边界
optical I/O	Ayar Labs ; Intel 研发/演示	SILITH 有平台展示 ; 台湾设计/封装生态需按项目核查	标准/代工生态横跨区域	光 IO 组件、展示、可规模采用是分开证据层级。 [N037] [N038] [N079] [N080]
硅光代工	GlobalFoundries	不把 TSMC 逻辑制程领先直接等同全部硅光客户采用	以色列 Tower ; 比利时 imec 为研发平台	GF/Tower 公开 PDK 与平台 ; imec 研发工艺不当成收入规模。 [N036] [N066] [N067]
工程衬底/材料	Coherent 含化合物与磁光材料能力	需要分晶体、外延、器件和贸易来源核查	法国 Soitec Photonics-SOI ; 日本住友相关材料能力	仅材料平台/产品证据, 未审各厂库存与客户供货表。 [N039] [N050] [N068] [N069]
光纤/插芯/连接	Corning	长飞 : 预端接干线与高密度布线 ; 亨通 : 棒纤缆与光网络解决方案	日本住友/古河 ; 意大利 Prysmian	光缆、连接器、光模块是不同产品和收入层级, 不能重复计算。 [N045] [N050] [N051] [N060] [R046] [R047] [R048]
装配与对准	FormFactor/Keysight 与伙伴组成测试体系	罗博特科相关 ficonTEC 业务需合并口径核查	德国 ficonTEC ; PI 精密运动/对准	多轴设备类别有竞争, 不给独家结论。 [N042] [N043] [N044] [N078]
外延装备	不以美国设备大厂列表替代激光外延专用机核查	国产化水平需看 InP/GaAs 具体工艺和客户验收	德国 AIXTRON MOCVD	设备功能已核 ; 厂商市占、客户配方与良率未知。 [N041]
DCI/光线路	Ciena ; Marvell DSP ; Lumentum 组件	新易盛有 800ZR 展示 ; 华为光线路另需项目核验	其他光传输厂商待补逐产品来源	网络系统、DSP、模块、光纤在不同收入层级。 [N046] [N049] [N065]

韩国在本章光互联企业地图中没有获得足以对比的逐产品证据 ; 资料覆盖不足不能视为相关产业能力缺位。电力章节引用 LS ELECTRIC 与 Hyosung 的官方资料 , 分别是配电产品/服务与电网设备及 SST 展示。
[\[N071\]](#)[\[N072\]](#)

与 CASPA 对话最相关的五个切口

- **Lumentum** : 分开问 EML、CW/外置光源和 OCS。EML 是已存在的器件路径 ; 外置激光源承接硅光/CPO 需求 ; MEMS OCS 改变的是网络光路。三者客户验证、BOM 价值和收入爬坡不应混算。[\[N033\]](#)[\[N034\]](#)[\[N070\]](#)
- **Ayar Labs** : 问 TeraPHY 与 SuperNova 的系统边界、计算封装接口、光源配置和客户部署阶段。官网能支持其完整光 IO 组件体系 , 不能替代客户对训练吞吐或单位成本的验收。[\[N079\]](#)

- **Semtech**：问 200G/lane TIA/driver、LPO 的模拟链路 with 主机 SerDes 适配、ACC 延长铜链路的条件。它在这些链路中的位置是信号完整性芯片，不能简单套“激光器涨价”逻辑。[N030][N031]
- **Credo**：问 AEC、光 DSP、retimer 与 SerDes IP 各自收入和产品周期；大集群更关心持续链路稳定及故障定位，而不只是一次跑满端口。公开产品组合可核，具体客户经济效果仍需客户证据。[N089]
- **SILITH**：2026-03-18 的官方摘要称 OFC 展示 400G/lane 硅光收发构件；现有证据范围限于该摘要。因此现场重点是索取演示条件、完整数据、样品/量产时间、制造/封装伙伴和客户验证；本报告不把它写成量产事实。[N038]

OCS 是另一个容易漏掉的组件层

补充一个材料到器件的可验证案例：源杰。 2026 半年报区分 CW 70/100mW 持续大批量交付、100G PAM4 EML 批量出货，以及 200G PAM4 EML 客户验证中。工艺讨论涉及 EML 异质外延对接、量子阱分别设计、寄生电容及高湿可靠性所需的钝化/镀膜；这是公司披露的技术难点，不能直接当成无竞争的护城河。应继续验证各代良率、认证转批量与毛利现金流，而非把全部新型号算入当期量产收入。[R064]

天孚 2026 半年报则将无源、有源与光模块代工组合，并列 FAU、多通道光引擎封装、高密度互连与 ELS 方案。它说明产业角色可以横跨部件与制造服务；财务分析须拆产品/工艺和客户，不应把“无源器件”标签套给集团所有业务。[R065]

Lumentum R300 的 MEMS 阵列在端点间建立光路，以减少中间 OEO 转换；它与逐包解析/转发的以太网交换 ASIC 承担不同功能。产品页面提供 300×300 端口框架，但实际网络仍需控制光路配置、故障切换和流量调度，端口数量不能直接换算成同数量 GPU 集群吞吐。[N070]

推论： OCS 的价值取决于通信图是否可组织、连接重配开销与任务阶段是否匹配；灵活性、控制软件和调度可能与镜面制造同样重要。应要求客户级拓扑与作业完成时间对比，避免只比较交换设备本体瓦数。[N070][N022]

将“瓶颈”拆成六种可验证命题

命题	为什么可能成立	最有价值的验证	可能推翻它的证据	依据
光源供应限制 光模块扩张	更高速率对激光功率、温度、寿命和制造一致性提出要求	已认证 SKU 产能、良率、交期、客户订单/备份供应	新厂爬坡、替代方案通过认证、库存与价格转弱	[N033][N034] [N035]
电子 DSP/ SerDes 决定 系统余量	高速电通道损耗与均衡影响光电链路	同通道长度/温度下 BER、功耗、抖动和恢复时间	主机 SerDes 改善、模拟链路满足需求、设计转向	[N029][N030] [N065]
硅光封装而非 晶圆成为限制	PIC/EIC/激光/光纤组合需要装配、对准和测试	每小时良品、固化漂移、封装终测良率和返修	被动对准、可拆卸连接、设计可测性显著降低装配成本	[N036][N042] [N078]

命题	为什么可能成立	最有价值的验证	可能推翻它的证据	依据
CPO 拉高专业设备内容	光引擎贴近大芯片，测试与维修边界改变	量产设备订单与客户产线验收、单位端口资本开支	CPO 量慢、既有设备可升级、设备复用提升	[N032][N042] [N044]
铜连接保持经济优势	短距连接可省光电转换，AEC/ACC 改进距离	同距离总成本、功耗、线束和稳定性	距离/带宽密度越过电预算、布线或散热代价上升	[N029][N031]
高端设备能力带来超额利润	失效代价、配方/算法与客户集成可能提高更换成本	复购、服务收入、验收周期、价格与毛利现金流	客户自制、竞争扩容、议价/集中度、技术替代	[N042][N043] [N078] (本文命题)

这里的“可能”是研究假说，不是已经发生的供不应求或买入结论。以高性能演示推断高毛利，以客户名单推断大额收入，都跳过了设备验收、交付、收入确认和现金回收。[\[N012\]\[N038\]\[N042\]](#) (本文分析方法)

用一个模型约束所有设备投资故事

本文建议按链条建立：**目标架构部署数量×每系统组件数量×单组件价格×供应商份额×可确认收入比例**。再扣除采购、制造、研发、售后、折旧和营运资本。不是所有系数都能从公开资料取得：没有客户 BOM/合约时应保留区间或未知，不把假设包装成市场数据。

对光装配设备，进一步拆为：**当年新增良品需求+单机年良品产出**。年良品产出取决于 UPH、工作小时、稼动率、良率和产品切换，而设备收入还取决于 ASP、存量设备升级与验收。由此可见，模块数量增长与装配设备收入不必同比例，自动化速度提高也会降低同产量所需设备台数。[\[N042\]\[N043\]\[N078\]](#) (本文模型)

对光源/芯片，关注良率和认证；对模块，关注代际 ASP 下降与客户集中度；对交换平台，关注软件、支持与端口配置；对材料，关注单位含量、纯度/一致性和替代；对数据中心电力，关注并网、交期与工程收入确认。这个拆法使不同公司能沿同一逻辑比较，而不强求用同一估值倍数。[\[N012\]\[N022\]\[N033\]\[N036\]\[N048\]](#) (本文分析框架)

论坛信息怎样进入思维框架

以下补充本模块的论坛案例；统一筛选方法见“社区异议与深入阅读”。

r/datacenter 关于直冷改造的讨论，给出管道净空、余热风冷、二次液体三类具体问题；与 OCP 的 TCS 边界、材料兼容和混合冷却定义能相互印证。保留的是问题清单，未采用匿名用户的价格或全行业渗透判断。[\[N052\]\[N010\]](#)

r/Semiconductors 的 1.6T 硅光讨论转述管理层发言，但讨论者对“更有份额”与“已经主导”有不同解释。本文据此保留三种变量：激光技术、模块集成路径、客户/代际混合；未将转述当原始财务口径，也未采用讨论中的份额推断。[\[N053\]\[N033\]\[N036\]](#)

需进一步核实的产业证据

以下项目缺乏充分可比的公开证据：统一条件下的光耦合良率/UPH/MTBF；主要客户完整 BOM 与认证名单；2026 年各 CPO 平台实际客户验收与数量；SILITH 量产与财务表现；关键 InP 外延、衬底及封装分档良率；韩国光芯片逐产品能力与跨国出口许可影响。代表性企业目录不能代替全部 SKU 的性能比较。

04 从 SRAM 到磁带 AI 的内存与存储全景

先按数据距离与任务分层

选择介质前先问四件事：数据是否正在参与计算；需要多少容量；必须多久取得；丢失后能否重算。

HBM、DDR、SSD 与 HDD 的连接位置和工作语义不同，“存储需求增长”不能直接推出所有层级收入同比例增长。[\[S026\]](#)[\[S039\]](#)[\[S040\]](#)[\[S041\]](#)

层级	系统位置与技术	常见 AI 数据	约束与验证对象
寄存器、片上 SRAM/cache	计算芯片内部；嵌入式 memory IP 与逻辑工艺协同	运算中间值、重复利用的块与缓存	面积、漏电、读写稳定性、端口数；不能用买 DIMM 的方式增加片上 SRAM。 [S032]
HBM	多颗 DRAM die 垂直堆叠，通过宽接口靠近 GPU/ASIC	当前参与运算的权重、激活与 KV cache	容量、带宽、每 bit 能耗、封装与散热；HBM 仍是 DRAM。 [S026] [S028]
DDR/LPDDR 主存	CPU/系统主存，DIMM 或特定模块/板载形式	数据准备、主机程序、可迁移的模型数据	通道与插槽、每通道带宽、容量、RAS、功耗。 [S027] [S036]
CXL 扩展或池化内存	主机经 CXL 接入设备内存，采用内存访问语义	可分层放置的较冷页面、共享资源池	控制器、交换、拓扑、操作系统页放置；规范支持与具体平台部署须分开。 [S039]
本地 NVMe SSD	PCIe 直连或通过本机交换结构	模型加载、数据暂存、检查点、局部缓存	随机/顺序、读/写比、尾延迟、耐久、热与容量。 [S019] [S020]
网络 flash/存储系统	NVMe-oF 等把存储节点接到计算集群	多节点数据集、共享检查点、分布式缓存	NIC/DPU、链路、拥塞、服务端软件、元数据；不只看单盘。 [S020] [S025]
HDD 容量层	磁盘、磁头、机械伺服及控制器组成的系统	大型原始数据、对象库、历史模型与日志	每 TB 成本、容量密度、随机 I/O、重建与可用性。 [S045] [S047]
磁带归档	磁带介质、驱动器、机械带库与目录软件	长期保留、离线副本和合规归档	装载/定位延迟、顺序流式吞吐、恢复流程、介质迁移。 [S049]

材料和连接也是分层的。SRAM 重点是逻辑晶体管与互连；DRAM cell 有访问晶体管和存储电容；HBM 在 DRAM 之上加 TSV/堆叠/封装；NAND 借助非易失 cell 阵列；HDD 与磁带依靠磁性介质。它们不是同一种“存储芯片”按速度排序后的可互换商品。[\[S030\]](#)[\[S032\]](#)[\[S026\]](#)[\[S046\]](#)[\[S049\]](#)

SRAM 为何 AI 芯片会“被缓存面积吃掉”

SRAM 常用于逻辑芯片内的 cache、buffer 和 register file 等结构。它的速度和近距离数据复用价值，使设计者愿意为其占用面积；但面积、功耗、容量、端口数和最小工作电压相互牵制。Synopsys 把小容量/高速需求与大容量/密度需求明确区分，因此“把更多数据都放 SRAM”通常不是免费的架构解法。[S032]

产业链应拆成三层：芯片设计者决定缓存层级及访问方式；memory compiler/IP 提供与工艺相容的可配置结构；代工工艺实现 cell 与互连。Synopsys 的产品目录把 SRAM、register file、TCAM 分开，说明“有存储 IP”也需要继续核查存储类型、工艺节点和接口。此处不据目录推断某 AI 芯片一定采用该 IP。[S059]

研究判断：若新增 SRAM 能减少从 HBM 往返搬运，可能提高系统能效；但扩大芯片又可能损害成本与良率。评估 SRAM 密集架构，应检查实际模型、命中率、片外流量和芯片面积，不能只用每颗芯片 SRAM 容量比较胜负。[S032][S040]

DRAM 电容缩放与主存配套芯片

DRAM 的器件难题是缩小 cell 的同时仍储存足够可辨识电荷，并控制漏电和外围感测差异。更高更窄的电容结构推动硬掩模、深孔刻蚀与共形薄膜；外围逻辑又有独立的高 k 金属栅、布线及低 k 介质需求。不能把“DRAM 采用高 k”统称为同一个位置。[S030]

在工艺分析中，把电容介质的电容/漏电目标、外围晶体管栅介质的栅控目标、后端低 k 材料的互连耦合目标分别列出。ASM 的 ALD 应用包含 MIM capacitor 与 HKMG，显示同类设备方法可以服务不同器件位置；器件研究中可看到 ZrO₂/Al₂O₃/TiN 电容膜系与原子层刻蚀的组合，说明介质、电极与图形修整需共同设计；这是论文实验结构，特定商业节点的配方、厚度与客户导入本报告未取得公开统一 BOM，故不作逐厂套用。[S031][S067]

一根服务器 DDR5 RDIMM 除 DRAM 还包括 RCD、PMIC、SPD hub 和温度感测；MRDIMM 则引入 MRCD、MDB 等。RCD 处理命令、地址和时钟分配，数据缓冲芯片处理数据路径，PMIC 处理供电。内存接口公司的技术价值，是在信号完整性、功耗与平台兼容性上支撑更高速度/容量，并不等于生产 DRAM cell。[S036][S037]

配套环节	可核实的代表	公司位置与尽调重点
服务器内存接口与缓冲	中国大陆澜起科技、美国 Rambus、日本 Renesas	官方资料列有 DDR5 RCD/DB、SPD/TS/PMIC 等不同组合；逐项比较产品，不能默认每家公司覆盖所有世代与全部料号。[S036][S037][S038]
标准与移动 DRAM	韩国 Samsung、SK hynix；美国 Micron；中国大陆 CXMT；中国台湾 Nanya	HBM/主存/移动产品和制造状态要分开；CXMT 官网列 DDR5/LPDDR5X，Nanya 投资者材料讨论 DDR5 等产品路线。[S026][S028][S029][S033][S034]

配套环节	可核实的代表	公司位置与尽调重点
利基型 DRAM	中国台湾 Winbond 等	Winbond 明确以 specialty/mobile DRAM 为产品线，不能把它视为拥有同等 HBM 量产能力。[S035]

CXMT 官网截至检索日已披露 DDR5 与 LPDDR 产品；这足以确认产品布局，不能据此认定与全球龙头在所有服务器容量、可靠性、HBM 或客户认证上等同。比较国产替代必须到具体规格、产量、平台验证及可交付性。[S033]

HBM 把“会造 DRAM”升级成“能交付高良率宽接口堆栈”

HBM 的基本结构是 DRAM die 垂直堆叠、用 TSV 建立穿硅通路，再与逻辑底片及外部处理器封装互连。真正的系统接口不是单一高速串行线，而是大量并行 I/O；拓宽接口提高总带宽，也放大供电、布线、测试与热管理压力。[S026][S028]

制造/集成步骤	技术任务	为什么会形成壁垒
DRAM 晶圆制造	cell、外围、I/O 与功耗设计	前道器件良率和电性分布影响后续可用裸片数。[S030]
TSV 加工及背面处理	穿硅连接、绝缘/阻挡、填充及露出连接结构	高密度通路、金属填充、薄片与机械应力必须协同。[S056]
裸片筛选与堆叠	将合格 die 与逻辑底片连接	若把缺陷 die 堆进去，其他已投入的合格 die 与封装价值也受损；已知良品 die 很关键。[S026]
微凸点/底填或后续混合键合路线	兼顾间距、机械稳定和导热	SK hynix 公布的 HBM4E 送样仍使用 Advanced MR-MUF，不能把混合键合写成当代所有 HBM 的必选方案。[S029]
与 GPU/ASIC 集成及测试	检查宽接口、热、供电、信号与整封装工作	单颗 DRAM 合格不等于堆栈合格，更不等于加速器系统客户认证。[S026][S056]

截至 2026-10-15 的状态切面。 Micron 3 月宣布 HBM4 36 GB 12H 已在 2026 年第一季量产出货；Samsung 2 月宣布 HBM4 商业出货，使用 1c DRAM 与 4nm 逻辑底片；SK hynix 6 月公布的是 HBM4E 12 层送样，不能把该消息当作 HBM4E 已量产。这三种表述均为各公司声明，不能直接组合成独立验证的市场份额表。[S027][S028][S029]

Micron 产品页给出的 HBM4 指标为 2,048-bit 级接口、每 pin 超过 11 Gb/s、每 stack 超过 2.8 TB/s；其 12H 的能效比较对象是自家 12H HBM3E。Samsung 的“最高”带宽与稳定 pin 速度也有各自边界。对外比较需要固定容量、堆叠高度、速度档、功耗口径与热条件，不能把不同峰值拼成统一跑分。[S026][S028]

研究判断： HBM 的门槛具有乘法特征：前道合格裸片、堆叠互连、热和整封装认证共同决定可售输出；示意上可写成“投入数量 × 各阶段良率”，但缺陷相关性、冗余/修复会影响真实计算。投资时应问良率发生在哪一阶段、可否修复、损失由谁承担，而非只问“有多少 HBM 产能”。[\[S026\]](#)[\[S056\]](#)

04A SSD、网络存储、HDD 与磁带怎样服务 AI

SSD 的核心竞争力在控制系统

SSD 控制器结合 NAND 通道、片上 SRAM、可选 DRAM、ECC/LDPC、数据路径校验与固件。FTL 把主机逻辑地址映射到物理位置；映射颗粒越细，通常需要管理的表项越多。Solidigm 指出大容量 QLC 盘的 DRAM 成本/功耗约束会推动更大的 indirection unit，因而需注意应用写入粒度与盘内部管理单位的匹配。

[S001][S017][S025]

DRAM-less 不等于 NAND-less，也不等于整机没有内存。它通常表示 SSD 没有独立外置 DRAM；可以利用片上资源、NAND 中的映射以及 HMB 等设计。慧荣 SM2504XT 公开支持 HMB，SM2508 则是另一个通道/缓存配置。应比较具体产品在随机访问、容量、功耗和断电场景中的表现，不能只按“有无 DRAM”定优劣。[S001][S017]

外形与散热也是系统接口。M.2、2.5 英寸 U.2/U.3 与 EDSFF 的 E1/E3 家族，涉及 PCB 空间、连接、功率预算、维护和热设计。E1.S、E1.L、E3.S 的名字不能替代 PCIe 代际、lane 数、单/双端口或盘容量规格；采购要把外形、接口、背板与服务器认证分列。Kioxia 将 EDSFF 明确作为改善旧外形限制的规范家族。[S073]

固件能力	要解决什么	应向厂商索取什么
LDPC 与数据保护	NAND 老化、错误位和内部数据通路风险	支持 NAND 世代、寿命末端纠错条件、不可恢复错误指标；不要把纠错能力当作介质零错误。[S017]
垃圾回收、磨损均衡与 OP	块擦除和页写入引发的内部搬迁	指定读写比例、盘占用率下的写放大、稳态吞吐与寿命；不把新盘 burst 当长期能力。[S018][S019]
PLP	掉电时保护在途数据与 FTL 一致性	是否保护已确认写入、用户数据及元数据，断电注入测试条件、电容健康管理；电容与恢复固件需配合。[S016]
QoS 与调度	多队列、后台任务及混合负载下维持响应	延迟分布及 p99/p99.9、预处理、容量档、温度、固件；平均值无法表示尾部。[S019][S023]
系统侧写入整形	避免小随机写对大 IU QLC 不利	应用块大小、IU、缓存层、合并策略和恢复路径；Solidigm CSAL 的公开方案把小写聚合成较大顺序写。[S024][S025]

耐久的读法。DWPD 是每天写入整盘次数，TBW 是生命周期允许的累计主机写入量，二者需连同容量、保修年限和工作负载理解。教学计算：假设 10 TB 可用容量、1 DWPD、5 年且每年 365 天，累计为 18,250 TB；这只是定义推算，不是某款盘承诺。真实比较仍应回到厂商额定测试与盘内写放大。[S071]

SNIA PTS 提供可比较的固态存储测试方法，强调统一测试环境与预处理。建议活动现场直接索要“满盘、稳态、混合负载、固定功耗下的延迟分布”，并问高温、寿命后段和掉电恢复后是否仍满足同一服务目标。这里是本报告提出的验证清单，不是声称已经完成实测。[S019]

从 NAND 成本到每个有效数据任务的成本

ScaleFlux 的公开产品说明把压缩/解压缩引擎集成到 SSD 控制器，用减少实际写入的方式改善容量使用和盘内工作。这里最容易误读的是“有效容量”：压缩后可装多少数据，取决于数据本身是否可压缩。已经压缩的视频、加密块和高熵数据，不应套用数据库文本的压缩比。[S021][S022]

