

高盛闭门会：鲁宾 BOM 分析：晶圆代工、硅电容与中国基础模型，TMT 周度网络研讨会 20260526

AI 驱动成熟制程满产，Rubin 机架价值量倍增与 AI 软件商业化提速

摘要

- 成熟制程晶圆厂产能利用率持续提升，AI 电源 IC、硅光及先进封装驱动下半年有望满产，联电评级上调至增持。
- Rubin 机架 ASP 较 GB300 翻倍至 780 万美元，HBM 成本占比升至 25%-30%，PCB、MLCC、ABF 价值量分别增长 233%、182%、82%。
- ODM 厂商附加值预计增长 35%-40%，虽毛利率因成本上升降至 1.9%，但利润绝对值增加，2026 年后转向寄售模式将改善现金流。
- MediaTek 2nm TPU (HumanFish) 2028 年出货量预期由 100 万片上调至 250 万片，采用英特尔 EMIB 技术，ABF 供应商已基本确定。
- 硅电容市场供需缺口仍超 10%-20%，APMemory 凭借技术领先及力积电产能支持，预计 2027 年底前维持近 100% 份额，2028 年面临 Samco 竞争。
- 中国 AI 软件转向推理经济性与商业化：智谱 AI 架构降低 33% 光模块开支，金山办公 WPS AI 3.0 带动定价权回升，金蝶 AI 套件目标收入 10 亿元。

Q&A

市场对成熟制程晶圆代工厂的最新看法是什么，尤其是在 AI 需求背景下，其产能利用率、定价能力以及相关公司的评级有何变化？

尽管过去一年市场担忧消费端需求下行会对成熟制程节点产生负面影响，但近两个季度，尤其是本次财报电话会议后，亚洲主流晶圆代工厂，包括联电、世界先进、华虹及中芯国际，均表示其产能利用率在持续提升且市场份额在扩大。这一驱动因素已超越消费需求，被认为与 AI 相关，例如 AI 电源 IC、硅光芯片或先进封装相关的半导体需求。目前 AI 电源 IC 市场规模可能已达到 150 亿美元，并有望随 AI GPU 同步增长。基于此，预计下半年成熟制程晶圆厂将达到满负荷运转。因此，联电的评级被上调至增持。相反，对于 Synology 等消费类模拟芯片设计公司，由于其可能无法完全转嫁新增的晶圆代工成本，其评级被下调至减持。对于环球晶圆等化工企业，尽管市场预期下半年晶圆价格上涨，但调研显示

涨幅将非常温和，因为面对台积电等大型客户时提价难度极大，故其评级被下调至持有。对于先进制程领域，台积电目前仍是超配标的，但其正面临来自英特尔晶圆代工业务日益激烈的竞争，这可能有助于美国政府争取苹果或特斯拉等美国客户。同时，随着存储行业进入上行周期，三星晶圆代工似乎将获得更多预算以补贴其业务竞争，历史上三星曾与 AMD、高通和特斯拉合作。因此，市场可能低估了竞争对台积电的影响，短期内其股价实现第二次回调并非易事。

MediaTek 在 2 纳米 TPU 产品上的最新进展如何，其出货量预期和生态系统支持情况是怎样的？

MediaTek 的 2 纳米 TPU 产品，代号为“Human Fish”，采用了英特尔的 EMIB 技术。近期调研显示，相关生态系统即将进入硅产能阶段。从桥接芯片的生产以及 ABF 基板的供应来看，目前供应商包括 EB Tan、Shinko、Austria Microsystem 和 Unimicron，甚至 SEMCO 也可以成为该产品的 ABF 供应商。基于这些进展，原先预计 2028 年 2 纳米 TPU 的出货量仅为 100 万片，现已上调至约 250 万片。MediaTek 预计 2028 年该芯片的市盈率为 15 倍，显示未来仍有上行空间。

