

算力尽头是光

AI光通信的技术、产能与2026—2028供需账本

蜜蜡

蜜蜡 著 · 串行写作与逐章复核 · 事实、规划与情景分别标注

导读

AI光通信的技术、产能与2026—2028供需账本

一家公司宣布AI收入翻倍，另一家公司宣布光模块工厂扩产。把两条新闻摆在一起，很容易得到一个结论：光一定越来越缺。

但如果新一代网络把两个800G端口合成一个1.6T端口，带宽翻上去，盒子未必多。如果工厂说的是年底月产能，全年也未必已经交出十二个月的货。更麻烦的是，有的“光”藏在模块里，有的搬到了交换芯片旁边，还有的只是让光改变行进方向。

这本书围绕一个问题展开：**AI多搬走的一份数据，最后会变成谁的几件产品、多少收入？**

从一根线走到一张账本

前五章从等待数据的GPU出发，拆开模块，解释CPO、硅光、MPO、LPO和OCS，并追到磷化铟与薄膜铌酸锂。种子中提到的“薄膜磷酸锂”，在这条高速调制技术路线里通常应为“薄膜铌酸锂”；两种材料不能混写。

第六章读NVIDIA与Alphabet截至2026年9月9日已披露的最新季度资料，辨别财报能告诉我们的增长，以及它没有披露的产品出货结构。第七、八章走进供应商，重点核验龙头和剑桥、东山、新菲光、华工、联特、汇绿、恩达通等扩张者。

第九、十章建立2026—2028供需与价格情景表。最后两章把中美上市公司的业务放回材料、芯片、模块、连接和测试的位置，给出按路线划分的优先观察对象，以及推翻判断的条件。

先认识书中三种数字

- **已披露事实：**季度收入、已发生交付、公开产品资料。可以追溯，但仍要辨别业务与统计口径。
- **企业规划：**产能目标、预计量产、客户验证进度。是带有条件的计划，不能提前确认为全年销售。
- **作者情景：**用于计算的未来产品占比、有效供给、合同价格。明确标为假设，允许读者替换，不伪装成公司指引或行业数据库。

公开资料尚不足以拼出全球同口径、分速率、去重后的有效产能。因此，本书没有给出一个冒充事实的全球精确缺口。第九章的数字是未充分校准的压力测试，目的是展示哪些条件造成结构性短缺，哪些条件使它消失。第十章也把公开零售价与假设的大客户成交价分开。

引用尽量回到公司公告、监管文件和产品资料；少数产能消息仅见媒体采访，书中单独标明。私人公司的客户未公开，就保留未知；参股一家光企业，也不把它全部收入算进上市公司。

大白话

可以先读懂光，也可以从公司名单倒着查。但读到一个很大的数字时，请先问：它数的是盒子、端口、芯片，还是钱？

现在从第一章开始。先看一张GPU，为什么会等另一张GPU。

第一章 · 一张GPU为什么等另一张

一台昂贵的计算机最令人心疼的状态，不是算错，而是等待。想象一百位厨师共同做一道菜：每个人切得飞快，却必须等隔壁传来配料和称重结果。再买一百把刀，未必能让晚餐提前。AI机房的光通信，解决的正是这条传菜通道。

算得快，把搬运变成了主角

GPU是擅长同时做大量相似计算的处理器。训练大模型时，工作往往拆给很多张GPU；各自算出的结果需要汇合，才能继续下一步。带宽是单位时间能运送多少数据，像道路每秒通过多少辆车；延迟是一次传递要多久，像第一辆车到终点的时间。拓宽道路不一定让第一辆车更早到达，两者不能混为一谈。

假设每一轮计算用8毫秒，交换数据用2毫秒，而且两者不能重叠，那么一轮需要10毫秒。换成快一倍的芯片，计算降到4毫秒，一轮仍需6毫秒，整体只快了约1.67倍。若交换也降到1毫秒，才得到完整的两倍提升。这是一个教学算例；真实系统会让部分计算和传输同时进行，但等待越多，网络升级越值钱的因果不变。

推理是用训练好的模型生成答案，也并不天然“不需要网络”。模型可能大到一张卡装不下；不同部分可能住在不同机器；一次请求还可能调用多个模型。反过来，若小模型完全放在一台机器里，外部通信就少。不能把“推理占比提高”直接翻译成光模块需求上升或下降，必须追问：每次请求的数据到底跨了几台机器？

先分三层，再数光

Scale-up是把一组处理器连得像一台大机器，强调低延迟和密集交换。Scale-out是把更多机器或机柜接入一个集群，强调规模。Scale-across则把相隔更远的数据中心连起来。它们是通信范围与组织方式，不是规定必须用铜还是光的材料说明书。

短距离铜连接很有竞争力：省去了电变光、光再变电的器件，安装和维护也熟悉。距离拉长、速率提高之后，铜线损耗和补偿电路的代价增加，光纤才越来越划算。因此一台全新的

AI机柜里，可能既有铜线，也有跨柜光纤，更远处还有连接园区的光网络。把所有内部连接都计为光模块，会先把需求夸大一遍。

英伟达的互连产品页同时列有可插拔光器件、铜缆及共封装光学，正说明现实采购是按距离和系统约束配套的组合。[1] 研究产业时，先画清哪些线越过机柜边界，再统计这些线两头的器件，会比先猜“每卡配几个模块”稳得多。

一根线，不等于一个模块

光模块是把电信号与光信号相互转换的一组封装器件，常见外观是插在交换机或网卡上的银色盒子。一条两端都用可插拔模块的双向光链路，通常需要两只模块。双向，并不意味着再乘两倍：每只收发模块本来就同时包含发送和接收功能。

看一个刻意简化的集群：1,024台服务器，每台一条光链路接到第一层交换机，共1,024条；交换机之间另有512条光链路。一共1,536条，两端可插拔对应3,072只模块。这还没有加备件，也没有存储网络。若厂商给出的数字已经是两端端口数，就不能再乘二。

如果某些交换机端改成集成光引擎，光链路仍在，两端的光电转换仍在，却不再都是银色盒子。这时“可插拔模块只数下降”和“光通信带宽增长”可以同时为真。产业变化经常先改变采购清单的名字，再改变最终支出。

大白话

先数哪些数据需要走远，再数每根线的两个端点，最后判断端点长成什么产品。别从GPU收入直接跳到光模块只数。

本书要结清的账

后面会沿这条线打开模块、比较技术，再回到季度报告与厂商产能。我们会把资料分成三种：已经发生的事实、公司宣布的计划、为计算而设的情景。没有公开披露的客户名字、产品份额和合同价，会明确留空。留空并不是没有分析，而是让下一次新信息进来时，知道该填在哪个位置。

本书资料核验截至2026年9月9日；2026年尚未结束，2027—2028年更不能被写成已知结果。贯穿全书的问题只有一个：当AI预算继续增加，哪一种光学能力最先成为瓶颈，又由谁

把解决瓶颈的价值留下来？

资料与核验

[1] NVIDIA, Optical Transceivers and Cables for AI Networking, 产品页, 访问于 2026-09-09: <https://www.nvidia.com/en-us/networking/interconnect/>。本章数量均为教学算例, 不是厂商出货预测。

第二章 · 打开那个银色小盒子

把一只800G光模块放在桌上，它看起来像一件金属配件。但机壳里真正忙碌的是一场接力：整理电信号，给光写上数据，让光穿过纤维，再把它读回电信号。理解这场接力，才知道一张扩产公告到底扩了什么。

800G是八条车道，还是四条？

800G表示标称每秒800吉比特，1.6T是它的两倍；这里的比特不是字节，一个字节通常有8比特。产品名称中的速率一般不能直接当作应用能拿到的净数据吞吐量，编码和纠错要占掉一部分传输能力。也不能把发送与接收相加，宣称一只800G模块变成1.6T。