研究判断：对压缩 SSD，比较单位应为“给定数据分布下的有效 TB、有效 IOPS 或任务完成成本”，同时报告物理 NAND 容量。现场要求同一原始数据集、相同压缩状态、相同功耗和耐久口径；还要测解压缩是否在读路径形成瓶颈。仅展示零填充数据或压缩友好样本，不能证明普遍优势。[S021][S019]

Solidigm 的 CSAL 则说明另一个方向：通过主机/DPU 软件重排写入，降低与 QLC 物理管理单位不匹配的成本。两者的价值来源不同：一个减少字节，一个改变到达介质的写入形态。它们可以互补，但都需要把控制器/DPU 资源、故障恢复与软件运营成本算进账。[S024][S025]

从单机 SSD 到全机架数据通路

本地路径可以是“SSD → PCIe → GPU 内存”，网络路径可以是“存储节点 → NIC/DPU → 网络 → 计算节点”。GPUDirect Storage 在支持的软硬件条件下减少 CPU 内存 bounce buffer，但 CPU 仍参与准备访问；减少额外复制并未取消 NAND 延迟、PCIe 拓扑或文件系统约束。[S040]

NVMe-oF 使用与本地 NVMe 相关的命令/队列体系，把访问延伸到网络。官方 transport specifications 分开定义 PCIe、RDMA 和 TCP；NVMe/TCP 利用标准 IP/Ethernet 基础设施。产品写“支持 NVMe”时，要继续确认是在本机 PCIe、远端网络目标，还是仅指其盘的主机接口。[S020]

2026 年 NVIDIA CMX/STX 设计引入 pod 级 flash context tier，由 BlueField-4、网络及推理框架协同放置和预取 KV cache。其重要性在于把可重算上下文与需要持久保护的企业记录分开处理；本报告视其为厂商参考架构及方案说明，未将宣称的 token 性能倍数视为跨模型独立实测结果。[S041]

CXL 属于另一种边界：它提供内存语义、一致性与池化/共享能力，不能与 NVMe-oF 画成同一种块存储协议。CXL 2.0 池化侧重把容量分配给不同 host；CXL 3.0 进一步描述多个 host 一致性共享区域。实际可用程度要核对处理器、交换机、内存设备及系统软件。[S039]

HBF 和高 IOPS flash 新层级 但尚需产品证据

Sandisk 与 SK hynix 在 2026 年 8 月经 OCP 发布首份 HBF 技术规格，目标是 HBM 与 SSD 之间的新层。规格发布证明技术协作进入标准化阶段，不能单独证明已有大规模系统采用。2025 年 Sandisk 公布的目标

是 2026 年下半年提供初始 HBF 样片、2027 年初提供首批含 HBF 推理设备样片；截至检索，本节不把目标日期自动转成已完成交付。[S042][S043]

Kioxia 的 GP Series 是另一条路线：2026 年 3 月宣布开发让 GPU 访问 flash 的 Super High IOPS SSD，并计划 2026 年底向特定客户提供评估样片。对这类产品，需要继续看访问粒度、控制器、读写语义、耐久、软件改造及实际延迟，而不是将“HBM 扩展”理解为与 HBM 完全等价。[S044]

研究判断： HBF、KV cache flash tier 与更快 SSD 可能分担冷权重、可复用上下文或检查点等不同任务。判断替代关系要画出“读取谁、写入谁、是否可重算、是否需随机小写”的数据流；用统一“AI memory”估值，会把研发样片与已认证量产混在一起。[S041][S042][S044]

HDD 磁性材料、光子和机械控制的复合技术

HAMR 的核心矛盾是：更小磁粒需要更强的磁稳定性，但更强的材料又更难写入。Seagate 的公开 Mozaic 技术采用高磁各向异性 FePt 介质、特制玻璃基底上的薄膜/退火工艺，并以近场光热写入路径配合读写头。这里的激光/光子器件服务磁记录，不能与数据中心光模块直接视作同一市场。[S046]

2026 年 3 月 Seagate 宣布最高 44 TB 的 Mozaic 4+ 已向两家 hyperscaler 批量出货，并披露自研制造激光技术。该容量与客户范围是公司声明；没有据此认定零售渠道普遍可买。WD 则在 2026 年 2 月披露 40 TB UltraSMR ePMR 正处客户认证、计划下半年量产，并继续并行 ePMR/HAMR 路线；当时计划不等于检索日已全部兑现。[S045][S047]

日本 Toshiba 公开 FC-MAMR 路线，说明磁记录增密可以使用不同辅助写入方法。投资验证应分别看介质、磁头、近场/微波辅助、伺服、振动、良率及客户认证，而非假设所有 HDD 厂商在相同时间采用同一平台。[S048]

研究判断： 更高 TB/drive 会降低某些容量部署所需盘位，却可能增加单位盘承载的数据及故障重建压力。HDD 与 SSD 的竞争必须固定容量、可用带宽、冗余、故障域、空间和电费，不能只比裸盘美元/TB；本报告未制作缺少当前成交价的 TCO 结论。[S045][S019]

磁带 容量、介质与恢复流程共同构成产品

IBM 当前文档列有 LTO-10 的 30 TB、40 TB 原生容量，压缩容量分别按 2.5:1 标为 75 TB、100 TB。已压缩或加密数据不能假定再达到该压缩比。“100 TB 磁带”必须注明压缩假设，不能与 100 TB 原生盘直接并列比较。[S049][S069]

带库的价值来自介质、磁头/驱动、机械取带、目录和恢复运维协同。对 AI 数据档案，可把读取频率低但需要长期保留的数据纳入候选；需要即时随机访问的热数据则不能只按磁带介质容量来选型。离线副本是否有效仍取决于恢复演练、目录保留、密钥和代际迁移。[S049]

04B 存储公司版图与可验证的投资问题

按能力归位 再看公司所在地

下表是已核实公开业务的代表性产业地图，不是市占率排名或所有公司的穷尽清单。“美国公司”“韩国集团”“中国台湾业务”描述企业主要身份/业务根基；它不等于晶圆、封装、客户和材料全部来自当地。

产业位置	中国大陆	美国及跨区域集团	其他主要地区	现场必须区分的能力
大容量 3D NAND 器件与制造	长江存储 YMTC：Xtacking 阵列/外围分开加工与连接。 [S007]	Micron：G9 3D TLC NAND；Sandisk：与 Kioxia 联合开发 BiCS。 [S015] [S014]	韩国 Samsung、SK hynix； 日本 Kioxia。 [S061] [S042][S013]	NAND 设计/制造、样片与量产、良品 bit 成本；不能只按扩层宣传排名。
HBM 与领先 DRAM	CXMT 官网有 DDR5、LPDDR 系列布局；本报告不据此声称其有同等 HBM 量产。 [S033]	Micron：HBM4、主存及存储产品。 [S027]	韩国 SK hynix/Samsung；中国台湾 Nanya 及利基型 Winbond。 [S028][S029] [S034][S035]	“能做 DRAM”与 HBM 堆叠、可靠性、客户资格的差别。
NAND/SSD 主控与固件	联芸 MAXIO：MAP1602-C 为客户端 DRAM-less PCIe4×4/NVMe2.0 主控；该资料不证明企业盘认证。 [R060]	美国/跨区域的 ScaleFlux、InnoGrit 等提供控制器或 SSD 方案。 [S021] [S065]	中国台湾 Phison、慧荣的产品资料可核控制器架构、代工及纠错功能。 [S017] [S064]	是自研 ASIC、外购主控、固件定制，还是仅做 PCB/品牌销售。
企业级 SSD 产品	忆恒创源 Memblaze：PBlaze 企业 SSD。 [S063]	Solidigm、Micron、Sandisk、ScaleFlux。 [S023][S027][S043] [S022]	Samsung、Kioxia；Phison Pascari 为企业 SSD 产品路线。 [S061][S044] [S064]	NAND 来源、固件责任、PLP、QoS、供货承诺与认证范围。
内存接口芯片	澜起科技。 [S037]	Rambus。 [S036]	Renesas，日本。 [S038]	RCD、DB、MRCD/MDB、PMIC、SPD、TS 逐项核对；与 DRAM 晶圆生产分开。
高深宽比工艺设备	中国供应商详见制造设备章；此处不由产品类别反推某 NAND 厂实际订单	Lam、Applied Materials。 [S005][S006]	TEL，日本；ASM，荷兰，覆盖特定沉积能力。 [S004] [S031]	能做一种刻蚀/沉积不等于能替代全流程；比具体薄膜、孔形与产能。
NAND 薄膜/清洗材料	国内候选须连同纯度、供料系统和客户工艺验证查核，具体公司见材料章	Entegris：清洗/高纯与材料整合方案。 [S008][S009]	Merck，德国：薄膜先驱体与化学平台。 [S010]	目录、送样、认证和持续量产消耗是不同证据阶段。

产业位置	中国大陆	美国及跨区域集团	其他主要地区	现场必须区分的能力
HDD 整盘与技术方案平台	本报告未核到与三家平台厂同等规模的中国自有 HDD 全链制造者	Seagate、WD；制造环节跨国分布。 [S045] [S047]	Toshiba，日本。 [S048]	磁头、介质、伺服、辅助写入、客户认证与制造良率。
磁带和专用持久存储	对应系统集成与运营能力需单独评估	IBM LTO/带库；Everspin MRAM。 [S049] [S051]	磁带介质与驱动供应链另需按料号核验	MRAM 是专用持久存储器件；磁带是另一种完整介质/机械系统。

实体边界。 Sandisk 与 Western Digital 在 2025 年 2 月完成业务分离；SEC 文件把法律分离日记为 2 月 21 日，2 月 24 日是对外宣布/交易相关节点。研究 NAND/SSD 应对应 Sandisk，研究 HDD 应对应 WD，旧 WD 品牌产品不应使两个上市实体的业务混算。[\[S050\]](#)[\[S070\]](#)

Solidigm 属 SK hynix 集团旗下的美国业务实体；本报告用它观察企业级 SSD、QLC 和系统协同，不能把集团 NAND、HBM 和 Solidigm 成品 SSD 收入重复相加为新增市场。集团官方简介可以确认隶属关系，具体重组与合并口径需回到期内财报。[\[S062\]](#)

用两条支线补齐架构

GDDR 与 HBM 是不同的 GPU 内存选项。 Micron 的教育材料把 DDR、LPDDR、GDDR、HBM 都归在 DRAM 内；GDDR7 产品资料面向图形、AI 与 HPC。比较应看板级连接、位宽、封装、容量、功耗及产品需求，不能把“GPU 用 HBM”写成所有 GPU 的普遍条件。[\[S060\]](#)[\[S066\]](#)

MRAM 不是 NAND 的同义替代。 Everspin 的 STT-MRAM 通过磁性隧道结状态保存数据，并提供持久缓存/嵌入式等产品路线。它可在特定写缓冲场景减少对保持供电的依赖，但这不能推广成已在容量、成本、工艺和生态上全面替代 DRAM/NAND。[\[S051\]](#)

从技术瓶颈推导投资假说 而非从名词直接选股

假说	技术触发	收益可能流向	必须验证的证据	可推翻它的事实
NAND 扩层推动工艺强度	更高堆叠带来更难的深孔、薄膜和字线	取得关键步骤导入的设备/材料商	客户认证、步骤覆盖、单片耗时/消耗、增量设备与服务	更快设备减少腔室数量；工艺合并、份额变化或资本开支推迟。 [S004] [S005] [S006]
企业 QLC 提升密度	更多 bit/cell 与系统统写入优化	NAND 原厂、能管理介质的 SSD/软件商	实际负载下寿命、尾延迟、写放大、有效容量/TCO	随机小写、后台搬迁或固件问题抵消容量优势。 [S018] [S024]
HBM 价值向封装/测试扩散	更多 die、更多连接、更严格热边界	经过认证的前道、堆叠、材料、测试环节	分阶段良率、吞吐、维修/耗材、设备份额、客户锁定	技术换代改变键合路线、瓶颈解除、客户内制。 [S026] [S029] [S056]

假说	技术触发	收益可能流向	必须验证的证据	可推翻它的事实
长上下文增加 可复用 flash tier	HBM 容量与重算 成本约束	SSD、DPU、网络与 缓存软件组合	命中率、预取成功、读取粒 度、GPU 利用率、成本/token	上下文不复用、重算更便宜、模型/ attention 优化降低需求。 [S041] [S044]
HAMR 提高容 量密度	磁记录稳定性与 可写性矛盾	平台厂及已导入的介 质/光子/磁头环节	云客户认证、批量出货、良率 和可靠性、每 TB 系统成本	客户迁移慢、可靠性开销过大、SSD 成 本或系统架构变化。 [S045] [S046]

研究判断：技术壁垒只有穿过“性能可验证 → 客户采用 → 可复制制造 → 定价/利润 → 合理估值”这条链，才构成投资结论。独家结构、专利或高难运动控制只是起点；本报告未以“能做某工艺”直接给出任何公司值得买入的结论。

外部研究和论坛怎样进入框架

SemiAnalysis 的 HBM 研究公开部分按内存墙、容量/带宽、制造与路线图组织问题；本报告借鉴从系统约束向供应链下钻的方式，没有引用付费墙后的未读数据。Seeking Alpha 所阅文章摘要强调 NAND 周期、产品变化与估值之间的张力；其评级与作者判断不作为技术事实来源。[\[S052\]](#)[\[S053\]](#)

技术社区 r/NewMaxx 的 SSD Basics 提供 FTL、TRIM、维护、磨损和垃圾回收等概念入口，价值在于暴露新读者常混淆的位置。本报告据此把“主控、固件、映射、物理介质”分列，再以厂商和标准组织资料核验；论坛本身不是良率、故障率或厂商市占证据。[\[S054\]](#)[\[S001\]](#)[\[S019\]](#)

投资社区关于“Micron 业绩能否直接映射 Sandisk”的讨论提出了有价值的校准问题：HBM、DRAM 与 NAND 的收入结构不同。采用的是问题，不是帖内的价格、产能或股价预测；公司业务归属由官方产品与分拆文件验证。[\[S055\]](#)[\[S027\]](#)[\[S050\]](#)

适合现场直接问的七个问题

1. 给 **NAND 原厂**：这代改进主要来自层数、横向密度、cell 状态还是外围/阵列集成？可以给每片晶圆良品 bit 与制造周期的可比口径吗？[\[S003\]](#)[\[S057\]](#)
2. 给 **Solidigm**：大容量 QLC 的 IU 是多少？4K/16K 写负载下，采用 CSAL 前后 WAF、稳态尾延迟和 DPU 成本如何变化？[\[S025\]](#)
3. 给 **ScaleFlux**：“有效容量”对应什么数据集和压缩比？对已压缩/加密数据，物理容量、持续写入和功耗又是多少？[\[S021\]](#)
4. 给 **Sandisk**：HBF 当前属于规格、送样、客户验证还是批量交付？相比既有 SSD tier，软件如何改变读写与耐久管理？[\[S042\]](#)[\[S043\]](#)

5. **给 HBM 供应链：**贵公司的收入依赖哪一步？新微凸点/混合键合路线会增加还是减少自身设备/材料价值？[S029][S056]
6. **给材料商：**客户采购的是通用高纯材料还是经其结构验证的专用配方？替换供应商需要重跑哪些电性/可靠性窗口？[S009][S010]
7. **给系统厂：**当前 GPU 空转主要来自存储、网络、解码/预处理还是软件调度？改善单盘性能后，整任务时间和成本/token 是否实测下降？[S019][S040][S041]

05 晶圆制造 把电路图变成三维器件

04 晶圆制造与设备材料对应

工艺循环而非单向装配线；每一层器件/互连都可能重复图形、薄膜、刻蚀、清洗和量测。



AI 基础设施产业研究 | 截止 2026-10-15 | 本报告原创示意图 | 产品阶段与企业证据见正文及来源附录

晶圆制造与设备材料对应

先分清五种产业角色

芯片设计公司交付电路与版图；EDA/IP 公司提供设计、验证、签核工具及可复用模块；晶圆厂把版图变成器件；设备厂提供执行工艺的机器；材料厂提供晶圆、气体、薄膜前驱体等耗材。它们共享工艺接口，但不能互相替代：例如 Synopsys 的 Intel 18A 认证流程和 IP，证明设计生态与特定制程的适配，不代表 Synopsys 制造晶圆。[F017]

晶圆厂内部又分 FEOL（形成晶体管）、MOL（器件接触）、BEOL（多层金属布线）；封装阶段常另称后道，不能把“BEOL 布线”和“封装测试”混为一谈。东京电子的工艺图把沉积、图形化、刻蚀、清洗反复组合，随后才进入互连、晶圆探针测试及键合，这比“一台光刻机做出芯片”的理解准确。[F001]

制造主循环与工程验收点

下表是通用工艺能力字典，不是某款 GPU 的保密配方；真实器件会按层重复、增删与重排步骤。每一步的价值都应以最终电性、缺陷和良率验收，不能只看单台设备的最佳指标。[F001][F002]

工序	解决的物理问题	关键能力与应问的指标	可核实的设备例子
衬底、氧化与外延	提供晶体基础，形成绝缘或应变层	晶格缺陷、掺杂均匀性、厚度、选择性；Si/SiGe 外延还要控制组分与应变性	ASM 外延，TEL 炉管；外延晶圆供应和器件厂内选择性外延是不同业务。 [F006] [F019]
CVD/PECVD	用化学反应形成介质、半导体等膜	成膜速率、应力、覆盖率、颗粒及热预算；PECVD 用等离子体帮助反应	Applied Materials 介质沉积；TEL 批量 LPCVD；拓荆 PECVD。 [F002] [F011]
ALD/PEALD	在复杂三维表面控制极薄膜厚	表面反应控制、共形性、杂质、循环时间、批量一致性	ASM Pulsar/EmerALD/Synergis；不能把“有 ALD”理解为所有材料、所有形貌通吃。 [F005]
PVD、金属沉积、电镀	提供金属膜、阻挡层、种子层及填充	电阻、附着、空洞、填充形貌、金属污染	Applied Materials 金属沉积/电镀；北方华创 PVD。 [F002] [F009]
涂胶、曝光、显影	将版图转换为光刻胶图案	分辨率、套刻、焦深、线边粗糙度、随机缺陷与产能	TEL 涂胶显影；ASML EUV/DUV；TOK 配套光刻材料。 [F001] [F003] [F044]
刻蚀、选择性去除	把图案转移到下层，或移除指定材料	各向异性、选择比、深宽比、侧壁、底部损伤、终点控制	Lam、TEL；中微 CCP/ICP/TSV；北方华创。 [F007] [F010] [F009]
离子注入、退火	调整掺杂并激活，修复损伤	剂量、能量、均匀性、结深、热预算	Applied Materials 注入/热处理；北方华创相关系列。 [F002] [F009]
CMP	将表面重新变平，支持下一层图形化或键合	去除速率、选择性、dishing/erosion、划伤和片内均匀性	Applied Materials、荏原；抛光机与抛光液、垫、过滤器一起决定结果。 [F013] [F023]
去胶、湿洗、干洗及干燥	去除残胶、颗粒、金属和工艺残留	清洁率与材料损失的平衡、脆弱结构损伤、干燥水痕	TEL、盛美上海；清洗并非只有最后一道。 [F001] [F012]
检测、量测、缺陷复查	找到哪里失控，给设备反馈	CD、套刻、膜厚、形貌、缺陷捕获率、误报率、每小时晶圆数	KLA 光学/电子束、掩模与化学过程控制组合；检测缺陷和测量尺寸是不同任务。 [F008]

三项结构变化为何改变设备需求

从平面到 FinFET/GAA。GAA 的沟道被栅极环绕；以 Si/SiGe 纳米片 GAA 为例，工艺需要多材料层堆叠、释放沟道、选择性去除及三维表面覆盖。ASM 的技术材料把 Si/SiGe 外延、high-k 栅介质和 ALD 列为

相关能力。推论：设备价值向“在不伤害相邻材料的情况下，稳定加工某一层”移动；不能只按晶圆直径或标称节点评价难度。[F005][F006]

从正面供电到背面供电。背面供电把电源布线移到晶圆背面，为前面信号布线释放空间并改善供电路径；制造上涉及承载、减薄、背面对准、介质和金属沉积、刻蚀、CMP 等协同。Applied Materials 的原理页含“所有芯片仍正面供电”的历史化表述，本报告只引用其结构定义与工艺范围，不把该句作为 2026 年产业现状。[F004]

从平面存储到高层数 NAND。NAND 的难点不是只把平面线宽缩小，而是在厚叠层内加工深而直的孔，控制整片膜应力、字线电阻与填充。Lam 的 Cryo 3.0 面向 400 层及以上器件的工艺需求，其页面同时列出氧化物/氮化物叠层、钨/钼金属化及翘曲管理。这是设备适用性和厂商目标，不是所有客户已量产 400 层的证明。[F007]

光刻机背后还有哪些不可省略的层

ASML 的光源路线包括 KrF 248 nm、ArF 193 nm、EUV 13.5 nm；EUV 用激光激发锡滴等离子体产生光。波长不等于节点名称，不能把“13.5 nm 光”读成“只能制造 13.5 nm 芯片”。EUV 的传播与成像采用反射镜体系；DUV 主要使用透镜体系，光学系统、污染与真空环境的要求因此不同。[F003][F054]

图形化还依赖掩模基板、吸收或反射膜、写版、掩模检测、光刻胶、显影、计量与计算修正。信越的掩模 blank 是待写入图案的基材，和最终电路掩模不是一个产品；TOK 的 EUV/ArF/KrF 胶也不能视为同一配方跨光源直接使用。[F020][F044]

CASPA 对应切口：Multibeam。该公司的多柱电子束直写提供无掩模图形化，官网定位包含先进封装、硅光、化合物半导体和高混合度原型制造。推论：应以目标图形面积、写入时间、设备利用率、掩模摊销及对非平面表面的适应性，和传统方案做项目级比较；不能因“无需掩模”就推出能经济替代先进逻辑所有关键层的 EUV。官网的倍数性能宣传未在此转为独立验证结果。[F018]

晶圆厂也是一座精密化工厂

晶圆工艺机器需要真空泵、压力调节阀、质量流量控制、RF 电源与匹配、温控、超纯流体、尾气处理等子系统。Edwards 的集成方案将泵、排气加热、气体探测、冷却水与控制信号一起设计；荏原公开的工厂图将洁净室设备、sub-fab 真空/尾气/冷机与厂务设施分层。推论：工艺强度增加可给子系统和维护服务带来需求，但必须核对每腔价值量及机型配套，而不能按 GPU 数量线性外推。[F014][F074]

