SAFWAN ALMODHAYAN



OPINION AND ANALYSIS

Editorial date: November 1, 2026

The Price of a Door We Cannot Close

> The greatest loss is not always a machine we bought and then watched fail. It may be a machine bought too early, forcing us to build the entire future around keeping that purchase rational.

When a Factory's Schedule Became a Technology Decision

On February 27, 2025, Intel sent the State of Ohio a report about two factories, neither complete, whose site was no longer merely a project on a screen. Construction workers had finished underground utility corridors, begun work on cooling plants and ultrapure-water systems, installed part of the electrical infrastructure, and poured concrete for fabrication facilities. By the end of December 2024, about **$3.7 billion had been spent** on the site, with **another $3.2 billion contractually committed** to work in progress: roughly **$6.9 billion across completed expenditure and outstanding commitments**. The company's site manager then told the state that, under the revised plan, the first factory would not be complete until 2030, with operations expected in 2030–2031; the second was scheduled for completion in 2031 and operations in 2032. Without embellishment, the letter shows part of the future already embedded in the ground while the date on which it would operate remained adjustable.2

In its annual disclosure for 2025, Intel did not describe a semiconductor plant as a building that could be completed whenever funds became available. It raised a more sensitive question: would the next generation of manufacturing technology attract enough customers to justify new investment in highly specialized equipment and facilities? The company said development of its **Intel 14A** manufacturing node could be slowed or discontinued if it failed to secure a major external customer, and that it retained the option of designing some future

products for manufacture at an external foundry rather than producing everything internally. This was the company's statement of strategic alternatives, not evidence that it had decided to discontinue 14A.1

In the same report, Intel said it had slowed construction of advanced manufacturing facilities in Ohio and reviewed construction or expansion projects elsewhere. It warned that high fixed costs make it difficult to reduce existing capacity quickly when demand disappoints. In 2024 and 2025, respectively, it recorded **$3.3 billion** and **$950 million** in manufacturing-asset impairment and accelerated depreciation charges. These figures do not mean that all those amounts concerned Ohio or the 14A node. At company level, they show what can happen when the economics of specialized assets change before their accounting life has ended.1

This is not an incident about an AI data center, but about factories that make chips. The distinction matters: a semiconductor foundry is not a collection of GPU racks, though both share one property—enormous upfront spending on equipment that cannot be resold tomorrow like an ordinary used server. The company's letter also distinguishes cash already spent from money still contractually committed. We cannot add them and call the sum completed cash expenditure.2

In such factories, concrete is poured not only over a production plan but over assumptions about customers, manufacturing processes, competitors, subsidy conditions, and access to capital. A company acts on what it knows today, then bets on demand that will exist when a years-long construction project is complete. If demand changes before the work ends, the question is not “Do we believe in our vision?” It is “Which part of the commitment can we postpone without destroying the possibility of success?”

An uncomfortable paradox follows. A delay portrayed in the press as a missed deadline may partly protect a company from committing more money before the conditions for doing so are clear. It may also signal deeper trouble in the market, operations, or financing. Intel's decision does not prove its delay optimal or profitable. It shows what the company itself states: possessing an alternative **option** for manufacturing is part of the decision's value, not a footnote in a project schedule.

A Ten-Billion-Dollar Contract That Was Not Ten Billion Dollars of Spending

Now consider a decision in a world whose assets look nothing like a chip plant but whose reversal is similarly difficult: cloud infrastructure for the US military. The **JEDI** project sought to provide cloud-computing services for multiple classified and unclassified purposes through a single-provider contract. Microsoft won the competition in October 2019. The **$10 billion figure was a potential ceiling for task orders over as many as ten years if all options were**

exercised. It was neither a payment made on award nor an automatic commitment to spend that amount. The base period was two years, with a guaranteed minimum of $1 million in orders, according to the Department of Defense inspector general.3

At the time, a single provider was not self-evidently wrong to every public body involved. Reviewing a protest against the solicitation terms, the US Government Accountability Office concluded that the Defense Department's justifications for a single-award model were legally permissible and supported, including national-security considerations. That fact matters if we are to avoid rewriting history so that every centralized decision appears foolish even before circumstances changed.4