Lane是并行通道，像一条独立车道。800G可以通过8条标称100G通道拼成，也可以由4条200G通道拼成。电侧与光侧的通道数量还可能不同，由内部电路转换。于是“800G扩产100万只”仍没说清：它消耗的是100G还是200G光芯片，每只用几颗，哪一种测试设备？

PAM4是用四档信号强弱表示数据的方法，一次符号可表示两个比特。与只用高低两档相比，它在相同符号变化速度下装入更多信息，但四档更容易被噪声搅混。道路没有拓宽，只是每辆车多装了一箱货，对识别车辆的要求更高了。

沿着信号走一遍

DSP是数字信号处理芯片，负责把变形、时序不齐的高速电信号整理得更容易读。驱动器把信号变成足以推动激光器或调制器的电流或电压。激光器提供光；调制器根据数据改变光的强弱或其他特征。接收端的光探测器把光变成微小电流，TIA再把这个小电流转换、放大成后级能处理的电压。

不是每一种模块都有相同结构。某些设计直接调激光器，某些让激光持续发光再单独调制；某些减少DSP处理，把更多任务交给主机。Marvell的产品资料同时覆盖DSP模块和线性驱动模块需要的电芯片，说明“少一颗DSP”不等于“模块里不再需要电芯片”。[1]

还有容易被忽略的装配：把几微米尺度的光路准确对接，把发热带出机壳，让产品在温度变化后仍可靠。耦合就是把光从一个器件尽量少损失地送进另一个器件。若芯片性能很好，装

配却让一半光走丢，模块仍不合格。封装工厂的价值不是把零件塞进盒子，而是让它们在量产中保持精确合作。

MPO不是一种发光技术

MPO是多芯光纤推入式连接器，把一排光纤一次插接，像把许多独立插头排成一个整齐的插座。MTP是US Conec的MPO兼容连接器品牌，不是与MPO平行的另一种光传输原理。[2]连接器主要解决多根纤维怎样准确接上，不负责电光转换，也不是硅光的替代品。

DR通常指数据中心短距离单模并行光接口这一类规格；FR通常对应比DR更远的单模接口类别，常用不同波长复用节约光纤；准确距离、波长与纤芯数必须看具体型号。波分复用是让不同“颜色”的光共走一根纤维，接收后再分开。平行多纤方案少一些合分波器件，却可能用更多纤芯和高密连接器；波分方案节约光纤，却把成本移到光学器件和波长控制上。

因此同样是1.6T，DR8与采用合分波结构的产品不能只按速率比较价格。还应核对传输距离、接口、工作温度、是否带DSP和系统认证。同样的箱子容量，冷链车与普通货车不是同一种报价。

把上下游串起来

最上游提供晶体材料、衬底、特殊气体；中间把材料做成激光器、调制器、探测器和电芯片；再由光器件、封装和模块厂装配测试；连接器与光纤厂把端点接起来；系统商将它们装进交换机、服务器与机柜；最终由云厂商等部署。AOC是两头固定带光电转换器件的有源光缆，通常作为一整条采购。统计时若把一条AOC再拆成“两只模块加一条AOC”，需求就重复了。

大白话

速率告诉你运多少，接口告诉你怎么接，芯片结构告诉你用什么料。三张标签缺一张，产能数字和价格就可能对不上。

下一章将把CPO、硅光和LPO放到不同的坐标轴上。它们经常同时出现在一只产品里，新闻标题却喜欢把它们写成只能活一个的对手。

资料与核验

[1] Marvell光互连电芯片产品页；[2] US Conec：MPO与MTP常见问题。访问于2026-09-09；接口实例可交叉核对FS 1.6T产品规格。

第三章 · CPO不是硅光的同义词

如果有人问“未来汽车会选择电动、敞篷还是自动挡”，你会发现选项不在同一层：一个说动力，一个说车身，一个说变速方式。把CPO、硅光和LPO排成一场只能有一个赢家的比赛，也犯了同样的错误。

第一根轴：光电转换放在哪里

可插拔光模块把光电转换装进可以单独插拔的盒子，放在设备面板。高速电信号要从主芯片穿过电路板走到盒子。它的优点是采购、替换和维护方便；挑战是速率越来越高之后，这一小段电路也开始昂贵。

CPO即共封装光学，把光引擎与交换芯片或其他计算芯片放到共同的封装集成体系中，缩短高速电信号必须奔跑的距离。光引擎是负责光电转换的核心组件，不必自带可插拔盒子的全部外壳和管理接口。CPO改变的是位置与封装关系，不保证任何一种特定材料垄断，也不自动代表处理器之间已经全用光。

NPO是近封装光学，把光学器件放得很靠近主芯片，但不完全采用同一种共封装集成方式。它试图在电路长度、装配复杂度与可维修性之间折中。近年还有XPO等产品命名，必须看厂商实际接口、供电和维修设计，不能仅凭缩写中的字母推定统一标准或成熟度。

第二根轴：光学器件用什么做

硅光是在硅基工艺平台上制造导光、调制、探测等光学功能，把原来分散的小器件集成起来。它既能放进可插拔模块，也能做CPO光引擎。硅不擅长高效发光，所以系统常需要另外的激光源，下一章会追到磷化铟。这里先记住：买硅光，并不是从采购表里删除所有其他光学材料。

英伟达的硅光产品页明确把硅光用于CPO交换机；剑桥科技2026年中期报告则同时讨论硅光可插拔模块与其他封装方案。[1][2] 两类产品都可以叫硅光，卖法却不同。前者可能由平台商更深地主导封装采购，后者仍由模块厂交付盒子。

第三根轴：谁负责把电信号整理干净

LPO是线性驱动可插拔光学，通常移除模块中传统的重定时DSP链路，更多依赖主机与整条链路的协同处理，以降低功耗、延迟及器件成本。线性并非“没有处理”，而是尽量忠实传递信号的变化，让主机承担更多修正工作。

它像取消中途的精细整队站，让车辆直接上路。省了时间和站点费用，却要求出发时队形整齐、道路质量足够好、终点接收能力足够强。因此主机电路、光器件、连接损耗和不同供应商之间的互通验证更重要。它能覆盖的距离和使用条件应看具体产品测试，不能用一个固定米数概括所有LPO。

LRO一般指线性接收光学，常见结构在发送侧保留重定时，在接收侧走线性路径，也常与TRO等名称关联。它用部分处理能力换取部分节能，是工程折中；不同实现的功能边界要看数据手册。Marvell同时提供DSP和线性相关产品，体现了供应商对多种方案并行的准备。
[3]

把三个坐标填进采购表

- 硅光可插拔加DSP：材料平台变了，面板盒子和数字信号处理仍在。
- 硅光LPO：仍是可插拔，光学集成与减少模块内处理同时发生。
- 硅光CPO：光学向主芯片靠近，采购对象从部分盒子转为光引擎、外置光源和精密封装。
- 采用分立激光器的传统可插拔：仍可随着单通道速率升级继续改进，不因新缩写出现而立即消失。

若一条链路只有交换机端采用CPO、服务器端仍可插拔，传统模块只减少一端。若原来就使用集成器件，则不存在再替代一只盒子的增量。预测不能把CPO端口数量直接乘二当作可插拔模块损失。

节能也要统一分母。某厂商报告光互连部分节能数倍，不等于整个AI机房电费下降同样倍数，因为GPU、内存、供电与制冷仍在用电。每比特能耗是传一比特要花多少能量；每比特能耗下降时，总流量扩大仍可能让设备总功耗上升。

问位置、问材料、问信号处理，再问谁修、谁供货。做完这五个问题，新路线才会变成可计算的商业变化。

CPO最值得关注的不仅是少卖多少盒子，而是谁开始拥有新产品的接口和验证权。模块厂如果跟随客户进入光引擎封装，可能换一种形态继续增长；只有装配能力、没有关键工艺和客户关系的厂商，更容易被价格竞争挤压。这是产业推论，不是任何一家公司的既定命运。