RF 电源决定等离子体能量耦合的一部分，阀和泵共同维持工艺压力及洁净排气。MKS 的脉冲 RF 案例说明三维器件会提高控制要求；VAT 的产品定位是高性能真空阀。真实壁垒应看动态响应、颗粒、腐蚀寿命、维修周期、和主机工艺认证，而不是“零部件单价低所以容易替代”。[F015][F016]

区域设备公司与验证口径见下一节；材料体系与封装测试分别见第 06、07 章。

05B 设备公司地图 按工艺能力定位

本表是经过公开产品资料核对的代表性能力地图，不是份额排行榜，也不表示同一行各公司在相同节点、材料和客户的资格等价。设备选型最终要落到“哪一道工艺、哪种膜、哪种结构、哪家工厂已验收”。[\[F002\]](#)
[\[F009\]](#)[\[F010\]](#)

地区/公司	可核实的能力	产业位置与比较边界
美国 Applied Materials	沉积、外延、注入、热处理、CMP、刻蚀、异质集成	广工艺材料工程平台；不同产品线的客户、工艺份额须分别核实。 [F002]
美国 Lam Research	沉积、刻蚀、去胶清洗；NAND 高深宽比介质刻蚀	看 NAND 层数、结构和材料改变对应的工艺步骤，而不是仅看 NAND bit 增长。 [F007]
美国 KLA	晶圆、掩模、封装检测测量测及过程控制软件	控制良率的反馈环；不是量产晶圆代工厂。 [F008]
美国 MKS	RF、气体/压力与工艺子系统相关能力	对接设备主机或工厂；具体子系统与 OEM 的收入关系需要财报确认。 [F015]
美国 Multibeam	多柱电子束无掩模图形化	适合研究其先进封装、硅光及高混合度制造场景的经济性。 [F018]
荷兰 ASML	EUV/DUV 图形化及相关量测、计算光刻	需区分 EUV 和 DUV、装机与产能、光源功率与可交付吞吐。 [F003] [F054]
荷兰 ASM	ALD、Si/SiGe 等外延	共形膜与材料界面；应按具体材料、形貌和客户工艺认证比较。 [F005] [F006]
瑞士 VAT；英国 Edwards	真空阀；真空泵/尾气/集成 sub-fab	子系统可同时服务多种工艺，需按应用和认证程度拆分。 [F016] [F014]
日本 Tokyo Electron	涂胶显影、刻蚀、沉积、清洗、热处理、探针机等	探针机做接触定位与温控，不能单独替代电性 ATE。 [F001]
日本 Nikon；Canon	Nikon：ArF 浸没/干式及封装图形化； Canon：i-line/DUV/NIL	分别比较具体层的套刻、缺陷和经济性，不将 NIL 产品发布等同先进逻辑全面替代 EUV。 [F080] [F081]
日本 NuFlare；Lasertec	电子束掩模写版；EUV/DUV 掩模检测、blank/光罩膜等检测	写掩模、查掩模、在晶圆上曝光是三种环节；NuFlare 与 Multibeam 的直写目标不同。 [F079] [F075]
日本 SCREEN；DISCO	清洗/涂胶显影/退火/直接图形化；切割/减薄/抛光	将膜层工艺设备、晶圆减薄、裸片分离与耗材业务分开。 [F077] [F076]

地区/公司	可核实的能力	产业位置与比较边界
德国 ZEISS ; TRUMPF	半导体光学 ; EUV 光源驱动激光	是图形化系统的上游能力 , 与整机集成商的收入和交付口径不同。[F078][F083]
日本荏原 Ebara	CMP、干式真空泵、尾气处理	同时位于工艺设备和 sub-fab 层 , 收入统计不可全部算成 CMP。[F013]
日本芝浦机电 Shibaura Mechatronics	清洗/刻蚀/去胶、2.5D/3D 键合及混合键合	此处是半导体设备公司 ; TFC-6700/6800 为 C2W hybrid/fusion 产品 , 不宜笼统归为“芯片测试”。[F031]
中国大陆北方华创	等离子刻蚀、PVD、CVD、湿法、炉管、注入、RTP、EPI、键合等公开系列	产品目录证明覆盖 , 不能单凭目录推导所有先进工艺都已完成替代。[F009]
中国大陆中微 AMEC	CCP/ICP、TSV 刻蚀 , MOCVD ; 扩展薄膜设备	2025 年官方事实表将既有刻蚀应用和新薄膜方向分别表述 ; 研发、客户验证与批量订单应分开。[F010]
中国大陆拓荆 Piotech	PECVD、ALD、gap-fill、3D IC	官网确认产品家族 ; 具体工艺验收和量产占比需逐设备核实。[F011]
中国大陆华海清科 ; 中科飞测	CMP/减薄等装备 ; 晶圆缺陷检测、CD/膜厚/套刻/三维形貌量测	华海清科 2026 半年报确认装备方向 ; 中科飞测目录确认覆盖 , 不用“有产品”替代先进工艺验证。[F084][F085]
中国大陆长川科技	测试机、分选机、探针台、AOI	平台目录覆盖多种应用 , 不能直接推出已替代最高端 AI SoC 测试平台。[F086]
中国大陆盛美上海 ACM	单片、post-CMP 等清洗设备	本次来源直接支持硅片/SiC 衬底清洗 ; 其他新产品不能由该页自动外推。[F012]
中国台湾致茂 Chroma	SLT、FT、burn-in、热控、接触与自动上下料	位于测试及验证层 ; 不是所有测试都由致茂一家公司包办。[F035]
韩国 HANMI	HBM TC bonder 等后道设备	本次读取 SEMI 展商目录中的厂商自述 ; 份额宣称未独立审计 , 不据此给出排名。[F051]

四个容易被投资叙事省略的成本

工艺窗口。最佳样品证明“能做” ; 整片、跨片、跨腔与连续批次的稳定性才更接近“能制造”。GAA 的多材料选择性和 NAND 深孔控制会使单指标比较失真。这是从设备技术文档得到的研究推论 , 需用客户验收数据验证。[F005][F007]

单位良品成本。 建议比较：设备折旧、耗材、维护、停机、返工、占地和公用工程成本，除以最终良品量。ALD 膜质量提高若以循环时间增加为代价，是否划算取决于良率增量和产能约束。论坛关于 ALD 速度、批量能力和膜质量取舍的讨论可作为现场追问线索，匿名身份及具体数值不作行业事实。[F057][F005]

认证与服务。 工厂购买的是持续合格产出，采购价只是一部分。Edwards 明确把维护、翻新和集成服务列入供给；ficonTEC 的应用软件把定位、校准、测试和固化动作连接起来。推论：长期壁垒可能来自 recipe、数据和现场支持的组合，不能只看机械零件是否可买到。[F014][F036]

路径依赖与可替代性。 一台新设备需要与材料、前后工序和厂内数据接口配合。EVG 的 die-to-wafer 白皮书明确把清洗、carrier handling、临时/永久胶等纳入开发项目。推论：更换供应商的成本应以整条流程重新验证计算，但历史认证也不等于永久垄断。[F058]

制造壁垒的最短尽调清单

现场应逐项索要：①设备做的是哪一层及哪一材料组合；②结果是实验、客户验证还是量产验收；③精度属于分辨率、重复性还是绝对精度；④吞吐是 dry UPH 还是含校准/搬运/测试的有效 UPH；⑤缺陷、耗材寿命和维护间隔；⑥客户换线与新代际复用成本；⑦订单能否交付、验收、回款。这里是本报告的验证框架；例如芝浦的公开规格明确附有室温/local alignment/ 3σ 等条件，说明脱离测试条件抄精度会误导。[F059]

06 材料地图 同样叫“绝缘” 实际承担不同任务

把材料分成五本账

①衬底提供晶体和机械基础；②工艺材料被消耗或形成薄膜；③设备耗材维持洁净、温度与等离子体环境；④封装材料连接、保护和散热；⑤光电/功率功能材料承担发光、导光或高压开关。不能把电子级多晶硅、完成抛光的单晶晶圆、SiC 散热衬底和封装 ABF 膜都称为“芯片原材料”后直接相加市场规模。下表按所在结构及失效模式归类。[F019][F024][F053][F068]

材料/所在层	功能和工艺耦合	主要验收问题、成熟度边界
电子级多晶硅→单晶锭→切片/抛光/清洗	多晶硅是晶圆上游；晶圆再进入器件制造	纯度、晶体缺陷、平坦度、氧/碳、批间稳定；光伏级产能不能直接当电子级合格产能。[F061][F066]
抛光/外延/退火晶圆	不同表面与晶体条件服务不同器件	外延厚度、掺杂、电阻率及缺陷；外延片不是“尺寸更大的硅片”。[F019][F066]
SOI (硅/埋氧层/硅)	埋氧层提供隔离；光子 SOI 用于硅光器件平台	活性层、埋氧层厚度及均匀性与光学设计耦合；并非所有先进逻辑都采用 SOI。[F050][F066]
SiO ₂ 、SiN 与不同介质薄膜	在器件/存储叠层中承担隔离、掩膜或结构作用	成膜、选择性刻蚀、应力需匹配；同一化学名称不代表同一层的电学职责。[F007][F011]
high-k 栅介质、金属栅	在极薄栅堆栈维持栅控与漏电平衡	界面缺陷、厚度、杂质、阈值和热预算；ALD 是关键实施方法之一。[F005]
low-k 布线介质	降低相邻导线寄生电容	介电性能、机械强度和加工损伤间权衡；与 high-k 栅介质用途相反，不能按“k 越高越好”排序。[F060]
Cu、W、Ru/Co 及 barrier/liner	导电、接触、填充或界面保护	随线宽缩小，阻挡/衬层占截面比例变大；Applied 的 RuCo liner 方案是具体集成路线，不能扩展成“全行业铜即将淘汰”。[F060]
PVD 靶材：Cu/Ti/Ta 等	靶材经溅射形成薄膜	纯度、晶粒、致密度、颗粒与使用寿命；金属原矿价格不是溅射靶材的全部价值。[F022]
光刻胶、底层/反射控制、显影/清洗材料	成像并转移电路图形	分辨率、缺陷、刻蚀耐受、纯化及与 track 的适配；EUV、ArF、KrF 分开评价。[F044]
DUV 掩模 blank / EUV 光学材料	制造图形模板；EUV 采用反射光学	blank 缺陷和表面控制影响后续成像；DUV 透射式与 EUV 反射式体系不能互换。[F020][F054]
特气与 CVD/ALD 前驱体	反应生成绝缘、金属、半导体薄膜，或执行刻蚀	分子、纯度、储存输运、挥发/分解行为、残留和批次一致性共同决定是否可用。[F021][F045]

材料/所在层	功能和工艺耦合	主要验收问题、成熟度边界
CMP slurry/pad/brush/filter	化学反应与机械去除协同平坦化	去除率与选择性、划伤、颗粒团聚及清洗相容性；不能只按磨料硬度评估。[F023][F043]
石英、陶瓷、Si/SiC 设 备零件	炉管、舟、环、腔体内件、绝 缘和耐腐蚀结构	热/化学循环、颗粒脱落、清洗再生寿命；这类“耗材”不一定留在最终芯片内。[F068]

回答“芯片层之间用什么隔离”

必须先问隔离发生在哪里。晶圆内的绝缘薄膜、布线间 low-k、SOI 埋氧层，以及封装基板的 ABF/BT，尺度、介电任务、温度和制程完全不同。堆叠结构的层间电隔离无法靠一张通用隔离膜解决：该绝缘的区域需要介质，该导通的区域需要受控的金属互连，二者要一起制造和测试。[F005][F024][F046][F050]

NAND 叠层中出现的氧化硅/氮化硅，要区分工艺中间结构与最终器件；Lam 资料明确把叠层控制、深孔刻蚀、金属字线和应力控制放在同一个系统里。详细电荷俘获结构见存储章；这里的材料判断原则是：先确认它在该工序是永久功能层、牺牲层、掩膜还是保护层，再判断供应商及工艺壁垒。[F007]

对芯片三维封装，EVG 说明混合键合是在介质键合界面嵌入金属接点，形成机械连接和电连接；这和用有机 ABF 形成封装基板层间绝缘不是同一道工艺。推论：需要同时看绝缘漏电、Cu 接点、表面平坦度、颗粒、对准及热循环，不能只找一种“最先进绝缘材料”。[F029][F024]

封装与散热材料

材料	精确位置/作用	工程风险与状态
ABF 增厚绝缘膜	封装基板多层微线路间绝缘，连接纳米级芯片与较大尺度板级互连	味之素官方强调树脂/无机成分分散、低热膨胀与加工性；ABF 膜厂、基板厂、封装厂不是同一角色。[F024]
BT 树脂/层压材料	半导体封装基板材料体系	三菱瓦斯化学将 BT 定义为以 bismaleimide/triazine 为主的聚合树脂；与 ABF 比较必须匹配产品结构和工艺。[F046]
焊料/铜柱、底填、模封	电连接及机械保护	底填增强互连对热机械应力的耐受；与模封树脂、die attach、lid attach 分开核算。[F055][F056]
TIM、铜散热件、复合材料	将芯片热量传到顶盖/冷板等后续结构	体材料导热率之外还看界面热阻、厚度、泵出、老化及接触压力；Henkel 供应不同界面及保护材料。[F055]
玻璃芯基板	更大、更密封装的候选基板平台	Intel 2023 年公告是研发/测试与后续路线规划；不能由测试载具推出所有 AI 封装已全面采用玻璃。[F025]

材料	精确位置/作用	工程风险与状态
CVD 金刚石、Cu-diamond	近芯片热扩散材料路线	Element Six 提供相应方案；产品可获得不等于大规模 AI GPU 量产份额。优先询问界面、加工、良率、成本及可靠性。 [F052]
高导热 SiC 衬底	可用于热管理，与功率器件用 SiC 业务需分开	Coherent 2026 年 8 月 17 日公告 300 mm 高导热 SiC 开始客户送样；本报告标记为送样，不写成规模量产收入。 [F064]

光电与功率材料 跨入机房系统

InP 用于高速光通信相关激光器/探测器材料平台，SOI 为硅光提供器件基础；两者可在同一光引擎供应链出现。AXT 的 InP 衬底是晶圆材料，不是装好的激光器或光模块。它的总部在美国，制造在中国子公司；“美国公司/中国公司”单一标签不足以描述这一供应风险。[\[F062\]](#)[\[F072\]](#)[\[F050\]](#)

Si、SiC、GaN 在数据中心电源内可混合使用。Infineon 的 8 kW 参考设计用 SiC PFC、GaN 等 DC-DC 级，说明材料选择取决于电压、电流、频率、成本与拓扑，而不是高价材料完全替代硅。参考板峰值效率不等于机房全年端到端效率。[\[F053\]](#)

06B 材料公司分布与投资验证

区域公司地图

同一公司可能跨地区生产；以下地区主要用于识别公司和供应网络，不代表生产原产地、出口资格或客户认证。列举代表性公司用于形成可研究的蓝图，不作没有统一口径的“全球前五”排名。[\[F072\]](#)[\[F065\]](#)

地区	公司及公开能力	进一步要拆的问题
美国	Entegris：CMP、过滤、污染控制；Hemlock：多晶硅	工艺耗材、容器/过滤及材料产品分别核算；电子级和光伏需求分开。 [F023] [F067]
美国/中国制造网络	AXT/Tongmei：InP、GaAs、Ge 衬底；VGF 晶体生长	公司归属、制造地、原料来源与跨境交付路径分开；不采信官网未经独立验证的供应商数量/份额作为事实。 [F062] [F072]
中国大陆	沪硅 NSIG：300 mm 硅片及 SOI/定制晶圆网络	官网所列部分客户认证不等于所有产品节点全覆盖；300 mm 产能还要问正片/测试片、合格率及利用率。 [F041]
中国大陆	安集科技：Cu/阻挡层、介质、W、衬底及先进封装 CMP 化学品	TSV/混合键合新配方、已量产与下一代验证状态逐项确认。 [F043]
中国大陆	华特气体：电子特气及光刻混合气	2025-06-20 官网列举产品和认证，属于公司声明；不能把某种光刻气认证扩展为全部气体或全部客户认证。 [F042]
中国台湾	GlobalWafers：抛光、外延、退火、SOI、化合物晶圆与硅锭等产品线	台湾公司有全球生产布局；具体晶圆工厂、尺寸和客户认证须逐项核实。 [F065]
日本	SUMCO：抛光/外延/退火/SOI/再生晶圆；信越：掩模 blanks	晶圆厂家与掩模材料厂家存在不同壁垒，不能合并为“硅材料份额”。 [F066] [F020]
日本	TOK：g/i/KrF/ArF/EUV 光刻胶及配套材料	每个光源、胶种、分辨率、客户的认证独立；总光刻胶规模不等于 EUV 规模。 [F044]
日本	JX Advanced Metals：半导体金属靶材/高纯材料；味之素：ABF；三菱瓦斯化学：BT；住友电木：封装模塑材料	金属纯化、精密靶材、绝缘膜和模封配方应分产品研究。 [F022] [F024] [F046] [F056]
日本及全球网络	Ferrotec：石英、先进陶瓷、CVD-SiC 零件等	设备耗材不留在芯片里；收入随装机、开工、工艺腐蚀及替换周期变化。 [F068]

地区	公司及公开能力	进一步要拆的问题
韩国	SK 特气/材料产品体系：WF ₆ 、SiH ₄ 、Si ₂ H ₆ 、DCS 等沉积气体	此处确认化学产品能力；集团内实体、品牌及持股变更需在交易研究时重验，不在本表作集团估值归属。 [F045]
德国	Siltronic：硅晶圆及外延片；Wacker：电子级多晶硅	前者是晶圆、后者是上游高纯材料；合并产能会重复计算供应链。 [F019] [F061]
法国	Air Liquide：高纯气体、前驱体和刻蚀材料；Soitec：Photonics-SOI	气体厂与晶圆平台价值链不同；现场供气合约和罐装先进材料业务亦不同。 [F021] [F050]
欧洲及全球网络	Henkel：底填、封装与热界面材料；Element Six：CVD 金刚石/Cu-diamond	关注实际组件认证、长期可靠性和规模产能，材料优异物性本身不能证明商业优势。 [F055] [F052]

不应把“材料纯度”压成一个数字

研究时应把化学纯度、金属/颗粒污染、包装/输运、稳定供货和工艺相容性分开。Air Liquide 的先进材料方案同时包括分子、合格输运系统、超纯容器及严格批次规格；Entegris 提供从配方到过滤和监控的 CMP 组合。推论：小比例杂质、容器析出或颗粒，可能比材料主成分是否达到某个“几个九”更能解释客户更换困难。[\[F021\]](#)[\[F023\]](#)

三类可证伪的产业假设

工艺强度增加。 假设：更多三维表面/更复杂界面提高每片晶圆的材料消耗或价值量。验证：按客户节点、层数和具体工序看销量与单价，确认增长来自用量、份额还是涨价；反例：膜变薄、工艺整合、效率改善或替代分子降低消耗。此为研究假设，源于 ALD/叠层工艺的结构需求。[\[F005\]](#)[\[F007\]](#)

认证带来粘性。 假设：材料变更牵动前后工序和可靠性验证，支持稳定份额。验证：重复订单、返利后价格、客户集中度、新厂复制和新代产品胜率；反例：第二来源认证完成、配方专利到期或客户垂直整合。不能把客户名单当长期定价权。[\[F058\]](#)[\[F043\]](#)

新材料带来新价值池。 假设：玻璃、金刚石、高导热 SiC 为热与尺寸问题创造增量需求。验证：先看测试载具→送样→认证→量产→回款的阶段推进，再看单位成本和生命周期可靠性。反例：已有铜/有机基板方案继续改善，或界面/加工成本抵消物性优势。Intel 的玻璃展示、Coherent 的 SiC 送样与材料商目录不应被合并成“已形成成熟市场”。[\[F025\]](#)[\[F064\]](#)[\[F052\]](#)

07 先进封装与电测试 系统性能的另一半

先读连接拓扑 再读封装品牌

2.5D 通常把逻辑和 HBM 等裸片并排连接在高密度互连结构上；3D 则把裸片/晶圆垂直连接。一个产品可同时包含两种拓扑。TSMC SoIC 形成的三维堆叠还可进入 CoWoS 等封装，因此 SoIC 和 CoWoS 并非简单二选一。[\[F026\]](#)[\[F027\]](#)

技术/结构	如何连接	产业分工与边界
CoWoS-S	大面积硅中介层连接逻辑/HBM 等裸片	高密度横向互连；不能把中介层面积、封装尺寸与芯片有效硅面积混用。 [F027]
CoWoS-R	聚合物与铜 RDL 中介层	TSMC 官网注明 2023 年起量产；与硅中介层在工艺、布线和机械性质上不同。 [F027]
CoWoS-L	RDL 为基础，局部嵌入高密度硅互连等结构	官网注明首个 3.5 倍 reticle 尺寸产品 2024 年起量产；不是“整块超大硅中介层”的同义词。 [F027] [F063]
TSMC SoIC	晶圆级三维芯片堆叠，可 chip-on-wafer/wafer-on-wafer	选 KGD 的灵活性与同尺寸高良率晶圆堆叠适用性不同；不直接外推良率。 [F026]
Intel EMIB/Foveros	硅桥横向连接与垂直集成产品族	按具体代际区分 bridge、interposer、bump 与直接键合；公开路线图不等于每个尺寸已量产。 [F028]
Fan-out/RDL	重构晶圆或面板，在裸片外延伸布线	ASE VIPack、JCET XDFOI 有相应平台；须比较线宽间距、翘曲、芯片位移及批量良率。 [F047] [F049]
TSV	穿过硅的垂直互连	TSV 形成、减薄露出及背面金属化可由不同厂商执行；Amkor 明确其所述流程从已有盲 TSV 晶圆开始，不提供 foundry 晶圆内 TSV 成孔。 [F048]
微凸点/TCB	在热、力和时间控制下接合微连接	在适合的 pitch、芯片厚度和材料体系使用；设备名称相同不证明工艺等价。 [F030] [F051]
混合键合	介质直接键合与嵌入金属接点结合	需要表面处理、清洁、平坦度和精密放置的整体流程；不是高精度贴片机直接加一个功能。 [F029] [F058]

