

观点与分析

编辑日期: 2026年11月1日

那扇无法关上的门，代价是多少?

> 最大损失未必来自一台买来后坏掉的机器，也可能来自一台买得太早的机器：为了让当初的购买继续显得合理，我们不得不围绕它建造整个未来。

工厂的工期成了技术决策

2025 年 2 月 27 日，英特尔向俄亥俄州提交两座尚未完工工厂的报告。地下公用设施、冷却与超纯水设施、电力基础和部分混凝土工程已开始。截至 2024 年 12 月底，项目现场已支出约 37 亿美元，另有 32 亿美元形成在建合同承诺，即已发生支出与现有承诺合计约 69 亿美元。公司告知州政府，第一座工厂按修订计划到 2030 年完成，预计 2030 – 2031 年投产；第二座 2031 年完成，准备 2032 年投产。未来的一部分已进入土地，启用日期仍可调整。2

英特尔 2025 财年年报把问题表述得比“有钱就能建完”更复杂：下一代制造工艺能否找到足够客户，使高度专用的设备与厂房投资合理？公司称，若未获得大型外部客户，Intel 14A 节点的开发可能放缓或停止；部分未来产品也可交给外部晶圆代工厂设计生产。这是公司披露的战略选择，不证明它已经决定停止 14A。1

同一报告称，俄亥俄先进制造设施建设放缓，其他国家的建设或扩建也被复核。固定成本高企，使需求不及预期时难以迅速缩减既有产能。公司在 2024 年和 2025 年分别确认 33 亿美元与 9.5 亿美元的制造资产减值和部分加速折旧。这些金额并非都属于俄亥俄或 14A，而是在公司层面说明：专用资产经济性在会计寿命结束前改变，会发生什么。1

这不是人工智能数据中心案例，而是制造芯片的工厂。晶圆厂不等于 GPU 机架；共同之处是前置巨额投入，设备不能像普通二手服务器一样次日出售。公司文件也区分已支付金额与合同承诺，不能相加后统称已完成现金支出。2

此类工厂的混凝土浇筑在关于客户、制程、竞争、补贴与资本获取的假设上。公司以今天的信息，押注多年后建成时的需求。需求中途变化，问题不是是否继续相信愿景，而是哪些承诺可以推迟，又不摧毁未来成功的可能。

延期在新闻里可能像违约，也可能保护公司不在条件明朗前继续投入；它也可能显示更严重的市场、运营或融资困难。本文不判断英特尔决定最优或延期有利，只提取公司明确表达的一点：保留替代制造方案的选择权，本身就是决策价值的一部分。

一份“百亿美元合同”，并非百亿美元支出

美国国防部 JEDI 云项目虽不同于晶圆厂，却同样难以逆转。2019 年 10 月，微软中标单一供应商合同。100 亿美元是最长十年、所有选项均使用时工作订单的潜在上限，不是授标当日付款或自动支出承诺。基础期两年，工作订单最低保证仅 100 万美元，依据国防部监察长报告。3

当时选择单一供应商并非所有机构都认为明显错误。美国政府问责局审理招标抗议时，认为国防部提出的单供应商理由在法律上允许且有依据，其中包括国家安全考虑。不能在条件改变后，把过去每项集中决策都重写成一开始就愚蠢。4

国防部最终没有启动那个长期合同未来。2021 年 7 月，它宣布取消 JEDI 并着手终止合同，官方理由包括需求、市场和机构云经验变化，使原框架不再适合。诉讼也曾阻碍执行，因此不能把案例缩成单一技术原因；但需求变化确为公开理由，替代方案也揭示了所需改变。5

2022 年 12 月，JWCC 合同授予 AWS、Google、Microsoft 与 Oracle 四家公司。多供应商在工作订单上竞争，基础期三年，之后两个一年延长期，潜在上限 90 亿美元。100 亿与 90 亿的上限差，不是已证明节省 10 亿美元；期限、需求和合同结构都不同。2023 年国防部首席信息官将其描述为允许竞争和资源跟踪的多供应商采购工具，而不是统一运行全部系统的一套云环境。6

这不是“四家永远优于一家”的故事。多供应商会增加运营、身份、安全、数据传输与协调负担。更深的问题是：哪些事项只能锁定两年，哪些应对五年或十年保持开放？旧结构形成时有其理由，条件却在可能期限用尽前改变。重新选择权在签约时不显眼，在昨天不适合的替代方案今天最接近需要时才显出价值。