资料与核验

[1] NVIDIA硅光产品页；[2] 剑桥科技2026中期报告；[3] Marvell光DSP与线性驱动产品。均于2026-09-09核验。

第四章 · 光从哪里来，数据刻在哪里

“硅光会不会把磷化铟替代掉？”这句话像问：装了更好的水管，是否就不用水泵了。先把发光、调光、导光、收光四件事分开，材料之间的关系便清楚很多。

磷化铟：既是地基，也能长出发光器件

磷化铟，InP，是由铟和磷组成的化合物半导体材料，适合构建通信波段的激光器以及其他光电子器件。这里的“适合”不是把一块原料通电就能得到光模块。先长晶、切片和抛光形成衬底，也就是器件的地基；再通过外延，在地基上生长厚度和成分受控的薄层；最后做出器件、电极和光路，再测试封装。

EML是电吸收调制激光器，通常把发光部分和通过改变吸收强弱写入数据的部分集成在一起。CW激光器则持续提供光，数据由别处的调制器写入。硅光系统采用CW光源时，可能减少对某些分立EML的需求，同时增加对高功率、高可靠性CW光源的需求。这是InP产品结构的改变，不能简化为InP需求归零。

Coherent在OFC 2026展示的InP产品同时包括面向1.6T模块的200G EML，以及面向CPO和硅光可插拔的高功率CW激光器。[1] 同一家材料与器件厂，可以押注不止一条下游路线。不过两类器件的工艺、筛选和客户验证不同，产线能力并不自动一比一互换。

硅光：集成能力很强，不是万能材料

波导是芯片上引导光前进的细小通道，可以理解为光的微型水渠。硅光擅长把许多波导和功能集成在一个平台上，借助成熟的半导体制造体系复制。微环调制器用一个很小的环形光结构控制光，面积小，但温度和波长漂移需要管理；马赫—曾德尔调制器把光分成两路，再通过控制两路相位关系改变输出，通常更占空间，也有自己的性能优势。

制造硅光芯片并不等于制造最先进的逻辑处理器。关键可能在光损耗、耦合精度、温度稳定性和良率，而不只在制程名字有几纳米。良率就是投入一批产品后，最终有多少比例合格。若一个设计省了芯片面积，却让光纤对准更难，整机成本未必下降。

正确的名字是薄膜铌酸锂

种子中写的“薄膜磷酸锂”，在这里应改为薄膜铌酸锂，TFLN：把铌酸锂材料制成很薄的一层，用来做高性能光学器件。它的化学式是 LiNbO_3 ，关键元素是铌，不是磷。它也不是电池里的磷酸铁锂。看错一个字，产业链就会走进另一座工厂。

铌酸锂的重要本领是电场能迅速改变它对光传播的影响，便于把高速电数据写到光上。薄膜结构把光和电场约束得更紧，能够实现紧凑、低驱动电压或高带宽的设计。2026年发表的一项研究展示了宽波段高速调制能力。[2] 这是器件性能证据，不等于这些器件已成为所有1.6T量产模块的标准配置。

它的产业链从高质量晶体出发，经薄膜制备和键合，再到波导刻蚀、电极制造、封装及可靠性测试。键合是把两种材料层牢固且精确地结合起来。技术难点不只是“薄”：膜厚均匀、界面缺陷、刻蚀侧壁、光纤耦合和长期温度表现，都可能影响最终成本。

薄膜铌酸锂既可能作为独立调制器，也能与硅或氮化硅平台结合。异质集成就是让不同材料各做擅长的工作，再拼成系统。研究已经展示这类组合。[3] 因而硅光与薄膜铌酸锂也未必是两支只能选一支的队伍。

材料很关键，价值却不按重量分配

一只模块里的衬底原料可能重量极小，却能卡住交货；大量普通金属和塑料可能重量很大，却不决定性能。稀缺原材料涨价，也不代表所有材料企业都得到同样利润：客户采购的是规定纯度、缺陷密度和一致性的合格产品，不是元素周期表上的名字。

做供给模型时，应将衬底片数转换成合格芯片：晶圆片数乘每片可切芯片数，再乘前道良率、封装良率和客户合格比例。若晶圆直径翻倍，理论面积约四倍，但实际合格芯片并不必然四倍，边缘损失、工艺成熟度和芯片尺寸都会改变结果。

大白话

材料路线的受益要沿着“哪种器件、用多少、谁能合格交付”往下算。拥有矿、拥有晶体、拥有高速激光器量产能力，是三件不同的事。

判断2026—2028年的路线时，InP光源是有现实产品支撑的关键环节；薄膜铌酸锂是值得跟踪的高带宽候选，但具体收入取决于量产设计导入。把实验室纪录直接乘全球模块销量，会把潜力误写成订单。

资料与核验

[1] Coherent InP产品公告，2026-03-17；[2] Li等，宽波段薄膜铌酸锂调制器，2026-01-08；[3] 薄膜铌酸锂与氮化硅微转印集成研究，2025-04-05。本章产业判断为基于器件分工的推论。

第五章 · Google为什么给光修道岔

传统交换机像分拣站：把每一件包裹拿下来，看地址，再放到下一条传送带。Google的一类光网络则像铁路道岔：先把路接好，让一长列货车直接通过。区别不在“有没有光”，而在中间是否每次都要把光变成电来处理。

道岔究竟做了什么

OCS是光路交换机，建立输入光路与输出光路之间的连接，让数据以光的形式穿过中间节点。它通常不逐个读取数据包目的地址。电分组交换则把数据组织成一个个包，由电芯片读包、排队、选择出口。OCS是改变线路，电交换是处理经过线路的每一批货物。

MEMS是微机电系统，可以把微小可动镜面做进器件，通过转动镜面把光送向不同方向。这是实现OCS的一条路线，并非唯一方案。机械动作有重配时间，光路也有损耗；OCS不意味着零能耗、零延迟或任何流量都能瞬时换路。

如果两组机器将在较长时间内频繁交换数据，先花一点时间接通专用光路，之后直接传递，就可能比不断经过多级电分拣更划算。若每个小包都要随机去不同地方，或者需要复杂排队和拥塞处理，电交换仍有价值。真实网络常把两者组合，而不是在机房门口选一个阵营。

TPU给了道岔一个合适的任务

TPU是Google围绕机器学习设计的专用处理器。大量TPU连接在一起时，任务怎样分配、芯片怎样互通，能够与软件共同设计。拓扑就是谁与谁相连的关系图。与只能买通用设备的客户相比，拥有芯片、网络和调度软件的厂商，更有条件让任务适应网络，也让网络适应任务。

Google的TPU v4论文已说明OCS用于动态重配互连；它不是2026年才突然出现的概念。[1] Ironwood技术说明进一步介绍：64颗芯片组成一个基本单元，通过OCS连接成更大的系统，示例最大为9,216颗芯片；系统还能绕开不健康单元、接入备用单元。[2] 这些是该代系统的设计事实，不能原封不动套在未来所有TPU上。

ICI是TPU之间专用的芯片互连，DCN是数据中心网络。前者把处理器组织为紧密协作的计算系统，后者还能承接更大范围的连接。数模块时，应分别计算ICI越过基本单元的链路、普通网络链路和存储链路，避免把同一个光端点分别塞进两个市场。

第八代不再只是一把更大的刀

2026年4月，Google公布TPU 8t与TPU 8i：8t偏重训练，采用扩大的三维环面互连，单个superpod为9,600颗芯片；8i偏重推理与后训练，采用名为Boardfly的网络组织方式。[3] 三维环面可以想象为三维格子在边界处首尾相连，每颗芯片通过邻居转发到更远处；它不等于每颗芯片与其他所有芯片都有一根直连线。