一条典型先进封装链

输入与筛选： 晶圆探针测试取得电性地图→挑选已知良品裸片 KGD→与中介层/基底进行工艺和测试规划。KGD 指在既定覆盖、温度、电压条件下通过测试，不是“绝对无缺陷”；多裸片集成提高了前置筛选的经济价值。[\[F034\]](#)[\[F026\]](#)

中介层与背面：按路线完成 TSV 相关加工、暂时键合承载、减薄、TSV 露出、钝化、CMP、背面 RDL 和铜柱/焊点。Amkor 公开流程提供了一个可核对的工艺边界实例：同样写“TSV 封装”，承包的工序范围可能不同。[F048]

贴装和键合：分选、取片、对准、贴装；按选定技术执行 reflow/TCB/混合键合，随后进行底填或模封等保护。薄片裂纹、翘曲、污染和对准误差会耦合；芝浦的资料强调薄片拾取及热变形校正，EVG 则把清洗和载体处理列入整体开发。[F073][F058]

系统收尾：接入封装基板、散热结构与最终外部连接，进行电测、必要的 burn-in、SLT、可靠性筛选与失效分析。TSMC/Intel/OSAT 的平台、设备商的键合机、材料商的底填及基板厂的能力共同参与，不宜把所有收入都归入“先进封装设备”。[F028][F035][F055]

公司地图 制造服务与卖设备分开

层级	代表性地区/公司	可核实能力，及不能自动推出的结论
代工与集成平台	中国台湾 TSMC	SoIC、CoWoS 等；封装平台存在不代表第三方任意组合即可免验证迁移。[F026][F027]
代工与集成平台	美国 Intel Foundry	EMIB/Foveros 及系统组装测试；比较时锁定代际与产品状态。[F028]
OSAT	中国台湾 ASE	VIPack 集成 fan-out、2.5D/3D 与 CPO 方向；技术平台不能证明某个未公开客户订单。[F047]
OSAT	美国 Amkor、亚洲制造网络	TSV 后段、2.5D、fan-out 等；明确本表来源的 TSV 前后工艺边界。[F048]
OSAT	中国大陆 JCET/长电及 STATS ChipPAC	XDFOI 高密度 fan-out；STATS ChipPAC 是其全球业务网络相关实体，活动赞助名称不能直接当独立上市股票。[F049][F069]
键合设备	奥地利 EVG；荷兰 Besi	EVG 的 W2W/D2W bonding 与 Besi hybrid/TCB 能力；按各自承担的步骤比较。[F029][F030]
键合设备	日本芝浦机电；韩国 HANMI	C2W hybrid/fusion；HBM TC bonder 等；不得把某产品的市场宣传扩张为“封装全链独家”。[F031][F051]
中国设备扩展	北方华创；拓荆	公开有键合或 3D IC 系列；工艺范围、精度及量产资格仍须独立核实。[F009][F011]

电测试的插入点与设备生态

阶段	在排除哪类风险	需要哪些组件/公司角色
晶圆参数/工艺测试	晶体管、接触、互连等制造偏差	参数仪器/SMU、探针机、探针卡；与功能 ATE 的任务不同。 [F040][F001]
晶圆功能测试、KGD	避免把明显失效裸片带入昂贵堆叠	ATE 发激励并测结果；prober 定位/控温；probe card 电接触。 FormFactor 的 HFTAP 公开用于 DRAM/HBM 类 wafer KGD。 [F032][F034]
封装最终测试 FT	检查封装后功能与电性，进行存档	ATE、handler、DIB/load board、socket/contact、热控制和测试程序共同组成 test cell。[F033][F035]
老化 burn-in	在规定应力条件下筛除或研究早期失效	电源、温度、板卡/接触和长期运行；通过不等于全寿命零失效。 [F035]
系统级 SLT	运行接近实际系统的软硬件工作负载	高功率供电、散热、真实接口、自动搬运与诊断；补充结构/参数测试，不能宣称完全替代 ATE。[F035]

Advantest V93000 和 Teradyne UltraFLEXplus 属于 SoC ATE 平台；致茂在 SLT、高功率热控、handler、接触及系统集成方面提供方案；FormFactor 提供探针卡及测量产品；东京电子等的 prober 完成晶圆定位/温控与电测对接。公司能“给 GPU 测试提供设备”和公司“运营测试工厂”是两种收入模式，不能混为同一家代测服务商。[F032][F033][F034][F035][F001]

测试为什么可能增加 也可能被优化

单颗测试成本可按“同一期间的设备/接口折旧+维护+工程+能耗+场地成本”除以同期合格器件数量估算；并行站点、测试时间和有效利用率都重要。多芯片价值提高支持更早检出缺陷，但更好的 DFT、并行化、测试程序复用与数据驱动筛选又能降低单位测试成本。这是本报告的分析框架；Teradyne 和 Advantest 官方平台都将扩展、并行和复用作为设计目标，不能只用“芯片复杂度上升”推导设备收入同比例增长。[F032][F033]

建议现场问测试企业：哪个新增失效模式必须由此插入点发现；每颗测试时间/测试功耗增长多少；并测数量受热控、供电还是引脚限制；是否能在下一代产品复用；减少过杀和漏检有何统计证据。这些指标比笼统的“AI 测试受益”更接近可验证投资逻辑。[F033][F035]

07B 光子封装测试与罗博特科案例

先把“光测试”拆开

光引擎需要连接激光器、PIC、电子芯片、微光学和光纤等部件；生产中既要检查直流/高速电性，也要检查耦合与光学性能。ficonTEC 的产品范围覆盖器件微组装、主动/被动对准、粘接/焊接和电光测试；它属于制造设备环节，不等于生产光芯片、光模块或运营数据中心。[F036][F037]

测试/工艺	解决什么问题	要看哪些实际指标
被动视觉对准	通过标记与几何校准找到初始位置	标记可见性、摄像/实际运动坐标校准、预对准耗时；其目标之一是缩小主动搜索空间。[F036]
主动光学对准	一边移动，一边测光功率等耦合反馈	低损耗位置、搜索时间、重复性，多端口/多芯片适配；包含平移与旋转等自由度。[F036]
attach 与固化/焊接	把找到的光学最佳位置固定下来	固化/热收缩导致的 post-bond shift、过程跟踪和补偿、返工率。[F036]
晶圆级电光测试	在切片/昂贵组装前获取 PIC 地图和良率信息	光电双面/边缘接入、校准、温控、端面检查、端口切换和 ATE 互操作。[F037]
高速信号测试	判断链路是否满足速率与信号质量要求	BERT/示波器等仪器与封装夹具协同；数据率、调制格式和测试点条件必须明确。[F033][F040]
可靠性与失效分析	检查温度、老化或机械扰动后能否维持性能	测试方法、样本、应力时间及判据需对应具体产品；ficonTEC 列有 LIV、光谱、beam 和 burn-in 类测试任务。[F070]

为什么“XYZ 三轴移动”不足以构成壁垒

XYZ 表示平移自由度；光学耦合常还需要角度，ficonTEC 2025 年产品说明明确提到六轴主动对准。更关键的是形成闭环：视觉预定位→激光/PIC 上电→读光信号→搜索最佳点→固定→跟踪位移→复测。运动平台分辨率只说明指令尺度，不能自动证明最终耦合精度、固化后稳定性和每小时良品产出。[F036][F037]

从公开能力可提出六项“可能的护城河”，但必须拿生产数据验证：①光/电/机械/软件协同；②面向不同光端口及材料的 recipe 和校准库；③固化位移补偿；④高速仪器、ATE、自动上下料的接口；⑤跨机器/跨批次一致性；⑥现场服务和工艺迭代。上述是研究推论，不是已确认独占专利或客户没有替代方案。[F036][F037][F058]

最有价值的验收组合是：在同一器件、同一插损分布、同一可靠性门槛下，报告有效 UPH、首次通过率、重测率、校准/维护间隔和最终成本。ficonTEC 的 wafer brochure 列有具体对准时间和耦合重复性，但仅能按其产品条件理解；本报告不把单项宣传指标横向排成“绝对领先”。[F071]

PI (Physik Instrumente) 也公开提供六自由度、多通道光子主动对准机构、控制器和算法。这个可核实的替代能力说明：六轴本身并非一家公司的专属；应比较子系统与整机在同一客户流程中的实际分工、产能和长期稳定性，不把供应商或潜在竞争者自动判成一对一整机替代。[F082]

罗博特科 确认的事实与未确认的推论

实体和交易事实。 罗博特科智能科技股份有限公司，深交所 300757，英文 RoboTechnik。2025-05-08 的独立财务顾问核查意见确认资产过户，最终直接及间接持有 ficonTEC Service GmbH (FSG) 和 ficonTEC Automation GmbH (FAG) 各 100%；2026 半年报也明确记载 2025 年 5 月完成收购。[F039][F038]

技术事实。 ficonTEC 的 2025 年晶圆测试产品提供 300 mm 双面电光测试对接：上方进行 DC/高速电测试，下方以六轴主动对准进行光探测，配套真空热控 chuck、光纤阵列校准等功能。官网“行业首创”“主要厂商采用中”属于公司声明；本次没有独立客户验收报告来验证其独家性或份额。[F037]

最新经营事实。 2026 上半年集团营收约 6.0848 亿元、归母净利约 655.68 万元；报告披露日在手未确收订单约 33.86 亿元，其中光电子及半导体业务约 24.52 亿元。在手订单以报告披露日为时点，不是 6 月 30 日订单余额，也不是当期已确认收入；集团同时有光伏业务。推论：订单到验收、回款和利润的转化比只看“CPO 概念”更关键。数字依据半年报 p7、p11，未经本报告独立审计。[F038]

投资研究假设。 假设光 I/O、CPO、硅光器件走向量产，自动化电光对准与早期筛选的单位价值可能提高。反证路径包括：客户架构采用更容易装配的无源对准；封装标准化降低调试溢价；竞争方案提高吞吐；订单延期/撤销；毛利与现金转化未跟随增长。这里是技术和商业验证清单，不是买入结论。[F036][F037][F038]

四家公司不是同一种“测试设备股”

公司	更准确的产业坐标	可比/不可比事项
罗博特科/ficonTEC	光子微组装、对准贴装、电光器件/晶圆测试	与光子封装自动化及测试对接同行比；不能直接用 GPU 电测 ATE 的所有市场作其收入上限。[F036][F037]
致茂 Chroma	SLT、高功率热控、handler/接触和 FT/burn-in 等验证生态	2026-08-24 官方发布明确此布局；本报告不把媒体所谓 NVIDIA 独家传闻扩大成所有 GPU 测试独家。[F035]
芝浦机电 Shibaura Mechatronics	高精度贴片、2.5D/3D 与混合键合、清洗等	核心比较是键合精度、洁净、薄片处理和产能；不是和光子 BERT 直接竞争。[F031][F059]

公司	更准确的产业坐标	可比/不可比事项
联讯 Semight	光通信/光芯片及半导体测试仪器与系统	本次公开目录确认有硅光晶圆测试方向；官网简介也列有 WAT、SMU、BERT 等；本次只取得搜索索引内容，未引用具体吞吐/精度、客户及份额。 [F040]

把技术故事落到可追踪指标

本模块的季度跟踪重点是新工艺导入、客户验证转量产、有效 UPH/良率、验收周期，以及并购前后同口径的光电子业务变化。通用财务指标统一见“技术壁垒与投资验证”；罗博特科还需拆分光伏业务和跨国整合，避免混用集团与光子设备口径。[\[F038\]](#)[\[F035\]](#)[\[F031\]](#)

参会时可以直接问：“你们能把什么失效模式从昂贵封装后提前到晶圆上发现？在同样良率和可靠性目标下，与客户原流程相比减少了哪一道测试、多少返工或多少设备？哪项数据可让我们证伪这一优势？”这种问法比“你们有 XYZ 吗”更能识别真实工艺能力。[\[F037\]](#)[\[F034\]](#)

08 计算架构与芯片设计

先由工作负载选择硬件

Roofline 的基本方法是同时比较计算峰值与“内存带宽 × 运算强度”。本报告进一步把网络通信、内存容量和尾延迟作为系统约束：提升其中一个上限，只有在它正好限制工作负载时才会提高有效吞吐量。这是工程分析框架，不是对某一类芯片永远优胜的判断。[\[R011\]](#)

训练要保存与更新参数、梯度和优化器状态，涉及数据并行、张量并行、流水并行或专家并行。推理的 prefill 处理提示词，decode 逐步生成 token；其瓶颈会随 batch、上下文、模型结构、并行策略和内核实现变化，不能把“prefill 计算密集、decode 带宽密集”当作无条件定律。[\[R013\]](#)[\[R014\]](#)

教学计算：70B 参数若全用 BF16 存储，权重约 140 GB，尚未计运行时、激活与 KV cache。若某假想模型有 80 层、8 个 KV heads、每 head 128 维，KV 用 2 字节、上下文 32,768、batch 8，则 KV 容量为 $2 \times 80 \times 8 \times 128 \times 32,768 \times 8 \times 2 = 80 \text{ GiB}$ 。首个 2 表示 K 与 V。权重量化改变每权重的字节数，KV 量化改变每个 KV 元素的字节数，GQA/MQA 改变 KV head 数；这些是假定结构的算术，不代表任一厂商模型。

各类处理器分工

类别	主要能力	需要哪些配套	投资研究应追问
CPU	通用控制、系统与工具执行、数据预处理	DDR、缓存、PCIe、编译器、OS	agent workflow 增加的是核心数、单核能力还是内存带宽
GPU	高并行计算与可编程矩阵运算	HBM 或 GDDR 等显存（依型号）、封装、互联、编译器与内核	同模型/SLA 下的利用率、显存占用与每瓦有效输出
TPU/自研 ASIC	对目标工作负载进行架构与系统协同	编译器、算子覆盖、集群互联和长期软件投入	芯片收益能否抵过迁移、验证和维护成本
FPGA	可重配逻辑、接口与实时数据通路	RTL/HLS、时序收敛、片上 RAM、收发器	实时性与灵活性是否比极致密度更重要
DPU/SmartNIC	网络、存储、安全等基础设施处理	NIC、可编程核和/或专用加速单元、驱动与服务框架	真正卸载多少主机开销，是否新增维护复杂度
光模块 DSP	高速链路的电信号处理	ADC/DAC 或相应模拟链路 with 均衡；FEC 按架构可在主机 PCS 或相干 DSP 等位置实现	端口速率提高后功耗、可达距离及误码要求

以上按功能划分，具体实现可以重叠。[\[R004\]\[R006\]\[R015\]\[R016\]\[R019\]\[N027\]\[N065\]\[N090\]](#) DPU 用于基础设施卸载，DSP 用于信号处理，两者应分别辨识。光 I/O 已有与计算芯片共封装或邻近集成的方案及演示，具体量产依平台核实；光模块通常不等同于一颗 CPU。[\[N037\]\[N080\]](#)

最新系统规格的正确读法

平台	截止日可核资料	证据阶段与比较边界
NVIDIA Vera Rubin NVL72	官方发布列 72 Rubin GPU、36 Vera CPU，并配 NVLink 6、ConnectX-9 与 BlueField-4	2026-03-16 称芯片进入 full production，伙伴产品从 2026 下半年提供；不能据此认定每家云和每一区域已经部署。 [R004]
AMD MI455X / Helios	产品页列每 GPU 432 GB HBM4、23.3 TB/s 理论内存带宽；Helios 为 72 GPU 参考设计	页面给出工程测算和 2026 下半年部署预期；实际整机配置及可购性由 OEM/云验证。 [R005]
Google TPU7x Ironwood	规范表列每芯片 HBM 192 GiB、HBM 带宽 7,380 GB/s、双向 ICI 1,200 GB/s	GB/s 与 Gb/s、GiB 与 GB 不能混用；ICI 与 DCN 是不同网络。规范不是自有硬件零售报价。 [R006]
AWS Trn3 UltraServer	AWS 产品页说明 Trainium3 与 UltraServer 的集成方式	比较对象应是可取得的云实例、软件栈和端到端成本；不能把云价格等同芯片制造成本。 [R015]
Cerebras WSE-3 / CS-3	官方 WSE-3 材料列 900,000 核、44 GB 片上 SRAM	晶圆级系统的内存层级与 GPU 不同；片上容量、外部参数存储和可运行模型规模不能混为一谈。 [R007]

不同文件可能修订规格，例如某些 Rubin 数据表与技术博客对聚合带宽的数字或口径不同。本报告不挑最高数字拼成一个“超级配置”；使用某个型号的规格时，应保存具体文档版本，并向供应商取得最终配置表。[\[R004\]\[R032\]](#)

供应链中的设计工作

芯片从需求进入架构、RTL/IP 集成、功能验证、综合、布局布线、时序/功耗/物理验证，再到 tape-out；封装、电源和热设计需要更早协同。EDA 是设计与验证工具，foundry PDK 描述可制造规则，IP 是可集成的电路设计资产，三者都不能替代最终硅验证。[\[R017\]\[R018\]\[R019\]](#)

Synopsys DSO.ai 使用 AI 探索设计空间并优化 PPA；Cadence Cerebrus 提供设计流程优化；Siemens Calibre PERC 的核心是电路与版图相关的可靠性验证。它们针对不同任务，不能用“会生成 RTL”统一评价。判断 AI EDA 价值时应固定设计、工艺、约束和人工基线，衡量达到 signoff 的时间、质量与重现性。[\[R017\]\[R018\]\[R019\]](#)

08B 计算与设计公司地图

参会企业的架构路线

公司	能力与当前公开状态	关键验证问题
Cerebras 美国	晶圆级计算；WSE-3 把大量计算与 SRAM 放在单一晶圆级芯片中。 [R007]	真实模型的端到端延迟、并发与外部数据搬运；故障绕行和集群可用性如何
SambaNova 美国	RDU 以可重构 dataflow 配合 SRAM/HBM/DDR 层级；论文解释 SN40L 的内存与执行方式。 [R008]	编译覆盖、动态模型支持、模型切换和实测功耗；不用旧型号推定最新产品
d-Matrix 美国	2026-06-09 宣称 Corsair full production；资料列 TSMC N6、Alchip、SRAM DIMC、有机基板和 LPDDR5。 [R009]	特定模型下低延迟如何与吞吐、内存容量及 GPU 配对成本共同变化
Etched 美国	当前官网描述 HBM/SRAM hybrid、LVI 与 CSM，称 A0 已回片并在验证首个机架产品。 [R010]	索取截止日可用硬件、模型名单、独立测试和实际客户验收；不沿用 2024 Sohu 宣传数字
Rebellions 韩国	当前官网列 Rebel100、RebelServer、RebelRack、RebelPOD 与 SDK 入口，并宣称生产部署。 [R022]	模型/算子覆盖与客户持续运行证据；“最盈利推理”应拆为具体成本口径
SiMa.ai 美国	Modalix 面向 edge/physical AI；产品页有芯片、SoM、PCIe 卡与开发套件。 [R023]	整个感知到动作应用的功耗和延迟；开发板、设计导入与量产收入分别验证
Normal Computing 美国	官方 ASIC 页面仍以 CN101 及 CN201/CN301 路线说明物理随机/热力学计算。 [R024]	采样问题的精度与功耗是否包含控制、读出和主机；路线图不是通用 GPU 替代证明
Unconventional AI 美国	官方定位强调硬件与算法共同演化，公开首页尚不足以形成可比较的量产规格。 [R025]	用何种物理原语处理何类运算，校准、精度、接口与软件如何实现
Ricursive Intelligence 美国	官方定位是 AI for chip design 与 chip design for AI 的循环。 [R026]	现阶段交付模型、设计工具还是芯片；有无可复现 PPA 和 signoff 案例
Cognichip 美国	官方 ACI 定位为 physics-informed foundation model 支持芯片设计。 [R027]	训练数据授权、验证闭环、工艺迁移和客户 IP 隔离；营销承诺不能当硅成功率
Altera 美国	FPGA 以可重配逻辑、片上存储和高速接口服务不同系统需求。 [R016]	工具成熟度、时序收敛、长期供货、设计导入后的实际出货

这些路线的共同问题是数据移动与可编程性的取舍。SRAM 的访问方式、dataflow、专用 ASIC 或物理采样可以改善特定任务，但“低能耗原语”到“完整服务更便宜”之间仍有精度、软件、外部内存、网络和利用率约束。这是基于上表的研究推论，不是各公司的统一量产结论。

全球代表性计算供给

产业位置	美国	中国大陆	其他地区与边界
通用与加速计算平台	NVIDIA、AMD、Intel；资料分别覆盖 GPU/rack、Instinct、Xeon。 [R004][R005][R033]	服务器 CPU 有海光、鲲鹏；加速器有华为昇腾、寒武纪、摩尔线程、昆仑芯、壁仞。产品类别不证明节点、产量或客户份额。 [R058] [R059][R034][R035][R036] [R037][R038]	Arm 英国提供 Neoverse/CSS，并已扩展至自有 CPU 产品；不能再简单写成只卖 IP。 [R028]
云自研加速器	Google TPU、AWS Trainium；入口分别为云平台服务。 [R006] [R015]	华为等整栈能力见官方资料；其他云的未披露设计不填补客户关系。 [R034]	Rebellions 韩国是另一路推理芯片/系统供给；不能视作同类云实例。 [R022]
CPU 与接口 IP	Synopsys 等提供接口/IP 及设计工具。 [F017]	“自主”不等于全部 IP、工具、材料国产；实际边界按芯片与工具合同逐项核实	Arm/CSS 位于架构、CPU 核心与预集成子系统层；实际芯片由不同设计与制造链实现。 [R028]
EDA 细分能力	Synopsys、Cadence 的设计优化见上一模块；不同工具不能混用同一产效指标。 [R017][R018]	华大九天 ALPS 为 SPICE 仿真；概伦 NanoSpice Pro 为 FastSPICE；广立微 SmtCell 为参数化测试结构版图，不能将单工具写成完整流程替代。 [R061][R062][R063]	Siemens Calibre 等工具在物理/可靠性验证有不同角色；总部归属不代表工具只在单一地区研发。 [R019]

地区列表是研究入口，涵盖代表性公开产品，而非每一家都在同一代制程、同一软件成熟度或同一出口可用性下竞争。要判断替代能力，必须在同一模型、质量、SLA、软件与实际取得成本上做系统比较。