为什么财务计算器不够？

传统财务分析看净现值（Net Present Value, NPV）：折现后的预期现金流价值。Robert Pindyck 关于不确定性投资的研究指出，许多投资同时具有两个特征：投入后难以收回，以及投入前可以等待新信息。二者结合，使等待具有类似期权的经济价值，即使投资者并未交易金融合约。7

这就是实物期权（Real Options）。可作多种用途的土地、允许而非强迫扩容的合同、先建外壳后选生产线的工厂，都包含选择权。机构支付的不只有资产，也包括保留或放弃在知识变化后调整战略的权利。

设一个完全虚构、使用抽象货币单位的例子。需求强时项目 NPV 为 +150，需求弱时为 −70，概率各半。立即投资的期望值是 40。若先做成本 8 的测试，延期损失估为 12，且在这个理想例子里测试能完全揭示需求，机构只在强需求时投资：$0.5\times150-8-12=55$。差额 15 是这些假设下分阶段投资的价值。

15 不是普遍规则，更不是任何真实公司估值。测试完全揭示信息过于乐观，延期损失固定也很简化。信号弱、竞争者趁等待占领市场，或电力与土地窗口关闭，都可能让差额消失甚至反转。例子只说明：不能只问项目利润，还要问在关键事实显现前封闭选择的代价。

沉没成本（Sunk Costs）也不能成为新增支出的理由。已支付且消耗的钱不会因为继续而回来。“已经花太多所以不能后退”可能是在把旧损失变成新损失。下一道门槛应比较从此刻继续的未来价值与成本，而不是惩罚未来来为过去辩护。

在理解所有医院之前，系统已经买下

英国 NHS 国家信息技术计划从 2002 年起试图大规模更新患者记录并连接医疗机构。英国国家审计署 2011 年报告称，按原愿景普及电子记录的核心目标不会如期实现；详细病历系统延期、范围缩小，成本却没有相应下降。[8]

截至报告时，病历系统已支出 27 亿英镑。审计员根据当时交付结果认定物有所值不足；一个主要合同区七年内，急性医院机构计划的 97 套系统只交付 4 套。战略后来从全面替换改为连接既有系统，卫生部估计互操作性至少再需 2.2 亿英镑。[8]

不能说所有资金都浪费或项目毫无收益。后续审查引用卫生部全生命周期收益估计 107 亿英镑、成本 98 亿英镑，但收益能否实现存在重大不确定性，约三分之二预计在 2012 年 3 月之后才出现。后来的政府评估把 2002 – 2011 年经历称为昂贵且没有实现大多数目标。[9]

项目不缺宏大愿景，缺的是在假设变成合同、集成、工期和万人培训前，准确了解不同系统、文化与优先级的医院能统一到什么程度。退出不只是解约，还可能重做数据接口、重新培训和多年并行两套服务。

JEDI 在潜在上限用尽前更换采购框架；英国计划则在系统、合同与期待积累后才发现愿景与现实的差距。比较不是证明分阶段采购能挽回每一英镑的因果试验，而是在展示可逆程度不同，并追问：哪些信息本可用低于重建的成本更早获得?

芯片比容纳它的建筑更快前进

人工智能基础设施面临双重时间：技术与物理资产不按同一时钟变化。国际能源署 2026 年报告估计，数据中心 2025 年用电约 485 太瓦时，中央情景到 2030 年约 950 太瓦时。这是情景估计，不是已签合同的电力需求；它同时强调电网、设备和高带宽内存等供应瓶颈。[10]

与此同时，芯片代际可能在工厂和输电建设期间变化。NVIDIA 2026 年对 Rubin GPU 架构的厂商说明展示了自身规格和性能主张；它不是独立基准，却说明硬件路线图的节奏会快于建筑资产。11

因此不能把“AI 工厂”当成一个同步折旧的单一资产。土地、电网连接、冷却管线与建筑壳体可能使用多年；计算机架、互联、模型平台和应用则在更短周期更换。若全部同时锁定，快速层可能在慢速层建成前过时；若什么都不提前锁定，慢速资源窗口可能消失。