这里真正改变需求模型的是分化：训练系统和推理系统不再必须共享同一个“每芯片光模块数”。推理侧增加片上存储、缩短通信路径，可能减少某些跨节点搬运；更高并发、更大模型，又可能增加总带宽。哪个力量更大，要看工作负载与部署比例，季报收入本身没有提供这个答案。

Google还公布面向TPU和NVIDIA系统的Virgo网络。[4] 对光产业的启示是：芯片品牌占比变化和网络代际变化要分开建模。Google买入更多GPU，并不意味着它停止做自己的网络；更多自研芯片，也不代表其他供应商的光学能力消失。

OCS增加什么，减少什么

一条光路从发送端进入OCS，再出来到接收端，中间如果没有光电转换，就不能为了OCS的每个插口都新增一只收发模块。OCS端口是光连接端口，不天然等于光电转换端点。它可能减少某些中间电交换层及其模块，同时保留两头光学器件，并提高光纤、精密连接、镜阵列和校准系统的需求。

另外，经过更多连接器、镜面和光路之后，接收光变弱，系统就要重新算光功率预算，即从发出去到能够可靠读到，中间容许损失多少。更高功率光源、低损耗连接器可能受益，但不能仅凭OCS台数给任何一家激光器厂机械增加收入。

OCS卖的是“怎样接路”，模块卖的是“怎样把电变成光”。一路可以经过很多道岔，却只在始发站和终点站装卸货。

因此本书不会给OCS设一个凭空精确的全球渗透率。我们会把它作为影响电交换层数、光端点数与器件结构的变量，等待具体部署和采购数据来约束。这比把TPU芯片数乘一个网上流传的倍数，更能解释下一代结构变化。

资料与核验

[1] Jouppi等，TPU v4论文，2023；[2] Google Ironwood技术说明；[3] Google第八代TPU深度介绍，2026-04-22；[4] Next 2026基础设施公告。产品规格不是出货量，本文结构影响为推论。

第六章 · 两份最新季报，能推导多少光

季报像一张商场收款单：告诉你今天卖了多少钱，却未必告诉你多少人买了哪种鞋码。把收款额直接换成鞋的双数，总要偷偷塞进一个平均单价。由AI收入推光模块出货，同样容易在看似简单的一步里藏进许多假设。

先把最新一季放到同一条日历

截至2026年9月9日，NVIDIA最新已公布季度是截至7月26日的2027财年第二季度，公告日期8月26日；Alphabet最新季度是截至6月30日的2026年第二季度，公告日期7月22日。财年2027不是日历2027，也不能将两家不同结束日期的季度视为完全同步。

公司与口径	最新披露	能说明什么
NVIDIA总收入	962.21亿美元	系统销售规模，不等于GPU颗数
NVIDIA数据中心收入	890.23亿美元	强需求背景，包含的业务范围不等于光模块市场
Alphabet总收入	1,197.96亿美元	大部分仍来自多元业务，不能直接换成算力投入
Google Cloud收入	247.68亿美元，同比增长82%	云需求与变现增强，不是TPU出货收入
Alphabet购置物业设备支出	本季449.24亿美元；上半年805.98亿美元	实际资本支出现金口径，含服务器之外资产

数据来源分别为两家公告和申报文件。[1][2] 表内美元单位统一为“亿”，不与英文billion混用。资本开支，CapEx是购买未来多年使用的设备和设施，不是本期所有经营费用，也不是全部用于光模块。

NVIDIA：产品在换，分母也在换

NVIDIA最新公告称Vera Rubin进入全面生产爬坡，并说明Spectrum-6交换系统支持可插拔和CPO方案。[1] 这更支持两条路径并行，而不是宣布可插拔需求终止。NVL72指以72颗

GPU组织的机柜级系统这一产品形态；它的内部连接与跨柜连接必须分开看，不能把72颗卡上的所有高速接口都当成外购光模块。

从较旧平台向Blackwell、Blackwell Ultra，再向Rubin迁移，可能同时改变每柜计算能力、网卡速率、对外端口数和内部互连。若每柜价格上升30%，系统收入上升50%，在产品组合不变等强假设下，柜数增幅约为 $1.5 \div 1.3 - 1 = 15.4\%$ ，而不是50%。实际产品组合还在变，因此这个式子只能示范为何不能用收入同比直接代替数量同比。

更稳的传导是：分代系统数量，乘每代对外端口数，再按速率和可插拔比例分组。本文核验的最新公告没有给出一张完整的Blackwell、Ultra与Rubin季度颗数占比表，也没有逐客户光学物料清单。因此下文不会把自设的代际比例标成NVIDIA指引。

Google：自研与采购同时增长

Google的公开产品路线已同时包含Ironwood、第八代TPU以及基于Vera Rubin的A5X实例。[3] 它既是自研芯片系统的设计者，也是GPU系统的客户。研究需求时，Alphabet资本开支与NVIDIA收入可能代表同一台设备交易的两面，不能相加后再计算一次光需求。

财报能证明资本支出强度，产品说明能证明架构方向；两者都没有单独给出“今年买多少只800G”的答案。机房土建、电力、服务器、存储、网络在资本支出中的比例，以及付款与设备到货的时间差，都会改变当季模块订单。

另外，Google新一代TPU区分训练与推理架构。若推理系统占比提升，模型不能继续把Ironwood的端点配置全部沿用。若更高计算密度让相同任务少买芯片，每芯片带宽提高也未必使模块总只数同比例上升。这正是结构性占比比一句“资本开支大增”更重要的原因。

把缺失项放在明处

可以用以下式子： $\text{模块需求} = \sum \text{各平台部署量} \times \text{每平台光端点数} \times \text{可插拔比例} \div \text{每只模块承载端点数} + \text{备件与净补库}$ 。这里端点已经包括链路两头，不再乘二。不同速率必须各算一行，CPO光引擎另记一行。

做一个纯结构算例：旧平台每台贡献4个已计双端的可插拔等效端点，新平台贡献6个；新平台台数占比从20%升到60%，平均端点从4.4增到5.2，增长18.2%。若部署台数同时增

30%，则模块端点需求约增53.6%。反过来，若新平台只有3个端点，结构迁移会抵消一部分台数增长。四个系数均是假设，实际应由系统配置与订单替换。

大白话

最新季报增强了“有人愿意继续建”的证据，产品迭代告诉我们“建成什么样”。只有配置、数量和采购时间接上，才成为光模块需求。

本书对2026—2028的方向判断是：高速率升级、低功耗设计与更复杂光连接的机会扩大；对具体路线的影响分化。DSP可插拔看新增端口与速率，LPO看系统协同认证，CPO看部署比例和光引擎交付，OCS看拓扑采用。下一章从供给端检验：宣布扩张的工厂，究竟能交出哪些产品？

资料与核验

[1] NVIDIA FY2027 Q2公告，2026-08-26及10-Q；[2] Alphabet 2026 Q2公告与现金流表；[3] Google Next 2026产品公告。结构算例为作者模型，不是公司预测。

第七章 · 龙头的产能不是一个总数

工厂说“明年产能翻倍”，听起来像粮仓会多一倍粮食。但如果新机器年底才到，部分芯片没到，客户又只批准其中两款产品，那一年卖出的货可能只多一点。光模块供给最难的地方，不是找一个数字，而是弄清那个数字站在哪一道门后面。

五道门，不可跳级

规划产能是准备建成后的能力；设备安装后形成可用制造能力；产品工艺跑稳后，才有较可靠的合格产出；通过客户验证，才成为该客户能买的供给；拿到订单并交付，才成为销量。认证可能针对某一个型号、某一个工厂、某一种芯片方案，不自动覆盖所有产品。

有效供给是规定期间内，能按客户所需规格真正交付的合格产品。若公告给的是尚未考虑良率的设备能力，可以再乘良率；若给的已经是合格成品能力，就不能重复扣一次良率。任何表格先写清分母，再谈紧缺。