基准结果审查卡

每张性能图至少记录：型号和芯片数；模型版本与权重精度；输入/输出长度；batch 与并发；TTFT、TPOT 和 P99；吞吐是否受 SLA 约束；功率是否实测；软件/内核版本；是否包含主机、网络和存储。MLPerf 数据中心规则区分测试场景，也明确只有规定的功率测试才构成其验证功率，TDP 或电源额定值并不等同于实测能耗。
[R012]

若厂商只提供最高 tokens/s，现场可直接追问：“在我们约定的输入长度、输出长度和 P99 延迟上限下，整机墙插功率是多少；模型质量相同吗；运行多久；我们能复跑吗？”这比对照宣传倍数更能识别工程成熟度。

09 CASPA 会议与公司对应

活动基线

活动官网列日期为 2026 年 10 月 23–24 日，地点 Santa Clara Convention Center。以下为 10 月 15 日官网公布安排，属于预告；议程和出席者仍可能调整。本报告采用当前官网，不用较早社交宣传替代当前职务或上台名单。[R001][R002]

最值得带着问题听的场次

当地时间	场次与公司入口	本报告建议追问
10/23 13:40–14:00	Credo 的 \$100/Watt	每瓦的定义是 IT、网络还是设施输入；价格和持续服务价值如何区分
10/23 14:00–14:40	半导体展望，Bernstein、SemiAnalysis，Chipstrat 主持	哪个瓶颈已反映在价格，哪个仍会推迟真实交付
10/23 14:40–15:30	Lux Capital、Samsung Ventures America、Salience Capital、Seligman Ventures，Synopsys 主持	从研发到客户验收需要几轮融资，资本承诺如何覆盖周期
10/23 15:30–16:20	Cognichip、Rebellions、Upscale AI/Velaura AI	AI 设计工具、推理芯片与网络系统分别如何证明技术价值
10/23 16:20–17:10	SiMa.ai、Schneider Electric、d-Matrix	边缘与数据中心负载分别受何种电、热和延迟约束
10/24 09:55–10:15	Synopsys	AI EDA 到 signoff 的完整闭环；跨 die 热与供电如何验证
10/24 10:30–11:10	Solidigm、Sandisk、ScaleFlux，Temasek 主持	AI 存储的 QD、P99、耐久、掉电保护与系统接入瓶颈
10/24 11:10–11:50	Unconventional AI、Normal Computing、SambaNova	新原语改变哪一段工作负载；控制和读出是否计入功耗
10/24 13:00–13:20	Lumentum	EML、CW laser、CPO/ELS 的产能与可靠性验证怎样分别推进
10/24 13:20–13:40	Fortinet	自研安全 ASIC 的系统性能和软件订阅如何共同产生价值
10/24 13:40–14:00	Western Digital	HDD 容量、重建时间与数据生命周期；怎样与 SSD 协同

当地时间	场次与公司入口	本报告建议追问
10/24 14:00–14:30	Altera、Multibeam	FPGA 的部署边界与电子束技术的具体工序，不把创始人背景当当前产品
10/24 14:40–15:20	Ricursive Intelligence、Etched、Siemens	自动设计与新架构分别怎样通过真实硅和软件验证
10/24 15:20–16:00	Silith、Ayar Labs、Semtech	光 I/O、激光、封装与模块的边界；多少成本在光学对准和测试
10/24 16:00–16:30	Cerebras	晶圆级架构在实际模型、并发和可用性上的优势与限制
10/24 19:45–20:45	Intel 晚宴对谈	产品设计、foundry 与先进封装业务如何分别验证执行能力

时间与参与机构来源为官网议程；右列为本报告提问建议。[\[R001\]](#) SEMI、IEEE、GSA 另有产业、标准与协作主题，可用于询问互操作、可靠性与人才，而不将其当成器件供应商。[\[R002\]](#)

活动公司在产业链中的位置

主讲公司的技术资料分别进入计算、存储、光互联、能源与 EDA 章节。另需明确：Fortinet 的安全处理器属于安全与网络处理，不等同 GPU 后端 fabric；Rambus 的接口 IP/内存接口相关能力位于设计和连接层；Silvaco 处于半导体设计工具/技术平台生态。官方产品资料及证据范围详见来源附录，未核查的产品细节不从活动 logo 推导。[\[R039\]](#)[\[R040\]](#)[\[R041\]](#)

当前嘉宾页面给 Solidigm 的上台代表是 **Avi Shetty**，Vice President，Upscale AI/Velaura AI 的代表是 **Rajiv Khemani**，分别为 Executive Chairman/CEO。这些与较早宣传文案不同，现场优先核对。[\[R002\]](#)

赞助与展台名录

以下仅确认活动网页上出现该品牌，不据此认定其与某大厂有供应、投资或认证关系。[\[R003\]](#)

网页分类	品牌名称	产业研究处理
Event Sponsors	STATS ChipPAC、Solidigm	封装测试与存储，正文对应模块展开
Table Sponsors 已可识别大型/技术品牌	Meta、Western Digital、Siemens、Credo、Rambus、Silvaco、Normal Computing、Silith、TetraMem、Multibeam	分别为需求方、存储、设计工具、连接、IP、计算/光学等入口；小型企业仍需官网/法律实体核验
Table Sponsors 需进一步实体核实	Infraeo、DeepWeave、Urbot、Architect Labs、SufasTech、OpenNetwork、WGTech、Far East、Himalaya、T3W、Futurus、YouAndI、Playbook / OpenExt	保留官网原字符串；同名企业较多，尚不能确定对应的法律实体或产品线

网页分类	品牌名称	产业研究处理
Display Tables	Western Digital、EY、TGV Tech、Architect Labs、PuppyGraph、Phantafield、Wiskind、UpMarket	不把展台身份当成 AI 硬件供应商证明；未核准实体作为现场核查列表

活动还出现 CASPA、IEEE、GSA、GPT.DAO、VIP Club、CEO、SESE、PopPopTech、CIE-SF 等联盟或社区名称；讲者/主持机构还包括 UC Berkeley、San Jose 市、Silicon Catalyst、Reuters、SemiWiki、Semiconductor Engineering、EE Times、ABC7、Vista Robotics 与 Silicon Valley Vector。[R002][R003] 它们在报告中承担组织、知识、投资或媒体入口，不能放进芯片设备竞争表。

对新创公司的三句尽调

第一句：“你卖的是芯片、模组、系统、软件许可还是服务，客户在哪里验收？”第二句：“目前最难通过的性能或可靠性门槛是什么，哪些数据能公开复核？”第三句：“量产增长先增加设备、材料、工程师还是营运资本？”先确认这三点，再讨论融资、估值或市场规模，可以避免把研发项目误判为成熟供给。

10 技术壁垒与投资验证

06 技术难点怎样转化成投资验证

每个箭头都是尚需证据的跃迁；技术领先不自动推出利润增长，更不自动推出当前价格有吸引力。



AI 基础设施产业研究 | 截止 2026-10-15 | 本报告原创示意图 | 产品阶段与企业证据见正文及来源附录

技术壁垒的投资验证链

从工程难点到经济回报

本节提出可证伪研究假设，不给没有估值与现金流模型支撑的买入排序。投资判断至少经过五步：确认目标工艺/系统中的不可缺部件；证明供应商达成指标；证明客户长期使用且替换成本高；证明收入与现金可兑现；再看当前价格是否已经透支预期。前一步成立不能自动证明后一步。

技术护城河可以来自精度、材料知识、工艺窗口、复杂集成、软件与数据，也可能来自规模良率、维修网络和长期认证。评估时间“竞争者能否以同样成本、良率、吞吐量和稳定性达到同样指标”，比“能否做出样机”更接近商业竞争。这是本报告的评估方法；光子装配与测试的具体能力和验收边界见相关产品资料。

[F036][F037][F071][F082]

七类值得深入的假设

假设	候选产业入口	成立所需证据	可能推翻它的情况
高带宽互联使光学制造瓶颈外溢到设备	ficonTEC/罗博特科、光测试与精密装配同行	客户产线的重复订单、耦合良率、UPH、稳定性与售后收入	另一工艺显著减少对准次数；客户内制；替代设备达成同样成本
先进封装的瓶颈向前后工序扩散	键合、减薄、切割、翘曲检测、探针与 ATE	具体节点/封装的产能利用率、交期、客户认证与单位收入	HBM/logic 需求下降；封装架构改变；重复扩产造成供过于求
功率密度提高增加电力与热系统价值量	变压器、母排、PSU、VRM、CDU、冷板、连接器	已送电项目、系统验收、可靠性和每 MW 实际订单	交付推迟；标准化压价；客户总功率上限限制装机
存储价值由容量转向可预测服务	企业 SSD、控制器、固件、NVMe-oF 与缓存层	真实读写混合、P99、写放大、耐久、故障恢复及客户回款	数据复用减少需求；介质价格上涨侵蚀毛利；软件绕过性能价值
材料更换和结构升级创造单位工艺价值	ALD/刻蚀/CMP、高纯前驱体、膜层/键合材料	新材料量产导入、单位用量、认证时间与良率收益	材料用量降低；替代化学路线；仅有送样没有采购
定制计算通过减少数据移动改善 TCO	TPU/ASIC、SRAM/dataflow/近存计算	同模型质量、SLA、墙插功率与软件维护成本	模型结构变化；编译覆盖不足；外存和网络成本吞掉收益
互连速率提升增加测试的经济价值	BERT、光电晶圆测试、探针、热控与可靠性	在可接受误判/漏判下缩短测试时间并提高出货良率	测试时间随集成下降；客户合并工位；产品周期短于投入回收期

表中为本报告综合推论；支撑工程事实见对应模块。它不证明每家候选都拥有稀缺产能、已进入头部客户或具备估值吸引力。

订单怎样变成现金

研究应将意向/框架协议、确定订单、已交付、客户验收、会计收入、回款分列。长周期设备还应看合同负债、存货与发出商品、应收账款、验收条款、质保及取消风险。不能用订单总额乘毛利率就得到当期利润；跨期订单也不能与单季收入直接比较。Seeking Alpha 公开讨论中对毛利和库存的关注，适合转化成这类逐项核查，不适合直接继承目标价。[\[R021\]](#)

从设备价格到客户价值，可用**教学模型**量化：假设每年处理 100 万个组件、单组件投入 100 美元，设备改善后的最终合格率从 90% 到 92%，在其他条件相同且全部新增合格品可售时，多产生 2 万个合格件。是否值得购买还取决于新增毛利、设备折旧、耗材、工程维护、节拍、停机及原有设备替代成本。这里数字是演算假设，不是 ficonTEC 或任一客户实绩。

财务与估值所需最小表

维度	应取数据	必须区分
增长质量	分业务收入、客户集中、订单取消、复购、现金转换	行业景气与公司份额；出货与验收
盈利能力	产品组合、GAAP 与调整后毛利、研发、质保、服务占比	周期涨价与持久壁垒；并购会计与经营改善
资本强度	CapEx、营运资本、产能爬坡、折旧、自由现金流	建设支出与维持支出；会计年限与经济寿命
竞争风险	同类验证、客户双供、内部替代、架构迁移	独家工序与整体独家；产品兼容与实际导入
估值	市值、净负债、股权稀释、周期化盈利、情景现金流	好公司、增长快与当前价格便宜是三个判断

本表提供投资分析的核验维度。个股目标价需要经过核实的股价、分部财务和估值模型；产业技术优势不能直接转换为全行业供应商的统一投资排名。

组合风险与反向观察

看起来分散的 GPU、HBM、封装、光模块、电源和机房公司，可能共同依赖同一批 hyperscaler 的资本预算，相关风险并未因持有多个细分板块而消失。应检查需求确认是否来自同一份数据中心计划，避免把上游采购和下游融资重复算作独立终端需求。这是本报告的供应链风险推论。

更有价值的领先问题包括：新机房真正通电还是仅公布规划；存货增长对应可验收订单还是提前备料；单位系统所需端口/测试工时/材料用量是否变化；终端价格下跌会先伤害谁的毛利；新一代产品是否使旧库存折价。回答这些问题，需要季度更新，不靠一张静态产业图完成。

11 社区异议与深入阅读

如何筛选论坛评论

本报告按“机制具体、口径明确、可被独立证据检验、承认边界、回应反证”筛选评论；不把点赞数、自称职业或盈利截图当专业资质。能够把争议转为可测量问题的评论，更适合进入研究框架。匿名用户的客户、业绩、价格和份额主张，未被一手材料确认前仍是未核实陈述。

讨论及作者	值得吸收的内容	可检验的产业问题	不采纳的部分
r/ ValueInvestin g , Educational- Play-465 , 2026-08 讨论	将维持现有产能的支出与新增 产能支出分开，再检验现金回 报	在财务模型中分拆 sustaining/growth CapEx；会计折旧年限与经济回收分 别建情景。[R029]	线程内对所有 GPU 统一寿命或 残值的断言；没有统一可验证实 据
r/ chipdesign , saw235 , 2026-10 讨论	具体提到波形 debug 和报告生 成，提供了 AI EDA 可测试的 工作环节	现场询问覆盖率、误报/漏报、人工 复核和最终 signoff；该体验只代表 匿名个人自述。[R030]	“结论很少错”未给样本数/对 照/设计复杂度，不推为行业成 功率
r/ LocalLLaMA , 2024 Sohu 发布讨论	对架构专用化、模型未来变 化、基准精度提出不同意见	以当前官方产品资料重设比较边界， 追问模型集合与迁移成本。[R031] [R010]	老宣传的最高倍数及一台机器替 换多少 GPU，不作为 2026 实际 部署数据

以上是对研究有用的评论，不是已证实的高资历专家判断。技术区可优先跟踪 chipdesign、FPGA、networking、sysadmin、LocalLLaMA、Semiconductors；投资区可看 ValueInvesting、stocks 和相关公司讨论区。每次仍须按上述规则逐条筛选，单一社区往往具有职业或持仓选择偏差。

研究机构的方法借鉴

SemiAnalysis 的公开产品说明把 AI TCO、网络和数据中心容量模型分开，再以配置、连接比例、功率、建设进度和财务回报联系。相关分层方法适合检验配置、物理约束与经济回报的联系。本文算例与分类不代表该机构的付费容量、客户份额或价格模型。[R020]

Seeking Alpha 的公开 podcast 文稿把行业主题带回毛利、存货、订单和 ROI。这提供了技术观察到财务验证的桥梁；作者的荐股、估值与宏观判断仍是观点。本文只吸收问题结构，不采纳其特定收益预期。[R021]

带着问题看的视频与播客

资源	对应章节与建议片段	资料用途与证据范围
Dwarkesh Podcast , Asianometry 与 Dylan Patel , 2024-10-02	38:51 产业分层 ; 49:21 制造复杂度 ; 59:13 架构 ; 1:16:24 成本与电力 ; 2:06:10 产业机会	产业分层与制造经济性的延伸阅读 ; 历史产业数据不作为 2026 事实。[R042]
Branch Education 的 SSD/NAND 系列	用 3D 结构理解 cell、string、page 与控制栅 , 再回读本报告 NAND 工艺	结构学习的延伸阅读 ; 具体器件与工艺结论以本报告引用的一手资料为准。[R043][R044]
Google Cloud Introducing 7th Generation TPUs Ironwood	将芯片、pod、ICI 与数据中心网络的层级对应起来	官方视频作为入门补充 , 数值以当前 TPU7x 规范为准。[R045][R006]
SiMa.ai、d-Matrix、SambaNova 官方技术材料	观察 edge pipeline、存内计算、dataflow 如何具体改变数据移动	以文字技术资料与论文作为主证据 ; 演示中的收益必须复核测试条件。[R008][R009][R023]

TED 风格讲座适合建立直觉，设备、材料和量产壁垒更适合读技术论文、标准、产品手册和制造商技术分享。本报告因此以这些原始资料建立论据，而把视频作为有明确学习目标的补充。

现场记录应留下的证据

每次交流至少记下产品型号、阶段、工作负载、测试条件、客户验收指标和可公开资料链接。对不能公开的客户名称，记录其不公开的边界，不用猜测补全。会后将“某人说”“官方资料写”“自己推导”分别归档，再更新相应模块与来源编号，产业图才有持续研究价值。

来源附录

编号对应正文引用。资料检索日期为 2026-10-15；未列发布日期的资料，其日期尚未确认。read 表示引用依据来自来源正文，search-excerpt 表示证据仅限摘要。证据范围针对具体论据，不代表整份资料的全面审计。

[F001] [Tokyo Electron semiconductor production process products](#)

Tokyo Electron | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Deposition lithography etch cleaning wafer prober flow

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位；先进封装与电测试：系统性能的另一半

[F002] [Semiconductor manufacturing products](#)

Applied Materials | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Deposition removal implant thermal analysis integration product categories

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位

[F003] [Light and lasers lithography principles](#)

ASML | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 KrF ArF EUV wavelengths and tin droplet plasma

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位

[F004] [Backside Power](#)

Applied Materials | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Backside routing definition process requirements; historical all-frontside statement excluded

用于 晶圆制造：把电路图变成三维器件

[F005] [Atomic Layer Deposition](#)

ASM | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Surface controlled conformal film deposition high-k metal gate products

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位；材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F006] Epitaxy

ASM | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Si SiGe SiP growth strain selective epitaxy

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位

[F007] Cryogenic Etching

Lam Research | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 3D NAND deep holes oxide nitride stacks metallization wafer stress and Cryo target

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位；材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F008] 2025 annual report semiconductor process control

KLA | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Wafer reticle packaging chemical inspection metrology and yield control

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位

[F009] 产品与服务

北方华创 | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Public product families etch PVD CVD implant RTP epitaxy bonding

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位；先进封装与电测试：系统性能的另一半

[F010] 2025 company factsheet

中微公司 | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 CCP ICP TSV MOCVD products and thin film development scope

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位

[F011] Product portfolio

拓荆科技 | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 PECVD ALD gap fill and 3D IC product families

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位；材料地图：同样叫“绝缘”，实际承担不同任务；先进封装与电测试：系统性能的另一半

[F012] 晶圆制造清洗产品

盛美上海 | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Post CMP silicon and SiC substrate cleaning single wafer cleaning

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位

[F013] Semiconductor manufacturing equipment

Ebara | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 CMP dry vacuum pumps abatement ozone water and fab support roles

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位

[F014] Integrated semiconductor vacuum and exhaust solutions

Edwards | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Integrated pumps abatement lines water signals maintenance and cost

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位

[F015] Pulsed Radio Frequency solution for critical processes

MKS | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Pulsed RF response as a nonplanar semiconductor process control problem

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位

[F016] High Performance Vacuum Valves and Vacuum Systems

VAT | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Company vacuum valve product positioning only

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位

[F017] Synopsys and Intel Foundry accelerate advanced chip designs with IP and certified EDA flows for Intel 18A

Synopsys | 发布 2024-02-21 | 检索 2026-10-15 | official | search-excerpt

证据范围 Certified EDA IP design ecosystem and multi-die design role

用于 晶圆制造：把电路图变成三维器件；计算与设计公司地图

[F018] Applications of Multicolumn Electron Beam Lithography

Multibeam | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Maskless advanced packaging photonics high mix applications; marketing multipliers excluded

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位

[F019] Epitaxial wafers

Siltronic | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Additional monocrystalline layers on polished wafers and customer specified properties

用于 晶圆制造：把电路图变成三维器件；材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F020] Photomask blanks

Shin-Etsu Chemical | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Blank material as photomask base and pattern template

用于 晶圆制造：把电路图变成三维器件；材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F021] Electronics Advanced Materials

Air Liquide | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 ALD CVD precursors etchants purity canisters qualified delivery systems

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F022] Semiconductors and Sensors materials

JX Advanced Metals | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Semiconductor interconnect metals sputtering targets solder related materials

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F023] Chemical Mechanical Planarization

Entegris | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Slurry pads filters brushes monitoring and defect management

用于 晶圆制造：把电路图变成三维器件；材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F024] [Ajinomoto Build-up Film innovation story](#)

Ajinomoto | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 ABF substrate insulation formulation low expansion and processing

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F025] [Intel unveils glass substrates](#)

Intel | 发布 2023-09-18 | 检索 2026-10-15 | official | read

证据范围 Glass substrate test chip development and roadmap announcement

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F026] [TSMC SoIC](#)

TSMC | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 3D stacking KGD heterogeneous integration and integration with CoWoS

用于 先进封装与电测试：系统性能的另一半

[F027] [CoWoS](#)

TSMC | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 CoWoS S silicon R RDL L local silicon variants and stated volume production years

用于 先进封装与电测试：系统性能的另一半

[F028] [Advanced Packaging Innovations](#)

Intel | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 EMIB Foveros advanced packaging platform roles and generation distinctions

用于 先进封装与电测试：系统性能的另一半

[F029] [Bonding](#)

EV Group | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Hybrid fusion temporary permanent die to wafer bonding

用于 材料地图：同样叫“绝缘”，实际承担不同任务；先进封装与电测试：系统性能的另一半

[F030] Scientific publications packaging bonding

Besi | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 TCB and hybrid bonding routes and process context

用于 先进封装与电测试：系统性能的另一半

[F031] Semiconductor manufacturing equipment and hybrid bonder

Shibaura Mechatronics | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Cleaning etching bonding C2W hybrid and thin chip handling roles

用于 设备公司地图：按工艺能力定位；先进封装与电测试：系统性能的另一半；光子封装测试与罗博特科案例

[F032] V93000 SoC test systems

Advantest | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 ATE electrical stimulus measurement scalability reuse and test architecture

用于 先进封装与电测试：系统性能的另一半

[F033] UltraFLEXplus SoC tester

Teradyne | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 SoC ATE instruments interfaces parallelism and test reuse

用于 先进封装与电测试：系统性能的另一半；光子封装测试与罗博特科案例

[F034] HFTAP series probe cards

FormFactor | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 DRAM HBM KGD wafer electrical probe cards and limits

用于 先进封装与电测试：系统性能的另一半；光子封装测试与罗博特科案例

[F035] Chroma semiconductor validation ecosystem SEMICON Taiwan 2026

Chroma | 发布 2026-08-24 | 检索 2026-10-15 | official | read

证据范围 SLT FT burn in thermal control contact handling system integration

用于 设备公司地图：按工艺能力定位；先进封装与电测试：系统性能的另一半；光子封装测试与罗博特科案例

[F036] Capabilities for photonics assembly and testing

ficonTEC | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Passive active alignment calibrated vision optical feedback post bond shift process control