项目负责人不能只问站点能容纳多少设备，还要问电力、冷却、网络与建筑能否在处理器功率、热密度和互联方式变化时重新配置。未来负载可能在大型训练、实时推理和更小的专用模型之间改变；因此应把慢速资产设计成可容纳多种快速层，而不是把今天最流行的工作负载写进不可逆的混凝土。

当未来用途尚未确定究竟是大模型训练、实时推理、工业负载，还是靠近用户的本地处理时，冷却、电力与网络的灵活性本身就具有独立经济价值。设计原则很简单：让慢速基础设施容纳变化，只在需求理由足够清楚时锁定快速层。但这项原则并不意味着等待永远正确。

免除退出费用，不等于退出成本消失

2024 年 3 月，AWS 宣布符合条件的客户迁出服务时，可以按程序免除向互联网传输数据的费用；指引持续更新到 2026 年 10 月。它消除的是满足条件时的一类传输账单，不保证数据库能在另一云上运行，也不保证权限、密钥、基础设施即代码与恢复计划无需重建。12

英国政府关于云技术锁定的指南比“以任何代价避免锁定”更准确。供应商专有服务可能在运营、安全和更新方面创造高价值；总用功能较低的通用服务替代，可能经济上不合理。指南建议比较价值与可移植性，估算离开时间与成本，并实际测试若干关键组件的迁移。13

因此有三扇门。法律之门由排他或购买承诺关闭；技术之门由接口、数据模型与服务耦合关闭；人的门由团队技能、流程与运营记忆关闭。供应商可以在法律上打开第一扇门，另外两把钥匙仍然不见。扩张前不演练退出，合同会显得比系统实际承受能力更灵活。

能买下更大决策的小测试

有限试验之所以可能比立刻启动大项目更有战略性，唯一正当理由是它真能改变后续昂贵选择。没有现实停止或修改可能、只确认所有人既有信念的试验，是披着知识外衣的额外成本。

例如机构计划为数千用户建立延迟敏感推理服务。测试不能只看演示模型，而应以受保护的真实数据与请求样本，测试峰值负载、上下文长度、语言切换、95 分位延迟、每项完成任务成本而非每百万 token 价格、合理拒答率，以及模型更新后的行为；再把可比组件迁到第二供应商或替代架构，观察成本、时间和质量。“我们可以离开”才成为经验信息。

试验本身也会成为陷阱。英国技术采购指南把缺乏明确合同关口、试点不断延长扩张称为试点蔓延（Pilot Creep）。它要求在扩张资格、有限学习延期或引入新伙伴之间设置断点；试验前确定目标、指标、知识转移和合同期限，并保留理解系统的内部能力。14

知识不是供应商宣布自身成功的报告。必须有一个失败后会真正改变方向的标准，例如：独立团队完成数据和流程迁移、证明兼容并测量折扣结束后的成本前，不签扩张合同；或在不同于营销演示用户的群体中通过测试前不扩张。失败没有采购后果，条件就只是仪式。

这种建议并不陌生。美国 IT 模块化采购规则要求尽可能把大型系统拆成可互操作增量，合同不得绑定机构购买所有后续阶段；英国 AI 采购指南也把发现与概念验证阶段作为全面运行前降低风险的方法。15

试验规模不由发票金额衡量，而由它能纠正的决策价值衡量。无法预测规模化表现的小测试不是真实期权；能消除关键未知或推翻运营假设的聚焦测试，能为机构买下拒绝更大损失的权利。

当等待本身也成为锁定

最强反对意见是：延期试图保留的选择会随时间消失。稀缺电力站点被预订，变压器排队，人才流向竞争者，采购条件不再可得。国际能源署指出，AI 快速扩张正撞上较慢的能源与供应基础设施。站点窗口稀缺时，提前固定某些权利，可能优于等待所有技术细节明朗。10

期权理论也不保证竞争中的延期最优。关于期权行使博弈的研究表明，竞争者先投资的风险会缩小等待价值。期权价值受竞争反应影响，不只是“不确定性程度”的函数。16

设想一个沙特计算项目——只是概念例子，不描述现有项目。它先要取得多年可用的电力与冷却基础设施，之后才采购变化更快的处理器、模型平台和应用。一个错误是等处理器竞争结束再申请站点和并网，届时容量可能已无；另一个错误是还没看见真实生产负载，就把全部机架和多年单一平台服务一次签下。