主要厂商的证据账本

厂商	截至核验日能确认的事实或计划	时间与限制
中际旭创， 300308	2026半年报称800G、1.6T出货增长、占比提高，并继续扩充全球制造布局	已发生的交付方向明确；本书未取得完整的2026—2028分速率合格只数计划
新易盛， 300502	半年报光互联产品表列产能2,836万只、上半年产量1,091万只、销量1,119万只；涵盖混合产品	不能把总量全部算800G/1.6T；产能需保留原表口径，不自行当作半年可交货量
AOI，AAOI	2026 Q2公告称制造能力接近20万只/月，预计年底800G与1.6T合计约65万只/月	年底目标为公司预期；不能乘12当作2026销量
Coherent， COHR	2026年5月材料提出年底内部InP产出翻倍，2027年再超过翻倍；6月Sherman扩建公告提出晶圆能力扩大	光芯片与晶圆口径，不能与模块只数相加；不得把两个重叠扩产表述重复计算
Lumentum， LITE	FY2026 Q4公告提到OCS与云模块业务、1.6T采用成为增长来源	收入与产品推进证据；未披露可直接填入本书的逐年模块合格产能表

来源见章末[1]—[5]。表格故意不把不同单位装成一个“总供给”：中际旭创讲出货方向，新易盛给混合只数，AOI给月末运行能力，Coherent还包括上游InP。把它们简单相加会同时发生单位错误、产品错误和重复统计。

AOI说明了为什么年底不是全年

假设一家工厂在7月初达到每月20万只，到12月底线性提高至65万只，半年平均每月42.5万只，六个月对应255万只。这只是数学示范，未必等于AOI实际爬坡。若再假设有效利用系数80%，则合格且有订单可交的量是204万只。年末65万只乘12得到780万只，代表该运行速度维持一整年的能力，不是当年已经卖掉780万只。

AOI还曾公告取得某大型云客户的首笔1.6T批量订单，预计在产品认证后于2026年第三季度初开始交付。[6]这句话同时包含订单和条件。不能删去“认证后”，也不能凭“大型云客户”自动填入一个著名公司的名字。最新季度更新优先于早前较低的年底产能目标。

混合产能能空着，高端仍可能紧

新易盛原表的产量、销量与产能看似能算利用率，但产能统计周期、产品组合、年底能力还是期间平均能力若未明确，简单相除就可能失真。即使统一了周期，总产能还有高速与低速之分。能做100G的设备有空余，不意味着能补1.6T客户认证过的缺口。

中际旭创和新易盛的量产经验，重要之处在于能把设计、材料、工艺和客户要求同时组织起来。Coherent与Lumentum则提供另一种供给保障：若关键光源能内部制造或提前锁定，模块放量不完全依赖外购紧缺芯片。不过一体化也要承担研发、资本投入与工艺失败风险，不能只算好处。

供给最短的板，可能在测试台

以教学数字说明：每只产品最终测试若占用设备20分钟，一台设备每天连续24小时、每月30天，理论只能测试2,160只。加班不会让已经24小时运转的测试机多出时间；减少测试时长或增加并行测试位，才会提高上限。随后还要考虑维护、换线、复测和高低温筛选。

这也是为什么“厂房面积翻倍”比“指定型号的合格交付翻倍”弱得多。真正可跟踪的是设备到货、单位测试时间、良率、瓶颈芯片交期与客户放行。它们任何一个没跟上，新增面积只会

先变成折旧和现金流压力。

大白话

供给是一个串联系统：最慢的那道工序决定能交多少。看产能规划，先找它没有说出口的条件。

下一章转向二线和新进入者。它们可能带来最大的新增供给弹性，也最容易被把“送样”写成“全面认证”的叙事高估。

资料与核验

[1] 中际旭创2026半年报；[2] 新易盛2026半年报，第12—13页；[3] AOI 2026 Q2申报公告；[4] Coherent 2026-05-06投资者材料及2026-06-16扩建公告；[5] Lumentum FY2026 Q4公告；[6] AOI首笔1.6T批量订单公告。

第八章 · 二线厂商跨过哪道门

新供应商进入，最先增加的往往是样品和新闻，最后增加的才是稳定交货。市场会把中间几年压缩成一句“打入北美”。我们把这句话重新展开，看看用户点名的七家厂商分别走到哪里。

读表规则

以下A为公司报告、公告或公司正式交流记录；B为媒体采访报道，需要继续向公司披露交叉验证；U为本次核验未取得可确认数据。U不等于零，也不等于没有能力。客户认证只记录公开确认的范围，不猜保密客户名字。“二线”是研究分组，不表示所有公司的规模或技术水平相同。

公司与经营主体	产能与放量时间线	认证及证据边界
剑桥科技， 603083	A：2026中报披露嘉善提升800G产能；公司交流预计下半年1.6T两类核心产品大批发货	A：披露部分800G产品完成海外大客户认证、1.6T相关认证测试完成；并非全部新型号全部客户放行。2027—2028分速率只数U。[1]
东山精密， 002384；索 尔思光电	A：半年报披露自研200G EML进入量产、支持1.6T；B：7月调研报道指2025订购设备2026 Q2调试，2026新订设备2027年初起交付	A：具备既有光芯片与模块业务，不是从空白厂房起步；完整客户名单及逐型号认证比例U。设备交付不等于成品达产。[2]
新菲光，非本 表直接上市主 体	B：2026-07-08采访报道深圳800G月能力10万只、1.6T约4万只；北美1.6T约1万只/月	官网有产品和早期小批量信息；采访所述北美客户进展不扩写为实名认证。2027—2028量化交付计划U；产值目标不是订单。[3]
华工科技， 000988；华 工正源	A：2026半年报披露800G、1.6T规模化交付，并展示更高密度方案	A：有交付进展；不能把展会样品当量产。可比的2026—2028高速合格产能只数及客户逐项验证表U。[4]
联特科技， 301205	A：半年报产能表列283万只、产量165万只、销量160万只，口径含光模	A：产品表含硅光、LPO、薄膜铌酸锂等方案；存在产品不等于某客户已大批采

公司与经营主体	产能与放量时间线	认证及证据边界
	块及受托加工；已布局马来西亚	购。另披露境外400G及以上销量13万只，仍未拆800G/1.6T；分项产能和认证比例U。[5]
汇绿生态，001267；武汉钧恒	A/B：原约300万只/年，鄂州一期新增约150万只/年，2025年底投产；马来西亚2026年4月底逐步投产，预计新增100万—150万只/年；二期规划300万只/年，预计2027年底投产	A：7月回复确认高意订单相对稳定，但未明确确认1.6T量产。产能含混合产品，不能全算1.6T；各基地与型号认证比例U。[6]
恩达通；嘉元科技688388相关投资	A：嘉元2025年投资披露识别标的为光通信企业，正式交流提及100G、400G、800G和1.6T产品	本次未核得可比较的2026—2028产能只数、客户实名认证与年度交付承诺，记U；嘉元是股权暴露，不能把标的全部营收当其主营收入。[7]

三处最容易算错的地方

- 第一，买下旧工厂不等于全球新增同样产能。东山通过索尔思获得的能力，在收购前就可能已计入行业供给。只有收购后新增设备、改造效率或新增班次带来的净增加，才进入全球增量。上市公司业务突然变大，与全球总供给突然变大，是两回事。
- 第二，新菲光不是欧菲光。创始人背景或名字相近，不会自动建立上市公司对它全部利润的所有权。均胜电子也有相关投资信息，应按具体持股与会计处理评价，不能将新菲光工厂的产值按一比一装进任何“关联概念股”。[8]
- 第三，钧恒既卖光模块，也卖有源光缆和光引擎。它的客户可能包含其他光模块厂商。如果上游引擎、代工模块和品牌商最终销售都相加，同一个终端需求可能被算两遍甚至三遍。全球供给表应按最终完整产品只数去重，材料与引擎另列配套约束。