用于 设备公司地图：按工艺能力定位；光子封装测试与罗博特科案例；技术壁垒与投资验证

[F037] ficonTEC releases wafer level test cell to complement existing ATE

ficonTEC | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 300mm double sided electro optical testing six axis active alignment and ATE docking

用于 光子封装测试与罗博特科案例；技术壁垒与投资验证

[F038] 罗博特科 2026 年半年度报告

罗博特科/巨潮资讯 | 发布 2026-08-26 | 检索 2026-10-15 | official | read

证据范围 Identity acquisition H1 revenue profit and orders at report disclosure date pages7 9 11

用于 光子封装测试与罗博特科案例

[F039] 罗博特科标的资产过户情况独立财务顾问核查意见

东方证券/巨潮资讯 | 发布 2025-05-08 | 检索 2026-10-15 | official | read

证据范围 Completed asset transfer FSG FAG each100 percent ownership pages18 19

用于 光子封装测试与罗博特科案例

[F040] Semight company profile and product portfolio

Semight | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 BERT、采样示波器、SMU、WAT、硅光晶圆及 SiC KGD 测试能力；证据限于公开摘要

用于 先进封装与电测试：系统性能的另一半；光子封装测试与罗博特科案例

[F041] National Silicon Industry Group products

NSIG | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 300mm silicon SOI and customized wafer company network only

用于 材料公司分布与投资验证

[F042] [华特气体中芯宁波优秀供应商新闻](#)

华特气体 | 发布 2025-06-20 | 检索 2026-10-15 | official | search-excerpt

证据范围 Company stated specialty gases lithography mixtures and certifications; not independently verified

用于 材料公司分布与投资验证

[F043] [安集科技解决方案](#)

安集科技 | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 CMP slurries Cu barrier W dielectric substrate TSV hybrid bonding post etch chemistries

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F044] [Semiconductor manufacturing field](#)

Tokyo Ohka Kogyo | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Photoresist classes EUV ArF KrF and auxiliary chemical products

用于 晶圆制造：把电路图变成三维器件；材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F045] [Deposition gases](#)

SK materials product website | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 WF6 SiH4 Si2H6 DCS deposition gas capabilities; entity ownership not inferred

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F046] [BT materials presentation](#)

Mitsubishi Gas Chemical | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 BT resin chemistry bismaleimide triazine and package substrate use

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F047] [VIPack advanced packaging](#)

ASE | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Fan out RDL TSV 2.5D 3D CPO platform; client wins not inferred

用于 先进封装与电测试：系统性能的另一半

[F048] 2.5D 3D TSV packaging

Amkor | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 TSV reveal thinning temporary bond backside metallization; explicitly excludes foundry TSV formation

用于 先进封装与电测试：系统性能的另一半

[F049] JCET XDFOI advanced packaging

JCET | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Ultra high density fan out and heterogeneous integration platform

用于 先进封装与电测试：系统性能的另一半

[F050] Photonics SOI substrates

Soitec | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Silicon on insulator photonics platform and data communication use

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F051] HANMI SEMICON Taiwan 2026 exhibitor entry

SEMI/HANMI | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Manufacturer stated HBM TC bonder capability; self asserted market rank excluded

用于 设备公司地图：按工艺能力定位；先进封装与电测试：系统性能的另一半

[F052] Synthetic diamond semiconductor solutions

Element Six | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 CVD diamond and copper diamond thermal product availability; adoption not inferred

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F053] Scaling AI Data Center Power Delivery with Si SiC and GaN

Infineon | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 8kW reference PSU mixed Si SiC GaN PFC and DC DC stages

用于 材料地图：同样叫“绝缘”，实际承担不同任务

[F054] Lenses and mirrors lithography principles

ASML | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 DUV lens and EUV mirror optical architecture distinction

用于 晶圆制造：把电路图变成三维器件；设备公司地图：按工艺能力定位；材料地图：同样叫“绝缘”，实际承担不同任务

[F055] High tech underfill solutions

Henkel | 发布 2024-11-08 | 检索 2026-10-15 | official | read

证据范围 Underfill TIM liquid molding lid and stiffener attach material roles

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证；先进封装与电测试：系统性能的另一半

[F056] IT materials product FAQ

Sumitomo Bakelite | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Epoxy molding compounds for semiconductor encapsulation

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F057] How will GAAFET affect ALD intensity discussion

Reddit r/Semiconductors | 发布 未确认 | 检索 2026-10-15 | forum | search-excerpt

证据范围 Anonymous discussion of ALD speed quality and batch tradeoffs; hypotheses only

用于 设备公司地图：按工艺能力定位

[F058] Die to Wafer bonding white paper

EV Group | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Bonding development needs materials cleaning carriers and adjacent processes

用于 设备公司地图：按工艺能力定位；材料公司分布与投资验证；先进封装与电测试：系统性能的另一半；光子封装测试与罗博特科案例

[F059] Advanced Package Bonder TFC 6500 W

Shibaura Mechatronics | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Accuracy given with local alignment room temperature 3sigma and dry UPH conditions

用于 设备公司地图：按工艺能力定位；光子封装测试与罗博特科案例

[F060] [Interconnect materials engineering](#)

Applied Materials | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Low k dielectric and RuCo barrier liner as specific wiring integration paths

用于 材料地图：同样叫“绝缘”，实际承担不同任务

[F061] [Wacker Polysilicon annual report2024](#)

Wacker | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Semiconductor grade versus solar grade polysilicon distinction and high purity capacity

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F062] [AXT substrate products](#)

AXT | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 InP GaAs Ge VGF substrate and raw materials capability; market rank excluded

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F063] [TSMC2025 annual report](#)

TSMC | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 CoWoS L RDL embedded local silicon interconnect; reserve contextual source

用于 先进封装与电测试：系统性能的另一半

[F064] [Coherent customer sampling300mm high thermal conductivity SiC](#)

Coherent | 发布 2026-08-17 | 检索 2026-10-15 | official | read

证据范围 Customer sampling announcement dated2026-08-17; not volume adoption

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F065] [GlobalWafers products](#)

GlobalWafers | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Annealed compound SOI epitaxial ingot diffusion polished wafer product families

用于 材料公司分布与投资验证

[F066] SUMCO product lineup

SUMCO | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Polished epitaxial annealed SOI reclaimed wafer structures

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F067] Hemlock markets and technologies

Hemlock Semiconductor | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Polysilicon business capability only

用于 材料公司分布与投资验证

[F068] Ferrotec products and technologies

Ferrotec | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Quartz advanced ceramics CVD SiC parts and related equipment consumables

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F069] JCET2025 interim report

JCET | 发布 2025-08-20 | 检索 2026-10-15 | official | search-excerpt

证据范围 公开摘要列示 STATS ChipPAC Management 与 JCET STATS ChipPAC Korea 的子公司关系

用于 先进封装与电测试：系统性能的另一半

[F070] Photonics assembly and testing product overview

ficonTEC | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Align attach UV epoxy soldering welding LIV burn in beam spectral tests

用于 光子封装测试与罗博特科案例

[F071] Wafer Testline brochure

ficonTEC | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Wafer optical alignment coupling repeatability under vendor brochure conditions

用于 光子封装测试与罗博特科案例；技术壁垒与投资验证

[F072] AXT global operations and company overview

AXT | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 US headquarters Tongmei manufacturing operations in China

用于 材料地图：同样叫“绝缘”，实际承担不同任务；材料公司分布与投资验证

[F073] Shibaura semiconductor solutions

Shibaura Mechatronics | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Thin die pickup deformation compensation and assembly process tradeoffs

用于 先进封装与电测试：系统性能的另一半

[F074] Ebara products widely used in semiconductor factories

Ebara | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Cleanroom subfab machine room rooftop support and fluid connections

用于 晶圆制造：把电路图变成三维器件

[F075] Semiconductor related inspection products

Lasertec | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 EUV DUV mask blank pellicle and wafer inspection categories

用于 设备公司地图：按工艺能力定位

[F076] DISCO product and technology lineup

DISCO | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Dicing saw laser saw grinders polishers blades wheels

用于 设备公司地图：按工艺能力定位

[F077] SCREEN semiconductor product lists

SCREEN | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Wafer cleaning track annealing metrology and direct package patterning

用于 设备公司地图：按工艺能力定位

[F078] [ZEISS SMT products and technologies](#)

ZEISS | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Precision lithography optical systems upstream role

用于 设备公司地图：按工艺能力定位

[F079] [Electron beam mask writer](#)

NuFlare | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Writing reticle patterns distinct from direct wafer patterning

用于 设备公司地图：按工艺能力定位

[F080] [Nikon semiconductor lithography lineup](#)

Nikon | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 ArF immersion dry scanners and package lithography categories

用于 设备公司地图：按工艺能力定位

[F081] [Canon semiconductor manufacturing equipment](#)

Canon | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 i line DUV and nanoimprint system portfolio only

用于 设备公司地图：按工艺能力定位

[F082] [Active photonics alignment systems](#)

Physik Instrumente | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Six DOF multichannel active optical alignment mechanisms controllers algorithms

用于 光子封装测试与罗博特科案例；技术壁垒与投资验证

[F083] [Generation of EUV radiation](#)

TRUMPF | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 CO2 laser drive in EUV system ASML integration ZEISS optics

用于 设备公司地图：按工艺能力定位

[F084] [华海清科 2026 年半年度报告](#)

华海清科/巨潮资讯 | 发布 2026-08-22 | 检索 2026-10-15 | official | search-excerpt

证据范围 CMP thinning equipment portfolio; no market forecasts adopted

用于 设备公司地图：按工艺能力定位

[F085] [中科飞测产品目录](#)

中科飞测 | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Wafer defects CD film thickness overlay and 3D topography metrology

用于 设备公司地图：按工艺能力定位

[F086] [长川科技产品目录](#)

长川科技 | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Tester handler probe AOI categories; no AI customer certification inferred

用于 设备公司地图：按工艺能力定位

[N001] [Executive summary: Key Questions on Energy and AI](#)

IEA | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 2025 年 485TWh 与 2030 年 950TWh 预测及电力和供应链约束

用于 从电网到芯片：供电架构与材料；热量怎样离开机柜：液冷、建设与运维

[N002] [Why Scaling AI Compute Performance Requires a New Power Architecture](#)

NVIDIA | 发布 2026-08-11 | 检索 2026-10-15 | official | read

证据范围 800VDC 混合 power rack 2026H2 与 row power center 2027 时间边界

用于 从电网到芯片：供电架构与材料

[N003] [NVIDIA 800 VDC Architecture Will Power the Next Generation of AI Factories](#)

NVIDIA | 发布 2025-05-20 | 检索 2026-10-15 | official | search-excerpt

证据范围 800VDC 架构背景与路线图，效益为厂商目标

用于 延伸阅读与研究线索

[N004] [Open Rack V3 Power Shelf Rev 1.0](#)

OCP | 发布 未确认 | 检索 2026-10-15 | standard | read

证据范围 48V 公共母排、整流模块 N+1 热插拔、接口和材料要求

用于 从电网到芯片：供电架构与材料

[N005] [Infineon to revolutionize power delivery architecture for future AI server racks with NVIDIA](#)

Infineon | 发布 2025-05-20 | 检索 2026-10-15 | official | read

证据范围 Si SiC GaN 在 grid-to-core 供电的不同位置

用于 从电网到芯片：供电架构与材料；热量怎样离开机柜：液冷、建设与运维

[N006] [Q3 FY25 Investor Presentation](#)

Infineon | 发布 2025-08-05 | 检索 2026-10-15 | official | search-excerpt

证据范围 第 59 页 800V 54V 12V 及 VRM、储能示意

用于 从电网到芯片：供电架构与材料

[N007] [AI Data Center Energy Performance Framework: Introduction and Purpose](#)

ASHRAE | 发布 未确认 | 检索 2026-10-15 | standard | read

证据范围 液冷 W 温度分类、PUE/WUE 与全生命周期阶段

用于 热量怎样离开机柜：液冷、建设与运维

[N008] [XDU1350 Coolant Distribution Unit Application and Planning Guide](#)

Vertiv | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 CDU 液体监测、泵过滤、通讯、双侧环路配置

用于 热量怎样离开机柜：液冷、建设与运维

[N009] [CDUs for High-Density Liquid Cooling](#)

CoolIT | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 液到液 CDU 换热器隔离技术和设施水环路

用于 从电网到芯片：供电架构与材料；热量怎样离开机柜：液冷、建设与运维

[N010] [ACS Liquid Cooling Cold Plate Requirements Document](#)

OCP | 发布 2019-10-09 | 检索 2026-10-15 | standard | read

证据范围 冷板流道、快接、接液材料、质量与可靠性要求

用于 从电网到芯片：供电架构与材料；热量怎样离开机柜：液冷、建设与运维；网络与光互联企业地图、瓶颈与可证伪投资命题

[N011] [Thermal Interface Materials](#)

Boyd | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 TIM 降低热源与散热结构接触空隙、材料选型

用于 从电网到芯片：供电架构与材料；热量怎样离开机柜：液冷、建设与运维

[N012] [Eaton expands operations in Nebraska with new manufacturing facility](#)

Eaton | 发布 2026-04-08 | 检索 2026-10-15 | official | read

证据范围 中压开关柜功能与新工厂 2027H1 预计投产

用于 从电网到芯片：供电架构与材料；热量怎样离开机柜：液冷、建设与运维；网络与光互联企业地图、瓶颈与可证伪投资命题

[N013] [Data centers industry](#)

Siemens | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 设施电气、冷却、消防、安全与控制系统范围

用于 从电网到芯片：供电架构与材料；热量怎样离开机柜：液冷、建设与运维

[N014] [iBusway for Data Center](#)

Schneider Electric | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 插接母线向 IT 机架配电与铜导体

用于 从电网到芯片：供电架构与材料；热量怎样离开机柜：液冷、建设与运维

[N015] [Rack PDU: Power distribution units for server racks](#)

Eaton | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 机架 PDU 插座/输入/遥测产品角色

用于 从电网到芯片：供电架构与材料

[N016] [Reliable Generators for Data Centers](#)

Cummins | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 备用和连续主用发电产品及项目

用于 从电网到芯片：供电架构与材料

[N017] [Coolinside Full Chain Liquid Cooling Solution](#)

Envicool | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 冷板歧管快接 CDU 及液到空气方案

用于 热量怎样离开机柜：液冷、建设与运维

[N018] [Kehua Presents Coordinated Computing, Power and Cooling Approach at EPIC 2026](#)

Kehua | 发布 2026-08-17 | 检索 2026-10-15 | official | read

证据范围 Power Pod 整合变压器 UPS 配电与厂商披露项目

用于 从电网到芯片：供电架构与材料；热量怎样离开机柜：液冷、建设与运维

[N019] [FusionPower6000-G 3.0 Huawei PowerPOD User Manual](#)

Huawei | 发布 2026-07-25 | 检索 2026-10-15 | official | search-excerpt

证据范围 低压输入柜 UPS 旁路馈线柜与机房电气边界

用于 从电网到芯片：供电架构与材料；热量怎样离开机柜：液冷、建设与运维

[N020] [Minutes of 2025 AGM](#)

Delta | 发布 2025-05-29 | 检索 2026-10-15 | official | search-excerpt

证据范围 液冷模块 CDU 与磁性材料业务，厂商披露生产阶段

用于 从电网到芯片：供电架构与材料；热量怎样离开机柜：液冷、建设与运维

[N021] [DGX GB Rack Scale Systems User Guide: Networking](#)

NVIDIA | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 NVLink 机架内 scale-up 与 IB/Ethernet 机架间 scale-out

用于 全景图与系统连接；数据怎样移动：机内、机架、集群与跨中心网络；网络与光互联企业地图、瓶颈与可证伪投资命题

[N022] [AI Network Fabric Deployment Guide](#)

Arista | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 RoCEv2 leaf-spine 拓扑、配置和验证方法

用于 全景图与系统连接；数据怎样移动：机内、机架、集群与跨中心网络；把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N023] [Specification History](#)

Ultra Ethernet Consortium | 发布 未确认 | 检索 2026-10-15 | standard | read

证据范围 UEC1.0 于 2025-06-11 与 1.1 于 2026-10-12 发布

用于 数据怎样移动：机内、机架、集群与跨中心网络；把一条光链路拆到材料、晶圆、装配与测试

[N024] [About UALink](#)

UALink Consortium | 发布 未确认 | 检索 2026-10-15 | standard | read

证据范围 开放 scale-up 协议 1.0 于 2025-04 公开

用于 全景图与系统连接；数据怎样移动：机内、机架、集群与跨中心网络

[N025] [PCIe 7.0 Specification Version 1.0 Now Available to Members](#)

PCI-SIG | 发布 2025-06-11 | 检索 2026-10-15 | standard | read

证据范围 PCIe7.0 128GT/s 规格发布，不等于产品普及

用于 全景图与系统连接；数据怎样移动：机内、机架、集群与跨中心网络

[N026] [CXL Consortium Releases the CXL 4.0 Specification](#)

CXL Consortium | 发布 2025-11-18 | 检索 2026-10-15 | standard | read

证据范围 CXL4.0 于 2025-11-18 发布与内存互连技术边界

用于 全景图与系统连接；数据怎样移动：机内、机架、集群与跨中心网络

[N027] [BlueField-3 Networking Platform Introduction](#)

NVIDIA | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 DPU 卸载隔离网络存储安全与管理

用于 数据怎样移动：机内、机架、集群与跨中心网络；网络与光互联企业地图、瓶颈与可证伪投资命题；计算架构与芯片设计

[N028] [Aries PCIe/CXL Smart DSP Retimers](#)

Astera Labs | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 PCIe6.x/CXL3.x 重定时器延伸电连接

用于 数据怎样移动：机内、机架、集群与跨中心网络；网络与光互联企业地图、瓶颈与可证伪投资命题

[N029] [HiWire 345Tb DDC Solution Brief](#)

Credo | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 AEC、Retimer、Optical DSP、SerDes/IP 产品角色

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N030] [FiberEdge and DirectEdge Signal Integrity Technology](#)

Semtech | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 TIA 和激光驱动、LPO 模拟 PMD 产品

用于 数据怎样移动：机内、机架、集群与跨中心网络；把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N031] [CopperEdge Signal Integrity Technology](#)

Semtech | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 ACC 模拟均衡与铜缆距离依具体链路

用于 数据怎样移动：机内、机架、集群与跨中心网络；网络与光互联企业地图、瓶颈与可证伪投资命题

[N032] [Broadcom Delivers First 51.2Tbps CPO Ethernet Switch Platform](#)

Broadcom | 发布 2024-03-26 | 检索 2026-10-15 | official | search-excerpt

证据范围 Bailly CPO 制造封装与光纤装配厂商宣称

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N033] [EMLs](#)

Lumentum | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 InP EML 的 DFB 与 EAM 集成，内部晶圆厂

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N034] UHP Laser Sources for CPO

Lumentum | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 InP CW 激光与 ELSFP 外置可维修光源

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N035] Lumentum Showcases Next-Generation InP Chip Solutions at OFC2025

Lumentum | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 200G/lane DD EML 技术论文与 NVIDIA 生态伙伴披露

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[N036] Silicon photonics technology

GlobalFoundries | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 200/300mm 硅光制造 PDK、电子光子封装与光纤连接

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N037] In-Package Optical IO Unleashing Innovation

Ayar Labs | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 TeraPHY 光 IO chiplet 与 SuperNova 多波长光源

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题；计算架构与芯片设计

[N038] SILITH Showcases 400G/lane Silicon Photonics Platform at OFC2026

SILITH | 发布 2026-03-18 | 检索 2026-10-15 | official | search-excerpt

证据范围 400G/lane 收发构件展示；证据范围限于公开摘要

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[N039] Datacom Innovation in the AI Era

Coherent | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 GaAs VCSEL 与 InP 激光晶圆平台

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N040] Significant Advancement in VCSEL Performance for Next-Gen AI Networks

Coherent | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 氧化孔径 VCSEL 与 200G/lane 发展方向

用于 把一条光链路拆到材料、晶圆、装配与测试

[N041] MOCVD Technology

AIXTRON | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 金属有机化学气相沉积生长单晶薄膜

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N042] Photonics assembly and testing From Lab to Fab

ficonTEC | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 多轴亚微米对准固化焊接及测试良率周期要求

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N043] PINovAlign Automated Photonics Alignment System

PI | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 全自动 6DOF 光学对准，证明多轴不是某一厂商独有

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N044] NX5402A Silicon Photonics Wafer Test System

Keysight | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 晶圆级光电联合自动测试与校准

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N045] OEM Core Products Guide

Corning | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 高精度氧化锆插芯与光纤连接器

用于 热量怎样离开机柜：液冷、建设与运维；数据怎样移动：机内、机架、集群与跨中心网络；把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N046] DCI and scale-across networks

Ciena | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 DCI 相干模块、光线路系统、metro backbone 与 submarine

用于 数据怎样移动：机内、机架、集群与跨中心网络；把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N047] ELSFP Implementation Agreement 2.0

OIF | 发布 未确认 | 检索 2026-10-15 | standard | search-excerpt

证据范围 可插拔外置激光源接口标准

用于 数据怎样移动：机内、机架、集群与跨中心网络；把一条光链路拆到材料、晶圆、装配与测试

[N048] AI and Data Center Networking

InnoLight | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 100G 至 1.6T 光模块产品组合

用于 数据怎样移动：机内、机架、集群与跨中心网络；把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N049] ECOC2025 Demo Eoptolink

OIF/Eoptolink | 发布 2025-09-29 | 检索 2026-10-15 | official | search-excerpt

证据范围 800ZR 相干模块参加 OIF 展示；非出货量证明

用于 数据怎样移动：机内、机架、集群与跨中心网络；把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N050] Infocommunications Research and Development

Sumitomo Electric | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 多芯光纤、连接器与 DFB 激光研发

用于 热量怎样离开机柜：液冷、建设与运维；把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N051] Briefing Material of Furukawa Electric

Furukawa Electric | 发布 2025-12-04 | 检索 2026-10-15 | official | search-excerpt

证据范围 光纤带缆 MT 插芯 DFB 与冷却产品

用于 热量怎样离开机柜：液冷、建设与运维；把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N052] Direct to chip cooling implementation