聪明决策不是口头折中，而是按资产性质设定不同承诺日期。获取周期长于技术变化速度时，先锁定站点权或扩容选项；建设容纳多种配置的连接；不影响服务时推迟快速贬值采购，把需求分成基础预订与有条件扩展，并在销售或合作合同中明确谁承担闲置能力、谁享受缩减权。

管理直觉不能独自判断。要画出“学习时间—替代方案消失时间”曲线：一年后到来的信息，对一个月后就消失的站点没有决策价值；反之，在运营成本每季度变化的市场里签五年不可退出合同，会把今天的小价格优势变成明天的持续负担。

退出能力账

本文提出技术与工业项目的“退出能力账”（Exitability Ledger）。它是未经测试的原创分析建议，不是专业标准或盈利保证。每项难以逆转的决定前，记录六件事：依据的假设；何时获得新信息；此前会沉没多少资金；退出合同、基础设施与流程的成本；等待会失去哪个替代方案；新信息将允许停止、扩张还

是切换。

可先测“承诺—学习缺口”：实质支出变得不可回收的时点，到有足够证据检验支出假设的最早日期之间的间隔。长缺口不总是错误，工厂、道路和网络必须先建后用；但它会逼问：能否把部分不确定性移进期权合同、设计阶段、有条件采购或扩容前客户承诺？

每道关口检查三个不能互相替代的层次：

1. 资产：设备能否出售、复用或改造，真实损失而非仅会计损失是多少？
2. 合同与数据：是否有明确终止、导出、转移许可、取回文件、接口和文档的权利，是否真正演练过？
3. 机构：团队能否运行替代方案，还是供应商即使不拥有代码，也已拥有系统知识？

账本的严厉问题是：如果需求或技术假设被证明错误，我们怎么办？唯一答案若是“说服市场”，就不是备选方案，而是补偿性营销。若有清洁停止关口、已知退出账单和演练迁移路径，机构便买到一项无形资产：不继续走错方向的能力。

测试也可能推翻谨慎本身。若可移植架构在整个生命周期增加巨大成本，变化概率低，专用服务收益高，接受经过评估的技术依赖可以合理。英国指南也承认：绝对独立并非免费，且不总是最佳；关键是依赖成为测量后的选择，而不是合同结束日才发现的事实。13

成熟战略不会永远打开全部门。建设、研究与扩张终须关闭许多门；但必须知道哪扇先关，哪项权利值得购买以多开一会儿，哪扇会因犹豫而永远消失。

仓促项目与雄心项目的差别，不在工程图大小、合同价格或新闻稿长度，而在新信息反驳原计划时：机构能否改变主意，而不因为已经签字就宣布世界错了？

项目最昂贵的可能不是已花的钱，而是这些钱让我们后来无法作出的决定。

参考文献

1. Intel，2025 财年 Form 10-K，美国证券交易委员会，2026。
2. Intel，俄亥俄州站点经理函与本土化激励年度报告，2025 年 2 月 27 日。
3. 美国国防部监察长办公室，JEDI 云采购报告，2020。
4. 美国政府问责局，Oracle JEDI 招标抗议声明，2018。
5. 美国国防部，JEDI 云合同的未来，2021。
6. 美国国防部，JWCC 合同授标，2022；首席信息官办公室，云使用合理化备忘录，2023。
7. Robert S. Pindyck，Irreversibility, Uncertainty, and Investment，NBER 工作论文 3307，1990。
8. 英国国家审计署，The National Programme for IT in the NHS，2011。
9. 英国国家审计署，NPfIT 最终收益声明审查；NHS 数字技术使用，2020。
10. 国际能源署，Key Questions on Energy and AI，2026。
11. NVIDIA，Inside NVIDIA Rubin GPU Architecture，2026；为厂商规格与主张。
12. AWS，迁出 AWS 时免除互联网数据传输费，2024，后续有更新。
13. 英国政府，Managing technical lock-in in the cloud。
14. 英国政府，Digital, Data and Technology Playbook，2023。
15. 美国政府，FAR 39.103: Modular Contracting；英国政府，Guidelines for AI procurement。
16. Strategic Delay in R&D Competition，*Review of Economic Studies*，2002；Option Exercise Games，*Review of Financial Studies*，2002。