The department did not ultimately activate that long contractual future. In July 2021, it announced that it was canceling JEDI and beginning termination of the contract, explaining officially that its requirements, the market, and its own cloud experience had changed enough to make the earlier framework unsuitable. Litigation had also delayed implementation, so reducing the case to a single technical cause would falsify it. Still, changing institutional requirements were among the stated reasons, and the replacement revealed what kind of change was sought.5

In December 2022, **JWCC** contracts went to Amazon, Google, Microsoft, and Oracle. The framework became multi-vendor, with firms competing for task orders over a three-year base period and two one-year options, under a potential **$9 billion** ceiling. The difference between ceilings of $10 billion and $9 billion is **not a demonstrated saving of $1 billion**: neither ceiling is actual spending, and the duration, needs, and contractual structure differed. In an official 2023 memorandum, the department's chief information officer described the new framework as a multi-vendor procurement vehicle enabling competition and resource tracking, not a single operating cloud environment that ran every system itself.6

The lesson is not that four providers are always better than one. Multiple providers can add burdens in operations, identity, security, data movement, and coordination. The deeper question is: **what needed to be settled for two years, and what should have remained open for five or ten?** The old structure had reasons when designed, but the facts changed before its potential term was exhausted. The value of a right to choose again does not appear when an agreement is signed. It appears when the alternative unsuitable yesterday becomes the closest match for today's need.

Why the Financial Calculator Is Not Enough

Financial analysis teaches us to examine net present value: what are expected cash flows worth after discounting time and cost? If NPV is positive, an investment appears sensible. But Robert Pindyck's work on investment under uncertainty explains that many spending

decisions combine two properties: **money is hard to recover after commitment, while waiting may reveal new information before commitment**. That conjunction gives delay a financial value resembling an option, even when no one is trading a financial contract in a market.7

This is the economics of **real options**. Land that permits several uses, a contract that grants a right but not an obligation to expand capacity, or a factory whose shell can be built before its adaptable production line is installed all contain options. We do not pay only for assets. We also pay for—or surrender—the right to change strategy when our knowledge changes.

Consider a **wholly invented** calculation in abstract monetary units. An institution faces one investment. If demand proves strong, its NPV will be **+150**; if weak, **−70**; each state has a probability of one-half. Investing now, before learning, yields an expected value of **40**. The institution can instead run a limited test costing **8** and accept a delay estimated to cost **12**; in this hypothetical example, the test reveals demand perfectly before the large commitment. It then chooses the project only in the strong state: half of 150 is 75, less 8 and 12, for an expected value of **55**. Under the exercise's assumptions, the **15** difference is the value of staging the investment.

Fifteen is not a general rule, and the model is not a valuation of any company or project. Perfect information from a test is an extremely optimistic assumption; treating the cost of delay as known and fixed is another simplification. If the experiment provides a weak signal, a competitor can occupy the market while we wait, or the window for obtaining power and land closes, the difference may disappear or reverse. The point is to expose an error in the question. It is not enough to ask about the project's profit. We must also ask about **the price of closing off choice before important information arrives**.

Nor do sunk costs justify further spending. Money already consumed does not return because we finish the job. When a board says, “We have spent too much to turn back,” it may be converting an old loss into a new one. At the next gate, a sound decision compares the prospective value and cost of continuing with the alternatives available from that moment. It does not punish the future to acquit the past.

When the System Was Bought Before Every Hospital Was Understood

Britain's **NHS National Programme for IT** offers a more troubling example than a contract canceled before its full potential life was used. Starting in 2002, the programme sought to modernize patient records and connect health institutions at scale. A 2011 National Audit Office report said the central objective of deploying electronic records according to the original vision would not be achieved as planned, and that detailed-care-record systems were late and being reduced in scope without a proportionate fall in cost.8

By the 2011 report, **£2.7 billion** had been spent on care-record systems, which the auditors found were not delivering value for money on the implementation results then available. In one major contractual region, only **four of 97 systems** had been delivered to acute hospital trusts in seven years. Strategy then shifted from wholesale replacement of institutional systems toward building on and connecting what already existed; the department estimated that interoperability would cost **at least another £220 million**.8

It would be wrong to say all programme spending was wasted or produced no benefit. The programme included different components, and some projects and operational benefits continued. A later review of projected costs and benefits reported that the Department of Health estimated whole-life benefits of **£10.7 billion** against **£9.8 billion** in costs, while substantial uncertainty surrounded the realization of those benefits; about two-thirds of the estimated benefits were still expected after March 2012. A later government review described the 2002–2011 experience as costly and as failing to achieve most of its objectives.9

The paradox is that the programme did not lack a grand vision. It lacked what vision alone cannot buy: precise knowledge of what could be standardized across hospitals with different systems, cultures, and priorities before those assumptions hardened into contracts, integrations, schedules, and training for thousands of workers. Every local system insufficiently tested before a central choice was generalized added to the cost of retreat. Exit may mean more than ending a supplier contract; it may mean redesigning data interfaces, retraining people, and operating two services in parallel for years.