Reddit r/datacenter | 发布 2025-05-26 | 检索 2026-10-15 | forum | read

证据范围 管道空间残余风冷二次液体等匿名经验，非市场事实

用于 热量怎样离开机柜：液冷、建设与运维；网络与光互联企业地图、瓶颈与可证伪投资命题

[N053] Why does Silicon Photonic become more dominant in 1.6T

Reddit r/Semiconductors | 发布 未确认 | 检索 2026-10-15 | forum | read

证据范围 讨论 EML 与硅光份额转移的条件，非管理层原始口径

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[N054] CHx2000 The AI CDU

CoolIT | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 两 MW 级 CDU 产品含泵过滤传感控制的厂商说明

用于 热量怎样离开机柜：液冷、建设与运维

[N055] Cold Plates for AI Server Cooling

Boyd | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 冷板与环路歧管 CDU 协同设计

用于 热量怎样离开机柜：液冷、建设与运维

[N056] SiFotonics Homepage

SiFotonics | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 800G/1.6T 硅光产品发布，未核实客户量产

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[N057] [Digital 48V Datacenter Solution Q4 2024](#)

Monolithic Power Systems | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 IBC POL eFuse 与 Z 轴/垂直供电布局

用于 从电网到芯片：供电架构与材料

[N058] [AI Data Center Electronics](#)

TDK | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 磁性无源与保护元件跨 UPS PSU SST 供电链

用于 从电网到芯片：供电架构与材料；热量怎样离开机柜：液冷、建设与运维

[N060] [Energy data centre solutions](#)

Prysmian | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 中低压电缆建筑线消防控制及数据线缆

用于 从电网到芯片：供电架构与材料；热量怎样离开机柜：液冷、建设与运维；把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N061] [H3C Switches Product Resource Center](#)

H3C | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 800G 以太网交换机产品与 CPO 机型宣称

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[N062] [800G AI ML Fabrics](#)

Accton | 发布 2025-08-26 | 检索 2026-10-15 | official | search-excerpt

证据范围 台湾 ODM 800G AI 交换硬件

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[N063] [IEEE P802.3df Task Force](#)

IEEE | 发布 2024-02-16 | 检索 2026-10-15 | standard | search-excerpt

证据范围 800GbE 标准获批，接口与协议范围

用于 数据怎样移动：机内、机架、集群与跨中心网络

[N064] [NVIDIA Announces Spectrum-X Photonics CPO Networking Switches](#)

NVIDIA | 发布 2025-03-18 | 检索 2026-10-15 | official | search-excerpt

证据范围 CPO 网络平台发布与当时预期时间，不是 2026 验收证明

用于 把一条光链路拆到材料、晶圆、装配与测试

[N065] [Marvell Optical DSPs](#)

Marvell | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 PAM4 与 coherent DSP 不同产品及 Ara/Aquila/COLORZ 范围

用于 数据怎样移动：机内、机架、集群与跨中心网络；把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题；计算架构与芯片设计

[N066] [Silicon Photonics SiPho](#)

Tower Semiconductor | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Si 和 SiN 波导、Ge 探测器、Al/Cu 后端及混合键合平台

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N067] [Silicon nitride-based photonics](#)

imec | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 LPCVD 低损耗 SiN 光子工艺平台

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N068] [Photonics SOI Products](#)

Soitec | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 SmartCut Photonics-SOI 工程衬底产品

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N069] [Magneto Optic Faraday Rotator Garnet Crystals](#)

Coherent | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 铋掺杂稀土铁石榴石磁光材料与隔离器用途

用于 把一条光链路拆到材料、晶圆、装配与测试；网络与光互联企业地图、瓶颈与可证伪投资命题

[N070] [300x300 Optical Circuit Switch](#)

Lumentum | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 MEMS 光路交换 R300 产品角色，非客户规模验收证据

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[N071] [Hyosung Heavy Industries Showcases Future of Power Grids for AI Data Centers](#)

Hyosung | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 韩国变压器开关设备 SST STATCOM 等 AI 电网展示

用于 热量怎样离开机柜：液冷、建设与运维；网络与光互联企业地图、瓶颈与可证伪投资命题

[N072] [Remote Power Panel](#)

LS ELECTRIC | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 数据中心 RPP 配电产品，不能证明整体收入

用于 热量怎样离开机柜：液冷、建设与运维；网络与光互联企业地图、瓶颈与可证伪投资命题

[N073] [Cisco Advances Scalability and Openness at OCP 2026](#)

Cisco | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 G300 系统开始向客户出货的厂商披露

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[N074] [HPE 2025 Annual Report](#)

HPE | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Juniper 收购后网络与 AI 数据中心产品范围

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[N075] [Xinghe AI Fabric 2.0](#)

Huawei | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 CloudEngine 与 StarryLink 800GE 产品地图

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[N076] [Accelink 2025 Annual Report](#)

Accelink | 发布 2026-04-27 | 检索 2026-10-15 | official | search-excerpt

证据范围 800G LPO 与 1.6T 光模块系列的公司披露

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[N078] [Silicon Photonics Test MeasureOne Solutions](#)

FormFactor | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 FormFactor Keysight PI 三方光电晶圆/裸片测试体系

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[N079] [Ayar Labs Products](#)

Ayar Labs | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 TeraPHY 光引擎与 SuperNova 外部光源完整组件体系

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[N080] [Intel Demonstrates First Fully Integrated Optical IO Chiplet](#)

Intel | 发布 2024-06-26 | 检索 2026-10-15 | official | search-excerpt

证据范围 PIC 带激光/放大器与 EIC 结合的光 IO 演示

用于 网络与光互联企业地图、瓶颈与可证伪投资命题；计算架构与芯片设计

[N081] [NVLink The Scale Up Network for AI Factories](#)

NVIDIA | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 NVLink scale-up 与 Quantum/Spectrum scale-out 层级

用于 数据怎样移动：机内、机架、集群与跨中心网络

[N082] [Three Genius Ideas for AI Fabrics](#)

Arista | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 多平面网络与端点组织，供应商工程方法

用于 数据怎样移动：机内、机架、集群与跨中心网络

[N083] [Powering AI Centers with AI Spines](#)

Arista | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 leaf-spine 与 AI spine 设计和规模方法

用于 数据怎样移动：机内、机架、集群与跨中心网络

[N084] [UEC Launches Specification 1.0](#)

Ultra Ethernet Consortium | 发布 2025-06-11 | 检索 2026-10-15 | standard | read

证据范围 端到端 RDMA 以太网协议及拥塞与互操作范围

用于 数据怎样移动：机内、机架、集群与跨中心网络

[N085] [Free Space Optical Isolators](#)

Thorlabs | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 光隔离器抑制返回光反馈，与电绝缘不同

用于 把一条光链路拆到材料、晶圆、装配与测试

[N086] [Breakthroughs in Heterogeneous Photonic Integration](#)

imec | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 薄膜铌酸锂调制器转印到硅光平台属于研究演示

用于 把一条光链路拆到材料、晶圆、装配与测试

[N087] [InP Optoelectronics](#)

Coherent | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 InP 光探测器与激光产品组合

用于 把一条光链路拆到材料、晶圆、装配与测试

[N088] [ROADMs and Wavelength Management](#)

Lumentum | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 WSS OCM ROADM 模块与光路重配置功能

用于 把一条光链路拆到材料、晶圆、装配与测试

[N089] [Credo We Connect](#)

Credo | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 AEC DSP SerDes 及链路可靠性定位，厂商陈述

用于 数据怎样移动：机内、机架、集群与跨中心网络；网络与光互联企业地图、瓶颈与可证伪投资命题

[N090] [Semtech Data Center Solutions](#)

Semtech | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 retimed 和 LPO 的激光驱动与 TIA 400G 至 1.6T 组合

用于 数据怎样移动：机内、机架、集群与跨中心网络；把一条光链路拆到材料、晶圆、装配与测试；计算架构与芯片设计

[N092] [DGX GB200 Service Manual Safety](#)

NVIDIA | 发布 2025-09-18 | 检索 2026-10-15 | official | read

证据范围 机架机械承载接地布线与设备维护边界

用于 热量怎样离开机柜：液冷、建设与运维

[R001] [CASPA CEO Summit Agenda](#)

CASPA | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 活动日期、议程、时间与参与机构；预告可能调整

用于 CASPA 会议与公司对应

[R002] [CASPA Speakers Moderators and MCs](#)

CASPA | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 当前讲者、职务与所属机构，包括 Solidigm 及 Upscale AI 代表

用于 AI 基础设施产业全景与 CASPA 参会研究；CASPA 会议与公司对应

[R003] [CASPA Sponsors](#)

CASPA | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Event/Table/Display Sponsors 及联盟名称，仅证明网页出现

用于 CASPA 会议与公司对应

[R004] [NVIDIA Vera Rubin Opens Agentic AI Frontier](#)

NVIDIA | 发布 2026-03-16 | 检索 2026-10-15 | official | read

证据范围 多芯片多机架平台、NVL72 构成；宣称 full production 与伙伴下半年产品提供分开

用于 AI 基础设施产业全景与 CASPA 参会研究；全景图与系统连接；计算架构与芯片设计；计算与设计公司地图

[R005] [AMD Instinct MI400 Series GPUs](#)

AMD | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 MI455X HBM4 规格、Helios72GPU 参考设计、工程测算与提供时间口径

用于 AI 基础设施产业全景与 CASPA 参会研究；计算架构与芯片设计；计算与设计公司地图

[R006] [TPU7x Ironwood specifications](#)

Google Cloud | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 192GiB HBM、7380GB/s、ICI 与 DCN 区分及 VMEM/SRAM 层级

用于 全景图与系统连接；计算架构与芯片设计；计算与设计公司地图；社区异议与深入阅读

[R007] [Cerebras announces third generation wafer scale engine](#)

Cerebras | 发布 2024-03-13 | 检索 2026-10-15 | official | read

证据范围 WSE-3 厂商规格 900000 核 44GB SRAM；网页页眉与正文日期有差异以正文发布日期

用于 计算架构与芯片设计；计算与设计公司地图

[R008] [SambaNova SN40L Scaling the AI Memory Wall with Dataflow and Composition of Experts](#)

SambaNova authors arXiv | 发布 2024-05-13 | 检索 2026-10-15 | paper | read

证据范围 SN40L RDU 与 dataflow 和分层内存的论文说明，不代表当前所有型号

用于 计算与设计公司地图；社区异议与深入阅读

[R009] [d-Matrix Corsair AI Inference Platform Enters Full Production](#)

d-Matrix | 发布 2026-06-09 | 检索 2026-10-15 | official | read

证据范围 Corsair 量产自述、N6/Alchip、有机基板 SRAM DIMC 与 LPDDR5，以及 qualified customers 边界

用于 计算与设计公司地图；社区异议与深入阅读

[R010] [Etched Frontier Inference Clusters current product page](#)

Etched | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 当前官网 A0 回片/客户验证、LVI 和 HBM/SRAM CSM 路线；不等同独立量产验证

用于 计算与设计公司地图；社区异议与深入阅读

[R011] [Roofline An Insightful Visual Performance Model](#)

Williams Waterman Patterson UC Berkeley | 发布 2009-03-01 | 检索 2026-10-15 | paper | search-excerpt

证据范围 Roofline 框架中的运算强度、带宽与峰值计算关系；证据范围限于摘要

用于 计算架构与芯片设计

[R012] [MLPerf Inference Datacenter benchmark rules and metrics](#)

MLCommons | 发布 未确认 | 检索 2026-10-15 | standard | read

证据范围 场景/质量/可用性类别及墙插实测功率的解释

用于 全景图与系统连接；计算与设计公司地图

[R013] [Efficient Memory Management for Large Language Model Serving with PagedAttention](#)

Kwon et al arXiv SOSP 2023 | 发布 2023-09-12 | 检索 2026-10-15 | paper | read

证据范围 KV cache 分页与共享、调度共同影响服务内存效率

用于 全景图与系统连接；计算架构与芯片设计

[R014] [Inside vLLM Anatomy of a High Throughput LLM Inference System](#)

vLLM | 发布 2025-09-05 | 检索 2026-10-15 | official | search-excerpt

证据范围 prefill/decode、连续批处理、缓存及分布式服务执行路径

用于 全景图与系统连接；计算架构与芯片设计

[R015] [Amazon EC2 Trn3 UltraServers](#)

AWS | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Trainium3 与 Trn3 的云系统定位及软件/系统入口

用于 计算架构与芯片设计；计算与设计公司地图

[R016] Altera FPGA products

Altera | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Agilix 及 FPGA、接口、内存控制器与软件产品分类

用于 计算架构与芯片设计；计算与设计公司地图

[R017] DSO.ai AI Driven Design Applications

Synopsys | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 强化学习设计空间 PPA 优化及多 die 热信号协同

用于 计算架构与芯片设计；计算与设计公司地图

[R018] Cadence Cerebrus Intelligent Chip Explorer

Cadence | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 设计流程优化、强化学习与 PPA 目标，收益为厂商陈述

用于 计算架构与芯片设计；计算与设计公司地图

[R019] Calibre PERC and 3DPERC

Siemens EDA | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 电路与版图可靠性验证类别；证据范围限于摘要

用于 计算架构与芯片设计；计算与设计公司地图

[R020] SemiAnalysis Models and Research

SemiAnalysis | 发布 未确认 | 检索 2026-10-15 | analysis | read

证据范围 公开说明 TCO 网络数据中心模型结构与连接比例方法；未访问付费模型

用于 AI 基础设施产业全景与 CASPA 参会研究；全景图与系统连接；社区异议与深入阅读

[R021] AI Capex Surge Bottlenecks And The Race For ROI

Seeking Alpha podcast Rena Sherbill Amrita Roy Uttam Dey | 发布 未确认 | 检索 2026-10-15 | analysis | read

证据范围 公开访谈对毛利率、库存、订单及 ROI 的分析问题；不采纳目标价

用于 AI 基础设施产业全景与 CASPA 参会研究；技术壁垒与投资验证；社区异议与深入阅读

[R022] [Rebellions current products and infrastructure](#)

Rebellions | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Rebel100/Server/Rack/POD 及 SDK 入口，生产部署为厂商自述

用于 计算与设计公司地图

[R023] [SiMa.ai MLSoC family](#)

SiMa.ai | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Modalix edge/physical AI 与 SoM/PCIe/devkit 形态

用于 计算与设计公司地图；社区异议与深入阅读

[R024] [Normal ASICs Physics Based Architecture](#)

Normal Computing | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 热力学/随机计算定位 CN101 与后续路线；页内 benchmarking 和 roadmap 边界

用于 计算与设计公司地图

[R025] [Unconventional AI official homepage](#)

Unconventional AI | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 硬件算法共同演化定位；公开信息不足以作量产规格比较

用于 计算与设计公司地图

[R026] [Ricursive Intelligence Recursive Self Improvement](#)

Ricursive Intelligence | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 AI for chip design 与 chip design for AI 定位

用于 计算与设计公司地图

[R027] [Cognichip About ACI](#)

Cognichip | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 ACI 与 physics-informed foundation model 定位及 IP 保护自述

用于 计算与设计公司地图

[R028] [Arm Neoverse CSS N4](#)

Arm | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 CPU 子系统、Neoverse 平台及 AGI CPU 产品入口

用于 计算与设计公司地图

[R029] [ValueInvesting comment Educational Play 465 on sustaining versus growth CapEx](#)

Reddit anonymous user | 发布 2026-08-10 | 检索 2026-10-15 | forum | read

证据范围 匿名评论关于维持与增长资本支出分拆的建议，不证明财务数字或 GPU 寿命

用于 社区异议与深入阅读

[R030] [chipdesign saw235 comment on waveform debug and AI verification](#)

Reddit anonymous user | 发布 未确认 | 检索 2026-10-15 | forum | read

证据范围 匿名自述波形 debug 与报告生成体验，仅转成验证问题；评论精确日期不明

用于 社区异议与深入阅读

[R031] [LocalLLaMA Meet Sohu discussion](#)

Reddit community | 发布 2024-06-25 | 检索 2026-10-15 | forum | read

证据范围 专用 ASIC 的模型假设与测试条件争议；不使用旧宣传倍数当当前实绩

用于 社区异议与深入阅读

[R032] [NVIDIA DGX Vera Rubin NVL72 datasheet](#)

NVIDIA | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 不同系统资料的规格和口径需依文档版本逐项核查

用于 计算架构与芯片设计

[R033] [Intel Xeon 6 ARK products](#)

Intel | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Xeon CPU 产品类别，不声称单一最新型号或客户关系

用于 计算与设计公司地图

[R034] 昇腾产品形态说明

华为昇腾社区 | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 昇腾训练推理芯片与 Atlas 产品类别；未验证所有型号可购性

用于 计算与设计公司地图

[R035] 寒武纪产品技术

寒武纪 | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 思元加速卡与 NeuWare 等公开产品类别，首页有旧内容不可当最新旗舰目录

用于 计算与设计公司地图

[R036] MTT S5000 Linux 服务器驱动产品介绍

摩尔线程 | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 MTT S5000 与 S4000 产品类别与文档范围；证据限于公开摘要

用于 计算与设计公司地图

[R037] 昆仑芯产品信息

昆仑芯 | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 AI 芯片与加速卡类别；官网首页可能不含全部新产品

用于 计算与设计公司地图

[R038] 壁仞科技产品与软件平台

壁仞科技 | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 壁砺 OAM 模组/加速卡及 BIRENSUPA 平台类别，部署为自述

用于 计算与设计公司地图

[R039] Fortinet Secure Processors FortiASIC

Fortinet | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 安全与网络处理器定位，不等同训练后端 fabric

用于 CASPA 会议与公司对应

[R040] [Rambus Interface IP](#)

Rambus | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 接口 IP 类别及内存/连接生态位置

用于 CASPA 会议与公司对应

[R041] [Silvaco official site](#)

Silvaco | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 设计工具生态入口；不支持具体产品性能结论

用于 CASPA 会议与公司对应

[R042] [Asianometry and Dylan Patel How the semiconductor industry actually works](#)

Dwarkesh Podcast | 发布 2024-10-02 | 检索 2026-10-15 | analysis | read

证据范围 播客主题、原始时间索引与相关 transcript 段落，历史数据不当 2026 事实

用于 社区异议与深入阅读

[R043] [Branch Education video directory](#)

Branch Education | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 官方视频目录与延伸阅读；不作为视频内容的逐条技术证据

用于 社区异议与深入阅读

[R044] [How does NAND Flash Work Reading from TLC](#)

Branch Education YouTube | 发布 未确认 | 检索 2026-10-15 | analysis | search-excerpt

证据范围 视频标题与描述为 NAND 系列学习入口，未做完整字幕审计

用于 社区异议与深入阅读

[R045] [Introducing 7th Generation TPUs Ironwood](#)

Google Cloud YouTube | 发布 2025-04-09 | 检索 2026-10-15 | official | search-excerpt

证据范围 官方视频存在与主题；当前规格以 R006 为准，未完整观看

用于 社区异议与深入阅读

[R046] YOFC Data Center cabling products

长飞 YOFC | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 预端接干线光缆、高密度配线架、MTP-LC 等布线产品，非光模块份额证明
用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[R047] YOFC Trunk Cable

长飞 YOFC | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 预端接干线光缆产品类别；旧速率说明不作为最新市场份额
用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[R048] 亨通产业产品与解决方案

亨通集团 | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 棒纤缆光网络与数据中心业务位置，不采纳全球排名宣传
用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[R049] 天孚通信产品与服务

天孚通信 | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Mux TOSA Demux POSA Lensed FAU AWG 等光器件及服务，不证明全部 SKU 同规模量产
用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[R050] 腾景科技无源器件目录

腾景科技 Optowide | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 准直器、环形器、VOA、隔离器、WDM 等无源器件类别
用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[R051] 源杰科技数据中心光芯片产品

源杰科技 | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 PAM4 EML 与 CW DFB 光源产品目录，仅支持型号能力不证明全部批量阶段
用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[R052] [CWDM LAN4 AWG Chip Subassemblies](#)

仕佳光子 | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 AWG 波分复用/解复用芯片与子组件产品

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[R053] [盛科通信 2025 环境社会与治理报告](#)

盛科通信 CNINFO | 发布 2026-04-23 | 检索 2026-10-15 | official | read

证据范围 p3 12.8T/25.6T 和最大 800G 等交换芯片产品披露，不证明与更高容量平台全面等效

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[R054] [中科驭数 K2 Pro DPU](#)

中科驭数 | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 KPU 架构网络存储安全卸载、2x100GbE PCIe Gen3x16 NVMe-oF 产品能力

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[R055] [云豹云霄 DPU 产品发布](#)

云豹智能 | 发布 2021-12-15 | 检索 2026-10-15 | official | read

证据范围 历史 DPU 2x25G Virtio RoCE NVMeoF 等类别；不当最新 400/800G 产品

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[R056] [云豹智能官网](#)

云豹智能 | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 仅支持 DPU SoC 企业定位，不支持 400G 具体型号的完整规格

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[R057] [Planar Lightwave Circuit AWG Platform](#)

仕佳光子 | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 PLC/AWG 平台定位，与光模块及激光器分层

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[R058] [海光 CPU 产品](#)

海光信息 | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 数据中心 CPU 类别；历史核心数不代表当前全部型号上限，证据限于摘要

用于 计算与设计公司地图

[R059] [华为鲲鹏 920](#)

海思 | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Arm 通用 CPU 与 DDR4 PCIe4 多路互连的该型号资料，不代表所有最新鲲鹏型号

用于 计算与设计公司地图

[R060] [MAP1602 C SSD controller](#)

联芸科技 MAXIO | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 客户端 DRAM-less PCIe4x4 NVMe2.0 主控及 ECC/NAND 接口，不据此称企业级盘认证

用于 存储公司版图与可验证的投资问题

[R061] [Empyrean ALPS SPICE simulator](#)

华大九天 | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 模拟混合信号及存储电路并行 SPICE 仿真工具，性能倍数不当独立实测

用于 计算与设计公司地图

[R062] [NanoSpice Pro FastSPICE product brief](#)

概伦电子 | 发布 2024-06 | 检索 2026-10-15 | official | read

证据范围 FastSPICE 及 DRAM SRAM Flash MRAM 等存储电路验证定位

用于 计算与设计公司地图

[R063] [SmtCell parameterized layout tool](#)

广立微 | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 参数化测试结构版图工具，不代表完整芯片布局布线流程；证据限于摘要