把时间线折算成当年贡献

按上述披露，钧恒马来西亚项目的100万—150万只只是年能力目标。即使假设5月至12月立即满负荷，2026对应也仅约66.7万—100万只；真实爬坡更慢，还要按产品组合拆分。鄂

州二期若2027年底才投产，主要完整年度贡献更应放到2028年，不能在2027年初就加满300万只。

新菲光采访数字若全部成立且不重叠，800G年化运行能力是120万只，1.6T合计年化60万只。但这不是经审计的全年销量，也没有扣除利用率、良率与订单限制。本书仅把它列入媒体证据观察项，不把它当作确定的供给下限。

剑桥的“下半年大批量发货”比“推出样品”前进了一步，但仍要看后续季度收入、库存和回款是否一起变化。华工的更高速样品、联特的多材料产品，能说明研发覆盖；只有明确订单和持续交货，才能将技术储备转换为收入预测。

大白话

设备到货是开始，认证是门票，重复订单才是客流。新增产能最重要的日期，是开始稳定交货的日期。

因此目前最可靠的供需结论不是“这些新玩家一共填掉多少缺口”，而是2026下半年至2028存在多批新增供给窗口，但公开证据尚不足以统一成完整全球表。下一章将给出有明确假设的可复算账本，并展示缺口如何随这些窗口改变。

逐项来源

[1] 剑桥2026半年报公告转载与2026-002投资者关系记录；[2] 东山2026半年报，第11页，设备时间为2026-07-22调研报道；[3] 界面新闻2026-07-08工厂采访转载及新菲光官网。

[4] 华工2026半年报；[5] 联特2026半年报公告转载；[6] 汇绿2026-05-20正式交流记录、2026-07-09正式交流记录及财联社6月18日扩产报道；[7] 嘉元投资公告及嘉元投资者交流记录；[8] 均胜电子2026-06-01投资者交流记录。核验截至2026-09-09，来源较旧者不冒充当日最新产能。

第九章 · 给2026—2028做一张供需表

现在可以问：到底缺多少？负责任的答案分两层。公开资料支持高速产品需求继续增长，也显示多家厂商扩产；但完整、同口径的全球分速率有效供给没有公开，因此**本次研究不能识别全球真实缺口的唯一数值**。仍然可以建一张可复算的情景表，说明什么条件会短缺、什么条件会转为过剩。

先钉住外部锚点

LightCounting在2026年公开的DSP行业简报中预计：2026年800G出货将增长一倍以上，1.6T将从小基数进入数千万端口量级。[1] 这是研究机构预测，不是全年已经发生的销量。其完整数据库还分距离、形态、价格和客户市场，本书没有获得付费明细，不能声称下面表格来自该数据库。

本章只覆盖800G、1.6T和3.2T可插拔等效单端产品；每只记一个完整收发模块，链路两端分别计数。不含400G、AOC、CPO光引擎和OCS设备；这些不能悄悄混进只数。800G带宽当量则把1.6T记为2、3.2T记为4，用来比较运力，不能替代实物物料清单。

2026—2028中心情景：不是确定预测

下面所有精确数字都是作者设定的压力测试输入。2026量级只受前述公开“数千万端口”方向约束；后两年增长和供给量没有足够公开数据校准。它们用于展示结构迁移，不是投资者可以直接当成行业共识使用的估计。单位均为万只。

年份	需求：800G / 1.6T / 3.2T	1.6T实物占比	有效供给假设：同序	分产品供给减需求
2026全年情景	4,000 / 2,000 / 0	33.3%	4,200 / 1,800 / 0	+200 / -200 / 0
2027情景	4,500 / 4,000 / 0	47.1%	5,200 / 3,800 / 0	+700 / -200 / 0
2028情景	4,000 / 6,500 / 500	59.1%	6,000 / 7,000 / 350	+2,000 / +500 / -150

中心情景下，1.6T实物需求由2,000增至4,000再到6,500，年度增幅100%、62.5%；800G则先增12.5%再降11.1%。全表实物需求6,000、8,500、11,000万只，对应800G当量8,000、12,500、19,000万。数量、带宽、收入将出现三种不同增速。

特别看2027：总供给9,000万只大于总需求8,500万只，但1.6T仍短缺200万只。旧产品多700万只，不能直接填补新产品少200万只。2028总量更宽松，也仍可出现3.2T早期短缺。这是结构性缺口，不是所有光模块都供不应求。

范围比中心数字更有用

固定中心供给，令每种需求上下浮动20%，只看1.6T，2026供给减需求范围为-600至+200万只；2027为-1,000至+600万只；2028为-800至+1,800万只。负数代表潜在缺货，正数代表潜在多余能力。同一套产能规划，足以同时容纳“紧缺”和“过剩”的故事。

而供应也会变。若2027年中心情景1.6T有效供给因认证延期再少10%，从3,800变成3,420万只，缺口扩大到580万只。若需求同时下降20%至3,200万只，则又变成多220万只。**在关键输入误差远大于缺口时，发布一个精确缺口百分比并没有意义。**

把平台结构接回模型

可为NVIDIA旧平台、Blackwell系列与Rubin分别列部署量与端点参数；Google则拆为Ironwood、TPU 8t、TPU 8i及GPU系统。下面只做Google类平台的归一化示范，不是它的真实采购份额：2026—2028旧代/训练新代/推理新代的台数占比分别设为60/25/15、30/40/30、10/45/45；每个标准化部署单元的光端点假设分别为4、6、3，三年平均便为4.35、4.50、4.45。

这个例子解释为何“更新一代”未必使每单元模块越来越多：训练侧增加带宽，推理侧结构优化，两股力量抵消。标准化单元总数仍需另乘增长率。实际取得平台配置后，应替换端点参数，不能把这些数字照搬为每颗TPU模块数。

真正有证据的放量窗口

窗口	公开事件或计划	模型处理
2026上半年	头部1.6T交付增长；钧恒马来西亚4月起投产	保留已发生与爬坡区别
2026下半年	Rubin相关系统推进；剑桥预期1.6T大批发货；AOI年底月能力目标	按月份累计，不乘12冒充全年
2027	东山设备后续交付报道；Coherent进一步扩InP计划；钧恒二期年底投产目标	新增能力须扣除重叠、分清芯片与模块
2028	若前期计划兑现，可有更多完整年度贡献；3.2T及相干1.6T需看客户采用	前半句是条件推论；后一项为机构路线展望，非交付承诺

本章没有将七家二线厂商的不完整总量硬凑成全球供给；它们应逐项进入底层表，与品牌商及代工关系去重，再更新中心供给。实际成交量通常是可交供给与采购需求的较小者，还受库存释放影响；有缺口不代表需求全部变成当年出货。

资料与可复算范围

[1] LightCounting 2026 DSP公开简报；时间线的公司依据见第六至八章；远距离路线参见LightCounting 2026市场展望。本章中心表、±20%区间与平台占比全部为作者情景，不是这些来源披露的精确预测。

第十章 · 价格下降，为什么利润还会涨

一只模块去年卖1,000元，今年卖800元，企业就一定更差吗？若制造成本从700元降到500元，单只毛利仍是300元；销量翻倍，毛利总额也翻倍。价格、成本、产品组合与销量，必须放在同一张纸上。

先辨认看到的是哪一种价格

ASP是平均售价，通常以某类产品收入除以销量得到。它可能同时受到规格升级、客户组合和促销影响，不等于同一只产品的年度降价幅度。合同价是买卖双方在数量、质保、付款和交付条件下谈定的价格；网上可买到的少量零售价，还可能含渠道服务、兼容性支持和库存成本。