With JEDI, the procurement framework could be changed before its potential maximum was exhausted. In the British programme, the gap between vision and reality became clear only after systems, contracts, and institutional expectations had accumulated. This is not a causal experiment proving that phased procurement would have saved every pound. It reveals different degrees of reversibility and leaves the proper question open: **what information could have been discovered early for less than the cost of rebuilding?**

When the Chip Moves Faster Than the Building That Will House It

In AI, the problem doubles because technology time and physical-asset time do not run on the same clock. In a 2026 report, the International Energy Agency estimated that data centers consumed about **485 terawatt-hours of electricity in 2025**, with a central scenario reaching roughly **950 terawatt-hours in 2030**. These are scenario estimates, not a ledger of electricity already under contract. The agency also described steep change in power density: from 2020 to 2025, AI-server power density increased by about **eleven times** in its analysis, with another near-fourfold increase expected by 2027.10

The issue is not merely buying an old GPU instead of a new one. In 2026, NVIDIA presented its Vera Rubin architecture as a system co-designing processors, high-speed interconnects, memory, liquid cooling, power distribution, and rack operations. The company reports large gains in performance and inference cost over earlier generations. These are **manufacturer claims and comparisons** that should be tested for a purchase decision, not treated as independent results applicable to every operating environment.11

When a new computing facility is built around assumptions about power, heat, connectivity, and workload, those assumptions may change before the equipment generation intended to fill it arrives. The building need not become useless; it may be reconfigured. But some elements are more flexible than others: floor layout, pipe routes, cooling units, electrical breakers, rack dimensions, fiber networks, and the pattern of connections among processors. Other costs never appear in a polished architectural rendering: training teams on particular operating tools, coupling software to a vendor's interfaces, and making long-term capacity agreements before actual usage is known.

It is not enough for a project manager to ask how many devices the site can contain. The manager should also ask how many different architectures it can contain before paid work must be demolished. Scalability is not the same as replaceability. A closed system may expand painlessly within one vendor and still make migration to another architecture resemble a new construction project.

The less certain the future mix of large-model training, real-time inference, industrial workloads, and processing near users, the more independent economic value lies in flexible cooling, power, and networks. The engineering interpretation of real options is simple: **design the slow infrastructure for change, and reserve the fast-changing infrastructure only when the reason for needing it becomes clear**. That principle has limits, to which we will return.

Exit Fees Are Not the Cost of Exit

In March 2024, Amazon Web Services announced qualifying waivers of data-transfer-out fees when customers leave its services, subject to defined processes and conditions; its guidance continued to be updated through October 2026. This is a material commercial development. Where the conditions apply, it removes one category from the transfer bill. It does not guarantee that a database will operate at another provider or that permissions, keys, infrastructure-as-code, and recovery plans will migrate without re-engineering.12

British government guidance on technical cloud lock-in is more precise than “avoid lock-in at any price.” It recognizes that vendor-specific services can offer substantial value in operations, security, and updates, and that always replacing them with lower-function common services

may be economically unsound. It recommends comparing value and portability, estimating the time and cost of leaving, and actually testing migration for some essential components.13

There are therefore three doors, not one. A **legal door** may be closed by exclusivity or a purchase commitment; a **technical door** by intertwined interfaces, data models, and services; a **human door** by team skills, procedures, and operational memory. A supplier may open the first door legally while the keys to the other two remain missing. Unless the exit journey is tested before expansion, the contract will look more flexible than the system can actually bear.

The Small Test That Buys a Larger Decision

Why can a limited experiment be more strategic than a major project begun at once? Not because the word “pilot” makes it innocent, or because brave managers should fear every large decision. There is only one defensible reason: **the experiment must be genuinely capable of changing a later, expensive choice**. A test that confirms what everyone already believes, with no realistic chance of stopping or changing the project, is extra cost dressed as knowledge.