用于 计算与设计公司地图

[R064] 源杰科技 2026 年半年度报告

源杰科技 CNINFO | 发布 2026-08-29 | 检索 2026-10-15 | official | read

证据范围 p14 CW70/100mW 和 100G EML 批量、200G EML 客户验证；p17-18 异质外延对接量子阱寄生电容钝化等公司陈述

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[R065] 天孚通信 2026 年半年度报告

天孚通信 深交所 | 发布 2026-08-19 | 检索 2026-10-15 | official | read

证据范围 p12-13 无源有源光模块代工与 FAU 光引擎封装高密度互连 ELS 方案类别

用于 网络与光互联企业地图、瓶颈与可证伪投资命题

[S001] ECC and Spare Blocks help to keep Kingston SSD data protected from errors

Kingston | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 SSD controller, SRAM/DRAM mapping tables, data-path protection and spare blocks

用于 从一块企业级 SSD 看懂整条产业链；SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S002] 3D NAND Anatomy and Scaling

Applied Materials | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Charge-trap cell materials, staircase, channel and metal fill; source contains silicon nitrate typo interpreted using SiN context

用于 从一块企业级 SSD 看懂整条产业链；NAND 的工艺、材料和设备怎样衔接

[S003] High-density 3D flash memory using high-precision wafer bonding brings new value to storage

Kioxia / licensed EE Times interview | 发布 2024-07-30 | 检索 2026-10-15 | official | read

证据范围 CBA separately processes CMOS and array; thermal budget, alignment, flatness and bit-density considerations

用于 从一块企业级 SSD 看懂整条产业链；NAND 的工艺、材料和设备怎样衔接；存储公司版图与可验证的投资问题

[S004] [Memory Channel Hole Etch Technology for 3D NAND over 400 Layers](#)

Tokyo Electron | 发布 2023-06-09 | 检索 2026-10-15 | official | read

证据范围 Technology demonstration of 10 micrometer deep channel etch in 33 minutes; not full-fab throughput

用于 从一块企业级 SSD 看懂整条产业链；NAND 的工艺、材料和设备怎样衔接；存储公司版图与可验证的投资问题

[S005] [Cryogenic Etching](#)

Lam Research | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Cryogenic channel etch, oxide/nitride deposition, W/Mo metal integration and wafer stress management

用于 NAND 的工艺、材料和设备怎样衔接；存储公司版图与可验证的投资问题

[S006] [Applied Materials Unveils Deposition and Selective Etch Systems to Advance 3D Chip Scaling](#)

Applied Materials | 发布 2026-06-15 | 检索 2026-10-15 | official | read

证据范围 SiN ALD and selective Mo etch; adopted wordline metallization and isolation applications

用于 从一块企业级 SSD 看懂整条产业链；NAND 的工艺、材料和设备怎样衔接；存储公司版图与可验证的投资问题

[S007] [About Xtacking](#)

YMTC | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Separate array and peripheral wafers, vertical metal connections, Xtacking 4.0 product family

用于 从一块企业级 SSD 看懂整条产业链；NAND 的工艺、材料和设备怎样衔接；存储公司版图与可验证的投资问题

[S008] [96 Layers and Beyond: 3D NAND Material and Integration Challenges](#)

Entegris | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 HAR film uniformity, nitride removal, contamination and process integration constraints

用于 从一块企业级 SSD 看懂整条产业链；NAND 的工艺、材料和设备怎样衔接；存储公司版图与可验证的投资问题

[S009] 2022 Scientific Report: Applications of Molecular Dynamics on Formulation Design for Complex HAR 3D Nanostructures

Entegris | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 PDF pp.17-19: hot phosphoric sacrificial SiN removal, SiO₂ redeposition, bath-life/defect tradeoffs

用于 从一块企业级 SSD 看懂整条产业链；NAND 的工艺、材料和设备怎样衔接；存储公司版图与可验证的投资问题

[S010] Advancing Semiconductors: Thin Film Solutions Offering

Merck | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Organosilane, organometallic, dielectric and precursor delivery/application development portfolio; not customer BOM

用于 NAND 的工艺、材料和设备怎样衔接；存储公司版图与可验证的投资问题

[S011] Recent Progress on 3D NAND Flash Technologies

Electronics / MDPI | 发布 未确认 | 检索 2026-10-15 | paper | search-excerpt

证据范围 Replacement gate charge-trap NAND and SiN storage layer; evidence limited to the published abstract

用于 NAND 的工艺、材料和设备怎样衔接

[S012] Material engineering to enhance reliability in 3D NAND flash memory

Device / Elsevier | 发布 未确认 | 检索 2026-10-15 | paper | search-excerpt

证据范围 综述摘要区分隧穿层、电荷俘获层、阻挡层、沟道材料及可靠性；证据限于论文摘要

用于 NAND 的工艺、材料和设备怎样衔接

[S013] Kioxia Commences Sample Shipments of 10th-Generation BiCS FLASH Devices

Kioxia | 发布 2026-07-03 | 检索 2026-10-15 | official | read

证据范围 BiCS10 1Tb TLC sample shipments, CBA and advanced stacking; not mass-production evidence

用于 从一块企业级 SSD 看懂整条产业链；存储公司版图与可验证的投资问题

[S014] Development of a 10th Generation 2Tb 4b/cell BiCS FLASH product

Kioxia | 发布 2026-07-15 | 检索 2026-10-15 | official | read

证据范围 Kioxia/Sandisk ISSCC 2026 2Tb QLC development and 332 stacked word-line features

用于 存储公司版图与可验证的投资问题

[S015] G9 NAND

Micron | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Micron has G9 3D TLC NAND and SSD products; interface-unit ambiguities not used quantitatively

用于 存储公司版图与可验证的投资问题

[S016] SSD Power Loss Protection: How Hardware and Firmware PLP Prevent Data Loss

Kingston | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Hold-up capacitors, voltage detection, volatile data/FTL preservation and firmware recovery; page labeled Apr 2026

用于 从一块企业级 SSD 看懂整条产业链；SSD、网络存储、HDD 与磁带怎样服务 AI

[S017] Edge SSD Controllers: SM2504XT and SM2508

Silicon Motion | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 SM2508 TSMC 6nm, controller NAND channels, ECC/LDPC, DRAM-less HMB product distinction and validation

用于 从一块企业级 SSD 看懂整条产业链；NAND 的工艺、材料和设备怎样衔接；SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S018] QLC NAND Technology Is Ready for Mainstream Use in the Data Center

Solidigm | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Page-write/block-erase, write amplification and workload-dependent QLC endurance

用于 从一块企业级 SSD 看懂整条产业链；SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S019] [Understanding SSD Performance Project](#)

SNIA | 发布 未确认 | 检索 2026-10-15 | standard | read

证据范围 Comparable device SSD measurements require specified platform/methodology; PTS and preconditioning context

用于 从一块企业级 SSD 看懂整条产业链；NAND 的工艺、材料和设备怎样衔接；从 SRAM 到磁带：AI 的内存与存储全景；SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S020] [NVMe over TCP Transport Specification](#)

NVM Express | 发布 未确认 | 检索 2026-10-15 | standard | read

证据范围 NVMe PCIe/RDMA/TCP transport split, queues and IP networking; 2026 revisions shown on current page

用于 从一块企业级 SSD 看懂整条产业链；从 SRAM 到磁带：AI 的内存与存储全景；SSD、网络存储、HDD 与磁带怎样服务 AI

[S021] [ScaleFlux FAQs](#)

ScaleFlux | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 主控集成的小块压缩与元数据处理；证据限于公开摘要

用于 SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S022] [ScaleFlux Reveals the CSD5000 for the AI Era](#)

ScaleFlux | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 FX5016 PCIe 5 主控与计算存储引擎；证据限于公开摘要

用于 SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S023] [Solidigm D7-PS1010 and D7-PS1030 Product Brief](#)

Solidigm | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Enterprise PCIe5 SSDs emphasize latency, QoS, endurance and AI pipeline qualifications

用于 从一块企业级 SSD 看懂整条产业链；SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S024] Platform optimization for performance and endurance

Solidigm | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 CSAL host FTL aggregates small writes using cache/capacity tiers, workload configuration and preconditioning disclosed

用于 SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S025] DPU-Native Distributed Flexible Storage Solution

Solidigm | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Large-capacity QLC FTL DRAM/IU tradeoffs, DPU resources, cache and NVMe-oF reference architecture

用于 从一块企业级 SSD 看懂整条产业链；从 SRAM 到磁带：AI 的内存与存储全景；SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S026] HBM4

Micron | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 DRAM die stack, TSV, good-die assembly/final test, 2048-bit bus and >2.8TB/s vendor specification

用于 从一块企业级 SSD 看懂整条产业链；从 SRAM 到磁带：AI 的内存与存储全景；存储公司版图与可验证的投资问题

[S027] Micron in High-Volume Production of HBM4 Designed for NVIDIA Vera Rubin, PCIe Gen6 SSD and SOCAMM2

Micron | 发布 2026-03-16 | 检索 2026-10-15 | official | read

证据范围 Micron says HBM4 36GB 12H volume shipments began calendar Q1 2026; product family includes SOCAMM2 and PCIe6 SSD

用于 从 SRAM 到磁带：AI 的内存与存储全景；存储公司版图与可验证的投资问题

[S028] Samsung Ships Commercial HBM4

Samsung | 发布 2026-02-12 | 检索 2026-10-15 | official | read

证据范围 Samsung commercial HBM4 statement, 1c DRAM/4nm base die, 2048 I/O and rated-speed boundaries

用于 从 SRAM 到磁带：AI 的内存与存储全景；存储公司版图与可验证的投资问题

[S029] SK hynix Ships Samples of 12-Layer Next-Gen HBM4E

SK hynix | 发布 2026-06-18 | 检索 2026-10-15 | official | read

证据范围 HBM4E sample status and Advanced MR-MUF; no assumption all HBM uses hybrid bonding

用于 从 SRAM 到磁带：AI 的内存与存储全景；存储公司版图与可验证的投资问题

[S030] DRAM: Cell Array and Periphery Logic Scaling

Applied Materials | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 DRAM cell capacitor/transistor, HAR capacitor etch, peripheral HKMG and low-k interconnect roles

用于 从 SRAM 到磁带：AI 的内存与存储全景

[S031] Atomic Layer Deposition

ASM | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Self-limiting conformal ALD, HKMG/MIM capacitor/dielectric applications and product platforms

用于 NAND 的工艺、材料和设备怎样衔接；从 SRAM 到磁带：AI 的内存与存储全景；存储公司版图与可验证的投资问题

[S032] AI SoC Memory Architecture Explained

Synopsys | 发布 2026-02-10 | 检索 2026-10-15 | official | read

证据范围 SRAM high-speed/small-block area/power tradeoffs and embedded-memory options

用于 从 SRAM 到磁带：AI 的内存与存储全景

[S033] CXMT Products

CXMT | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Official DDR5 and LPDDR product portfolio; does not establish HBM production or full server parity

用于 从 SRAM 到磁带：AI 的内存与存储全景；存储公司版图与可验证的投资问题

[S034] Nanya Q4 2025 Investor Conference

Nanya | 发布 2026-01-19 | 检索 2026-10-15 | official | read

证据范围 PDF p.19 discusses 128GB DDR5 RDIMM functional status and DRAM roadmaps

用于 从 SRAM 到磁带：AI 的内存与存储全景；存储公司版图与可验证的投资问题

[S035] [DDR/SDR Customized Memory Solution](#)

Winbond | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Specialty/mobile DRAM market positioning and product families

用于 从 SRAM 到磁带：AI 的内存与存储全景；存储公司版图与可验证的投资问题

[S036] [DDR5 Server DIMM Chipset](#)

Rambus | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 RDIMM RCD/PMIC/SPD/TS and MRDIMM MRCD/MDB; distinct command/address vs data roles

用于 从 SRAM 到磁带：AI 的内存与存储全景；存储公司版图与可验证的投资问题

[S037] [Memory Interface Products Selection Guide](#)

Montage Technology | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 DDR5 RCD/DB and supporting SPD/TS/PMIC products; evidence limited to the published abstract

用于 从 SRAM 到磁带：AI 的内存与存储全景；存储公司版图与可验证的投资问题

[S038] [DDR5 Solutions Product Selector](#)

Renesas | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Renesas DDR5 RCD, PMIC and SPD hub product offerings

用于 从 SRAM 到磁带：AI 的内存与存储全景；存储公司版图与可验证的投资问题

[S039] [Compute Express Link 3.0 White Paper](#)

CXL Consortium | 发布 未确认 | 检索 2026-10-15 | standard | read

证据范围 Memory semantics, CXL2 pooling vs CXL3 coherent sharing and fabric; not generic NVMe block transport

用于 从 SRAM 到磁带：AI 的内存与存储全景；SSD、网络存储、HDD 与磁带怎样服务 AI

[S040] GPUDirect Storage Design Guide

NVIDIA | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Direct DMA to GPU memory avoids CPU bounce buffer when supported; CPU still prepares communication

用于 从 SRAM 到磁带：AI 的内存与存储全景；SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S041] Introducing BlueField-4-Powered CMX Context Memory Storage Platform

NVIDIA | 发布 2026-03-16 | 检索 2026-10-15 | official | read

证据范围 Pod-level flash KV cache tier, DPU/NVMe-oF and software pre-staging; vendor architecture/performance claims

用于 从 SRAM 到磁带：AI 的内存与存储全景；SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S042] SK hynix Unveils First HBF Standard Specifications with Sandisk at FMS 2026

SK hynix | 发布 2026-08-04 | 检索 2026-10-15 | official | read

证据范围 HBF specification via OCP, tier between HBM/SSD and SK hynix NAND roadmap; not mass-adoption evidence

用于 SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S043] Sandisk to Collaborate with SK hynix on HBF Standardization

Sandisk | 发布 2025-08-06 | 检索 2026-10-15 | official | read

证据范围 Stated target HBF samples H2 2026 and AI device samples early 2027; targets not treated as completed delivery

用于 SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S044] Kioxia Announces New SSD Model Optimized for AI GPU-Initiated Workloads

Kioxia | 发布 2026-03-17 | 检索 2026-10-15 | official | search-excerpt

证据范围 GP Series Super High IOPS SSD development; evaluation samples to selected customers planned by end-2026

用于 SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S045] Seagate Delivers Highest Capacity Hard Drives with Mozaic 4+

Seagate | 发布 2026-03-03 | 检索 2026-10-15 | official | read

证据范围 Vendor says up-to-44TB HAMR volume shipments to two hyperscale customers, vertically integrated lasers

用于 从 SRAM 到磁带：AI 的内存与存储全景；SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S046] Mozaic: Nanophotonics and Superlattice Platinum-Alloy Media White Paper

Seagate | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 FePt alloy high anisotropy, crystalline underlayer/glass substrate, thin-film growth/annealing and recording technology

用于 从 SRAM 到磁带：AI 的内存与存储全景；SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S047] Western Digital Accelerates Storage Innovation for AI Era

Western Digital | 发布 2026-02-03 | 检索 2026-10-15 | official | read

证据范围 40TB UltraSMR ePMR qualification and H2-2026 planned volume; ePMR/HAMR dual roadmap

用于 从 SRAM 到磁带：AI 的内存与存储全景；SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S048] HDD Technology Evolution: Toshiba Article TSH186

Toshiba | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Toshiba FC-MAMR magnetic recording approach

用于 SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S049] IBM LTO Tape Drive

IBM | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 LTO10 30/40TB native vs 75/100TB compressed at 2.5:1 and drive/library operating context

用于 从 SRAM 到磁带：AI 的内存与存储全景；SSD、网络存储、HDD 与磁带怎样服务 AI；存储公司版图与可验证的投资问题

[S050] [Western Digital Form 8-K: Flash Separation](#)

Western Digital / US SEC | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 Legal separation date February 21 2025 and separate Sandisk flash entity; distinguishes Feb24 announcement/trading

用于 存储公司版图与可验证的投资问题

[S051] [Spin-transfer Torque MRAM Technology](#)

Everspin | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 MTJ magnetic-state persistent memory and product/application boundaries

用于 存储公司版图与可验证的投资问题

[S052] [Scaling the Memory Wall: The Rise and Roadmap of HBM](#)

SemiAnalysis | 发布 2025-08-12 | 检索 2026-10-15 | analysis | read

证据范围 Public article framing of capacity/bandwidth/latency/efficiency and HBM roadmap; paywalled data not relied upon

用于 存储公司版图与可验证的投资问题

[S053] [Has Sandisk Escaped Its Punishing Cycle?](#)

Seeking Alpha | 发布 未确认 | 检索 2026-10-15 | analysis | read

证据范围 Public abstract frames NAND cyclicalality, valuation and product mix; no rating or paywalled dataset adopted

用于 存储公司版图与可验证的投资问题

[S054] [SSD Basics v1.0](#)

Reddit r/NewMaxx | 发布 未确认 | 检索 2026-10-15 | forum | search-excerpt

证据范围 Community topic map of FTL/TRIM/wear leveling/garbage collection, used for questions only not hardware truth

用于 存储公司版图与可验证的投资问题

[S055] Does MUs quarter actually help SNDK?

Reddit r/stocks | 发布 未确认 | 检索 2026-10-15 | forum | search-excerpt

证据范围 Investor question separates NAND exposure from HBM/DRAM; no forecasts, prices or capacity claims adopted

用于 存储公司版图与可验证的投资问题

[S056] HBM: Materials Innovation Propels High-Bandwidth Memory Into the AI Era

Applied Materials | 发布 2024-05-20 | 检索 2026-10-15 | official | read

证据范围 TSV and HBM packaging materials-engineering steps; historical layer-count/market forecasts not used as current

用于 从 SRAM 到磁带：AI 的内存与存储全景；存储公司版图与可验证的投资问题

[S057] Kioxia Commences Sample Shipments of 9th Generation 512Gb TLC

Kioxia | 发布 2025-07-25 | 检索 2026-10-15 | official | read

证据范围 9th-gen cost-oriented 120-layer process based on 5th generation plus advanced CMOS/CBA; generation not layer ranking

用于 从一块企业级 SSD 看懂整条产业链；存储公司版图与可验证的投资问题

[S058] Pure-Metal Replacement Gate for Reliable 30nm Pitch Scaled 3D NAND Flash

IEEE | 发布 未确认 | 检索 2026-10-15 | paper | search-excerpt

证据范围 Research Mo pure-metal gate, blocking oxide/high-k liner; experimental result not universal production recipe

用于 NAND 的工艺、材料和设备怎样衔接

[S059] Synopsys Foundation IP Product Overview

Synopsys | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 嵌入式 SRAM、寄存器文件、TCAM 及工艺相关 memory IP 的类别；证据限于公开摘要

用于 从 SRAM 到磁带：AI 的内存与存储全景

[S060] [Introduction to Memory](#)

Micron | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 DDR, LPDDR, GDDR and HBM are DRAM categories; modules and interface form factors distinguished

用于 存储公司版图与可验证的投资问题

[S061] [Samsung Unveils Next-Gen 3D-Memory Vision at FMS 2026](#)

Samsung | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Samsung DRAM/NAND/enterprise storage portfolio, V10 BV-NAND introduction versus concept-model status

用于 存储公司版图与可验证的投资问题

[S062] [About SK: Investing in Industries Around the Globe](#)

SK | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Identifies Solidigm as SK hynix subsidiary with US memory operations

用于 存储公司版图与可验证的投资问题

[S063] [About Memblaze](#)

Memblaze | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Beijing company offers PBlaze enterprise NVMe SSD products; company application claims not independent adoption measurement

用于 存储公司版图与可验证的投资问题

[S064] [Phison Electronic Corporation 2025 Annual Report](#)

Phison | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Phison SSD controllers E28 and E31T, PASCARI enterprise SSD brand and own development portfolio

用于 存储公司版图与可验证的投资问题

[S065] Enterprise SSD Controllers

InnoGrit | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 Enterprise PCIe SSD controller offerings including IG5668; not customer shipment-volume evidence

用于 存储公司版图与可验证的投资问题

[S066] Micron GDDR7 Memory Product Brief

Micron | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 GDDR7 DRAM addresses AI/gaming/HPC, distinct from HBM; no extrapolation of system bandwidth without bus configuration

用于 存储公司版图与可验证的投资问题

[S067] Atomic layer etching of ZrO₂ thin films for DRAM capacitors using NF₃ plasma and TiCl₄

Ceramics International / Elsevier | 发布 未确认 | 检索 2026-10-15 | paper | search-excerpt

证据范围 Research ZrO₂/Al₂O₃/TiN capacitor films and ALE; lab structure is not evidence of particular commercial customer recipe

用于 从 SRAM 到磁带：AI 的内存与存储全景

[S068] BiCS FLASH Generation 8 with CBA Technology

Kioxia | 发布 2025-01-22 | 检索 2026-10-15 | official | read

证据范围 NAND package example uses 16 stacked 2Tb dies; package-die count distinct from cell layer count

用于 从一块企业级 SSD 看懂整条产业链；NAND 的工艺、材料和设备怎样衔接

[S069] Capacity of TS4500 Supported LTO Tape Cartridges

IBM | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 LTO10 30/40TB native and 75/100TB compressed, 2.5:1 capacity assumption

用于 SSD、网络存储、HDD 与磁带怎样服务 AI

[S070] [Western Digital Completes Planned Company Separation](#)

Western Digital | 发布 2025-02-24 | 检索 2026-10-15 | official | search-excerpt

证据范围 February 24 public announcement of flash business separation, distinct from February 21 legal separation date

用于 存储公司版图与可验证的投资问题

[S071] [Understanding SSD Endurance: TBW and DWPD](#)

Kingston | 发布 未确认 | 检索 2026-10-15 | official | search-excerpt

证据范围 DWPD = TBW divided by user capacity and warranty days; endurance depends on workload and write amplification

用于 SSD、网络存储、HDD 与磁带怎样服务 AI

[S072] [V10 BV-NAND: Pushing the Boundaries of AI Storage Beyond 400 Layers](#)

Samsung | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 V10 separates 3-stack HARC Merge from cell/periphery wafer bonding; etch, conformality, clean, stress integration

用于 NAND 的工艺、材料和设备怎样衔接

[S073] [EDSFF: A New Form Factor for Next Generation Servers and Storage](#)

Kioxia | 发布 未确认 | 检索 2026-10-15 | official | read

证据范围 EDSFF E1/E3 form factor family and legacy M.2/U.2 context; physical form distinct from lane/interface configuration

用于 SSD、网络存储、HDD 与磁带怎样服务 AI