公开样本，人民币/只	页面标价	用途与限制
FS 800G DR8，500米，双MPO-12	9,599元	同页面解决方案样本；不是头部云厂商采购价
FS 1.6T DR8，500米，双MPO-12	12,828元	与前项可作同一公开页面的粗比较，散热与兼容条件仍需核对
FS 1.6T 2FR4，2公里	16,038元	距离与光学结构不同，不能只按速率比较

价格为2026-09-09检索该销售商解决方案页所见，不保证下单仍有效，税运费及批量折扣以实际报价为准。[1] 同站其他页面的1.6T标价也有不同，说明SKU与渠道页面选择本身会影响样本。本文不把它们平均成“行业真实价格”。

给未来价格一个可修改的输入

本次未取得可核实的主流云客户2026—2028合同报价清单。以下美元价格是作者用于演算的假设，不是市场报价或机构预测；3.2T前两年空白表示不在此模型中计入，并不代表市场绝无样品或收入。

年份	800G假设ASP	1.6T假设ASP	3.2T假设ASP
2026	400美元	900美元	未计

年份 800G假设ASP 1.6T假设ASP 3.2T假设ASP

2027 320美元	650美元	未计
2028 260美元	480美元	1,200美元

若沿用第九章中心供需、无库存释放、出货按每种供需较小者计算，对应模型销售额为322亿、391亿、458亿美元。它们不是全球市场规模预测，只是前述未经完整校准的量价输入相乘。这个算例揭示：单品年年降价，整体收入仍可增长，因为新速率数量与占比上升。反过来，若需求减弱、降价加快，两股力量也会同时伤害收入。

对2027年的1.6T假设而言，3,800万只出货乘650美元为247亿美元。若价格再降10%，同样出货便少24.7亿美元收入。供应链研究里，多猜准几个月份的“产能”却忽略价格10%的波动，可能得出更不准确的盈利结论。

技术路线怎样改写物料数量

先固定一个教学样本：100万只1.6T模块；假设分立方案每只8颗EML，硅光方案每只2颗CW激光器。这里的2颗是演算假设，实际有不同光源数量、功率与共享方式。硅光份额由40%升至60%时，EML需求由480万颗降至320万颗，CW由80万颗升至120万颗。总模块不变，上游产品结构已经大变。

若模块总数再翻倍，EML变成640万颗，仍高于最初480万颗。于是“硅光替代EML份额”与“EML绝对需求增长”可以同时成立。这就是判断中美材料、芯片厂时，应同时算总量和份额的原因。具体型号的物料表比泛泛的路线渗透率更有用。

同理，固定100万只模块，假设全DSP方案占比从90%降到70%，传统完整DSP模块数量由90万降到70万。但若总体翻倍，则变成140万，高于起点。LPO减少单端数字处理，并不自动使所有DSP供应商收入下降，它们还可能卖更高速DSP、驱动器与TIA。

CPO则改变产品形态：被替换的那一端少一个可插拔盒子，但光引擎、光源、精密耦合仍需采购。MPO与其他高密连接器需求要看每链路纤芯数；若单通道速率翻倍、通道数减少，带宽上升也未必让纤芯需求同步增长。OCS还可能减少中间转换端点，却提高光路连接的复杂度。

毛利率之后，还有现金

毛利率是收入减直接营业成本后占收入的比例；它还没有减去全部研发、销售、管理和融资费用。扩产中，新设备折旧先来，订单和回款后到，毛利率上升也可能伴随现金紧张。联特2026半年报就同时呈现收入增长与利润受投入、汇兑等因素影响，说明“高速产品增加”不足以独立解释股东收益。[2]

大白话

价格不是利润，利润也不是现金。先把每只产品的收入与成本算清，再看扩产到底占住多少资金。

因此受益排序不能只看哪家技术更新。还要问：省下的成本由供应商保留多少？客户是否压价？产线折旧是否被产量覆盖？下一章用这些问题把上市公司放到具体环节，避免一张“光概念”名单把完全不同的生意混在一起。

资料与核验

[1] FS InfiniBand光模块解决方案价格与规格表，访问于2026-09-09；[2] 联特2026半年报。三年美元ASP、市场销售额与芯片数量均为教学情景，未以零售价冒充批量合同价。

第十一章 · 中美上市公司的技术地图

设想把一家光模块工厂搬到桌上：左边是晶圆和激光器，中间是芯片、装配线，右边是连接器与测试台。桌子后面站着决定网络结构的云厂商。所谓“光通信股票”，其实分散在这几个完全不同的位置。

本章按**公司出售什么、收入怎样发生**分类。一家公司可以跨行出现。美股上市也不等于美国制造：Fabrinet的重要生产基地在泰国，AXT的生产经营还涉及中国。下面是截至2026年9月9日的研究地图，不是已确认向NVIDIA或Google供货的名单。

先看整机与模块：谁拿到客户的订单

位置	上市公司／代码	对应路线与核验重点
平台与交换芯片	NVIDIA NVDA；Alphabet GOOGL／GOOG；Broadcom AVGO	前两家决定GPU、TPU与网络结构；Broadcom同时有交换芯片、DSP和CPO。平台增长不等于同一供应商份额增长。
高速模块主线	中际旭创300308；新易盛300502	800G、1.6T及硅光，见第七章。应跟踪高速产品收入、良率及现金回款，不能只看总产能。
模块与器件布局	光迅科技002281；华工科技000988	光迅有器件到模块产品；华工通过华工正源开展光业务。整家公司收入还包含其他业务。
扩张与客户导入	剑桥科技603083；联特科技301205；东山精密002384；汇绿生态001267	分别追踪高速新品、海外基地、索尔思、钧恒。第八章逐项区分认证与交付，不能把集团市值直接映射成光模块产能。
美国上市模块／器件商	Coherent COHR；Lumentum LITE；Applied Optoelectronics AAOI	前两家跨激光器、器件及模块，Lumentum另有OCS；AOI跟踪800G／1.6T爬坡。计划产能与合格出货要分开。
精密制造服务	Fabrinet FN	为设备商做光学封装与制造。不能把它代工的模块再与品牌商模块相加，否则同一只盒子算了两次。
间接股权联系	嘉元科技688388；均胜电子600699	分别结合恩达通、新菲光相关投资披露研究。参股价值、合并收入和实际持股比例是三本账，不能把

位置	上市公司／代码	对应路线与核验重点
被投企业全部利润算进上市公司。		

这张表的模块依据见第七章与第八章。补充核验包括光迅产品资料、Fabrinet 2026财年业绩、Broadcom CPO发布及均胜投资者交流材料。股权联系并不替代客户认证证据。

再看材料与芯片：谁控制光的源头

位置	上市公司／代码	产品对应与阶段边界
电信号处理	Marvell MRVL；Credo CRDO；Broadcom AVGO	高速DSP及相关电芯片。LPO减少模块DSP需求，但升级速率又提高电接口难度；必须看各自产品组合。
InP衬底	AXT AXTI	磷化铟等衬底，上游材料不是可直接发光的成品激光芯片。关注材料品质、扩产兑现和跨境交付。
激光芯片	源杰科技688498；仕佳光子688313；Coherent COHR；Lumentum LITE	源杰有CW产品；仕佳布局DFB、CW与EML等，具体产品处于不同阶段。后两家覆盖InP激光器。不能把研发中的EML当成已认证大批量供货。
铌酸锂与调制器	天通股份600330；光库科技300620	天通涉及晶体、晶片和薄膜产线；光库涉及调制芯片与器件。相干通信产品出货，不自动证明AI短距PAM4产品量产。
光引擎与精密器件	天孚通信300394；仕佳光子688313	天孚有光引擎配套、FAU等；仕佳有AWG等器件。 <u>FAU</u> 是把多根光纤精确排列的接口， <u>AWG</u> 则按波长分路，像给不同颜色的光安排不同出口。