与 GB300 机架相比，Rubin (VR200/V072) 机架的成本构成、平均售价以及 ODM 厂商的附加值有何变化？

从 ODM 采购的 Rubin (VR200/V072) 机架成本估计约为 780 万美元，而 GB300 机架成本约为 400 万美元，机架级别的平均售价几乎翻倍。若超大规模企业直接从内存供应商采购 SoC 内存，ODM 机架的 ASP 将降至约 670 万美元。在物料清单方面，最大的变化是内存占比从 Grace Blackwell 系统中的中高个位数大幅提升至 Rubin 系统成本的 25%-30%，这使得 GPU 占机架成本的比例从 60%-70% 降至约 51%。尽管新一代产品的 GPU 价值仍在增长，但下游组件也迎来显著增值。其中，PCB 是最大受益者，其价值较上一代 Blackwell 300 系统增长 233%，这得益于新板卡（如中板、BlueField、ConnectX）的引入以及现有板卡（如计算板、交换板）制造复杂度的提升。MLCC 价值增长约 182%，ABF 载板价值增长约 82%。对于 ODM 厂商而言，尽管市场担忧机架标准化可能降低其附加值，但自下而上的分析显示，每机架的 ODM 附加值（按美元内容价值计算）预计将增长 35% 至 40%。虽然由于 GPU 和内存成本上升，整体毛利率预计将从 GB300 的 2.7% 降至 VR200 的 1.9%，但利润绝对值将增加。附加值的提升不仅体现在机架级组装测试，还包括计算板及其他板卡的组装测试。需要注意的是，并非所有 ODM 都能处理从板级组装到机架级的全流程。

随着 AI 服务器机架规模和成本的扩大，ODM 厂商在营运资本方面面临哪些挑战，行业内出现了何种应对模式？

随着机架规模扩大，ODM 厂商将面临更大的资产负债表压力和营运资本问题。为应对这一挑战，ODM 与其超大规模客户之间关于寄售模式的讨论增多。鸿海率先提及此模式，随后广达和双赢也跟进，并提到 2026 年下半年可能有项目转向寄

售模式。如果客户更愿意承担营运资本负担，这将成为 ODM 现金转换的长期积极催化剂，并改善其风险回报状况。在 ODM 板块中，基于股价目标的上行空间，双赢是热门标的，其次是纬创、广达和鸿海。该板块 2026 年迄今表现逊于台湾整体股市，当前估值基于 2027 年市盈率接近 10 倍，风险回报具有吸引力。

Samco 进入硅电容业务对 AP Memory 构成了怎样的竞争威胁，AP Memory 又具备哪些竞争优势？

Samco 于 2026 年 5 月 20 日宣布与一家全球客户（被认为是英特尔）签订了价值约 10 亿美元的硅电容供应合同，期限为 2027 年 1 月至 2028 年 12 月。Samco 将采用与 AP Memory 类似的堆叠工艺，并可能与华邦电子合作进行晶圆代工，这使其成为 AP Memory 的直接竞争对手。尽管如此，市场对 AP Memory 的前景无需过早担忧。首先，市场将持续存在结构性供应短缺。需求增长主要来自英特尔的 EMIB 技术以及高端智能手机 SoC 从 2026 年第二季度开始采用硅电容。预计未来 2-3 年，即使计入 Samco 的新增产能，供应缺口仍将保持在 10%-20% 以上。其次，AP Memory 拥有差异化优势。技术上，其第四代架构电容比 Samco 已知路线图高出约 30%，并提供高度灵活的集群尺寸。在晶圆代工方面，AP Memory 与力积电的长期合作能提供更专属的产能和定制化支持。考虑到电容器在供应链中处于上游，新供应商签订合同需要 1 至 1.5 年的准备期，预计 AP Memory 在 2027 年底将保持近 100% 的市场份额。到 2028 年，Samco 有望占据约 30% 的总市场份额。

中国 AI 模型和软件行业近期有哪些值得关注的发展，它们分别反映了怎样的商业化趋势？

近期中国 AI 和软件行业正从功能发布转向更可衡量的商业化和效率提升阶段，主要体现在三个方面：一、智谱 AI 发布的 Z-Cube 网络架构。该架构旨在提升推理效率，据称可将交换机与光模块资本开支降低 33%，GPU 推理吞吐量提升 15%，并在 GLM 5.1 编码测试中将首 token P99 延迟降低 40.6%。这表明行业焦点正从训练能力转向推理经济性，通过系统级架构革新而非单点改进来提升效率，以支持更快的响应速度和更低的基础设施成本。二、金蝶发布的灵境 AI 操作系统。灵境定位为企业级 AI 操作系统，旨在帮助企业开发、构建和部署智能体与技能，并与 ERP 系统打通。其商业模式结合了按席位订阅和按使用量计费的信用制。金蝶有信心在 2026 年实现 10 亿元的 AI 套件收入目标，该目标不包含预计将贡献额外收入的灵境。这显示了 AI 正被打造为企业软件的深度工作流层。三、金山办公 2026 年 618 促销观察。数据显示，金山办公可能正在重新获得消费市场的定价权。与 2025 年同期相比，2026 年 618 活动的折扣幅度收窄，而平均标价提高了 2%；与 2025 年双 11 相比，折扣同样更温和，平均标价则提高了 18%。这表明在 WPS AI 3.0 推出后，公司对自身价值主张更具信心，并通过产品捆绑（如 WPS 超级会员捆绑生态权益，尊享会员突出 AI 功能）推动用户转化和高阶