硅光路线不能简单列一排“硅光公司”：设计、晶圆制造、光源、封装和模块可能由不同企业分担。Marvell已在2026年2月完成收购Celestial AI，不能再把后者当成独立上市标的；Credo的硅光并购也应按交易阶段核对，而不是凭旧标签把它永远归为铜连接公司。

核验入口：Marvell收购完成文件；Credo光DSP及硅光并购公告；AXT材料业务说明；源杰年报摘要；仕佳年报；天通交流记录；光库年报；天孚产品资料。

最后看连接、远距传输与测试

位置	上市公司／代码	为什么要单列
多芯连接与布线	太辰光300570；长芯博创300548；Corning GLW	前两家涉及MPO等连接产品，业务范围各有不同；Corning覆盖光纤与连接。高密度布线有机会，但单价与模块完全不是同一量级。
模块电路板	博敏电子603936	披露高速模块PCB业务。PCB是承载电路的板，不能因为装在1.6T模块里就享有整个模块售价。
远距与相干传输	Ciena CIEN；Cisco CSCO；德科立688205	Ciena的WaveLogic、Cisco的Acacia，以及德科立的长距模块、放大器等。跨数据中心投资是另一条需求线，不能套用机房内DR8数量。
测试与量产验证	Keysight KEYS；VIAVI VIAV	检验电信号、光信号与网络性能。速率升级会增加测试难度；收入还受设备重复利用、测试效率和其他业务影响。

这些位置的资料入口是太辰光年报、长芯博创披露的行业与业务材料、Corning连接趋势、博敏半年报、Ciena产品、Cisco业绩材料、Keysight测试方案及VIAVI产品披露。

大白话

名单的用处，是让你沿着收入往回找到产品。模块涨十元，不代表衬底厂也涨十元；一个订单从品牌商流到代工厂，也没有变成两个订单。

这是一张覆盖主要环节的观察地图，不是全球所有上市企业的穷尽清单。长途光缆、通信设备和晶圆代工还有大量参与者；若没有可核验的AI光业务关系，本书不凭概念补名单。下一章再从这张地图中选出最值得持续验证的对象。

第十二章 · 谁最有潜力，怎样推翻这个判断

如果只能带走一家公司名字，这本书很容易写：挑增长最快的，再给它配一个足够大的未来。但真正有用的答案应该能接受下一份季报的反驳。我们要选的是一条可以验证的利润路径。

本章的“潜力”指**技术需求转成持续业务的可能性**，不指从今天股价出发的预期收益。本书没有建立统一的当前估值、净债务和稀释后每股盈利模型，因此不能据此排列股票回报。一个优秀供应商，如果价格已经预支了三年成功，也可能不是有吸引力的投资。

先给出我的判断

以2026—2028年的业务兑现为目标，我优先研究已经有高速量产基础、能够跨越800G到1.6T、同时保留下一代封装能力的企业。在中国模块公司中，**中际旭创是我的首要研究基准，新易盛是必须并列比较的核心对象**。这是基于产品位置与现有交付基础的产业判断，不是两家未来股价的排名。二者都要面对客户集中、降价、扩产和路线变化，规模本身不构成永久优势。

如果问题改为“谁能从多条光路线中获得收入”，美国上市公司中我优先比较Coherent与Lumentum：前者值得看InP与器件到模块的纵向布局，后者值得看激光器、模块和OCS的组合。如果问题是“谁更能决定CPO怎样实现”，Broadcom比单纯贴着CPO概念的公司更值得先读。依据分别在第七章的公司披露及Broadcom的CPO产品公告。

研究方向	优先观察对象	成立条件	最有力的反证
1.6T可插拔持续放量	中际旭创、新易盛	高速收入与良率同步提升，降价没有吞掉全部成本改善	库存、应收增长持续快于收入，或客户订单份额下滑
第二供应商导入、扩产弹性	AOI、剑桥科技；再追踪华工、东山	产品认证变成连续批量交付，新基地按期提供合格产品	月产能目标多次延期，或收入增长但经营现金持续恶化
硅光与CPO带动光源、	Coherent、Lumentum；天孚通信、	高功率光源与精密连接的价值随新架构增加，并拿	新增部件价值被集成压缩，或合格供应商份额没有提升

研究方向	优先观察对象	成立条件	最有力的反证
封装	源杰科技	到相应订单	
交换芯片与电接口升级	Broadcom、Marvell、Credo	新增速率、交换容量和产品组合抵消部分DSP被省去的影响	LPO／CPO转移的价值多于新产品补回的收入
OCS扩展	Lumentum；新易盛作为产品进展观察对象	出现可核验的系统订单、验收与后续扩单	展示长期停留在样机，未形成可辨认业务；不默认已成为Google供应商
薄膜铌酸锂加速采用	光库科技、天通股份	从材料、相干器件进一步走到目标应用认证与稳定收入	AI短距产品仍在验证，只有概念与远期产线计划
高密度布线与测试	太辰光、Corning；Keysight、VIAVI	连接密度、测试复杂度确实增加单系统采购价值	采购周期延后、设备复用或效率提升压低增量采购

表中是作者推断；产品对应可回查第十一章的官方资料入口。其中OCS产品发布与OCS规模收入尤其要分开。所谓“跨路线受益”，也必须逐项数清售出的部件，不能把同一笔模块收入换三个技术名字再加三次。

有些机会很大，但证据还很小

联特、汇绿及新菲光等值得跟踪，不等于它们应当与龙头采用同一确定性。新业务基数小，增长百分比容易很高；认证、原料和良率中的任何一环，都可能让规划与收入相隔数个季度。恩达通的公开分速率产能与客户认证证据不足，我把它留在待核验栏，而不是补一个想象中的2027年出货量。

天通的案例尤其具体。公司在2026年5月的交流记录中，对市场流传的某项Lumentum送样认证说法作了否定。原始交流记录说明了一件事：材料能力、送样、通过认证、正式采购是四件事。把这四个台阶画成一条没有停顿的直线，模型就会系统性高估早期公司。

评价高弹性公司，我更关心“失败之后还剩什么”：旧业务能否支持现金开支？新厂能否切换产品？关键技术是否属于自己？如果必须每一件好事同时发生，才撑得起盈利假设，这个故事对时间误差就极其敏感。

下一份季报，只更新六格

1. 平台：Google与NVIDIA是否改变设备交付、网络产品及资本开支的实际节奏？只更新已披露事实，不把总收入直接换成GPU数量。
2. 结构：800G、1.6T、3.2T的实际销售占比有没有新证据？先区分台数占比、端口占比与收入占比。
3. 架构：更多流量走可插拔、LPO、CPO还是OCS？改变的是模块数量、DSP数量，还是光源与封装的位置？
4. 供给：厂房落成、设备进场、客户认证、批量验收各到哪一步？年底目标只进入年底月度能力，不倒填全年。
5. 价格：是不是同速率、同距离、同封装、同交付条件的价格？不要用公开零售价代替大客户合同价。
6. 现金：增长是否最终进入经营现金流？库存与应收是否提示提前备货、验收拖延或价格压力？

完成这六格后，再重算第九章。举例说，若一条计划7月量产的线延迟到10月，原先假设下半年的六个月贡献就不能保留；还要重新安排爬坡曲线。若客户从八颗EML方案改为硅光加外部CW光源，变化的不只是模块份额，而是整张材料清单。模型最该有的功能，是允许事实把旧判断改掉。

大白话

有潜力的企业，要站在需求确实经过的位置上，还要有能力按时把货交出去。技术名称只告诉你可能经过哪条路，收入与现金才告诉你车有没有到站。

回到开篇：AI增加的不只是计算，还增加计算之间的等待。光解决其中一部分等待，却要付出光源、调制、连接、交换和测试的成本。顺着这笔成本找到供应商，再顺着认证与交付找到收入，我们才有理由讨论谁会受益。最值得保留的不是某个永远正确的名字，而是一份能够逐季对账、允许自己被推翻的判断。