Suppose an institution plans a latency-sensitive inference service for thousands of users. It should not test only a model's performance in a demonstration. It should construct a protected sample of real data and requests, then test peak load, context length, language switching, 95th-percentile latency, cost per completed task rather than price per million tokens alone, justified abstention, and what happens when the model changes. It should migrate a comparable component to a second vendor or alternative architecture and observe cost, time, and quality. Only then does “we can leave” become information grounded in experience.

But the test itself can become a trap. The British government's technology procurement playbook calls the continuation and expansion of a trial without clear contractual gates **pilot creep**. It calls for break points at which an institution can assess readiness to scale, grant a limited extension for a defined learning objective, or bring in new partners. Before a trial starts, its objectives, metrics, knowledge-transfer work, and contract duration should be defined, and the institution should retain internal capacity to understand what is being built for it.14

Knowledge is not a vendor's report declaring its own vendor successful. A criterion must exist whose failure will actually change the course. For example: “We will not sign the expansion until an independent team conducts a data-and-workflow migration exercise, demonstrates compatibility, and measures operating cost after discounts expire.” Another could be: “No expansion until the service is tested with a group unlike the users in the marketing demonstration.” If failure has no effect on procurement, these conditions are ceremonies, not gates.

This proposal is not foreign to public procurement practice. The US modular-contracting framework for information technology calls for large system acquisitions to be divided, as far as practicable, into interoperable increments, with contracts designed not to bind an agency to every later phase. British AI procurement guidance likewise presents discovery and proof-of-concept phases as ways to reduce risk before full operation.15

Most importantly, **a trial's size is measured not by its invoice but by the value of the decision it can correct**. A small test that mispredicts performance at scale is not a real option. A focused test that removes a decisive unknown or refutes an operating assumption can buy an institution the right not to incur a larger loss.

When Waiting Becomes a Lock-In of Its Own

Here is the strongest objection to everything above: the option we think we are preserving through delay may evaporate over time. Scarce power sites may be reserved, transformer queues lengthen, the best talent may join competitors, and purchase agreements may be signed on terms that are no longer available later. The International Energy Agency explains that rapid AI expansion is colliding with slower energy and supply infrastructure, including constraints in high-bandwidth memory and grid and electrical equipment. When a site window is scarce, securing some rights early may be better than waiting for every technical detail to become clear.10

Nor does option theory guarantee that delay is optimal when firms are competing. Research on option-exercise games has shown that the risk of competitors investing first can reduce the value of waiting. Here option value depends partly on competitors' responses, not only on the degree of uncertainty.16

Imagine a computing project in Saudi Arabia—**a conceptual example, not a description of an existing project**. It needs an electricity source available over years and cooling infrastructure, followed by processors, a model platform, and applications that can be purchased on a faster technology cycle. One mistake is to postpone applying for a site and grid connection until every processor competition is settled; the required capacity may be gone when the technical choice is ready. The other is to sign today for every rack and years of services from one platform before observing the real production workload.

An intelligent decision is not a verbal compromise. It means **different commitment dates for assets of different kinds**. Reserve a site right or capacity-expansion option when the lead time to obtain it exceeds the technology's rate of change. Install connections that permit more than one configuration. When service will not be disrupted, postpone purchases that lose value

quickly; divide demand between base reservations and conditional expansion options. Examine sales or partnership contracts so it is clear who bears the cost of unused capacity and who benefits from a right to reduce it.

Managerial intuition alone cannot identify the better course. The institution needs a **curve of learning time against the time at which an alternative disappears**. Information arriving in a year has no decision value if the relevant site vanishes in a month. Conversely, an inescapable five-year agreement in a market whose operating costs change every quarter may turn a small price advantage today into a continuing burden tomorrow.

The Exitability Ledger

I propose an operating tool for technical and industrial projects called the **Exitability Ledger**. This is an original, untested analytical proposal, not a published professional standard or a guarantee of profitability. Before any hard-to-reverse decision, it records six things: the assumption on which the decision rests; when new information about that assumption will arrive; how much money becomes sunk before then; the cost of leaving the contract, infrastructure, and operating process; the alternative that will disappear if the institution waits; and the stop, expand, or switch decision the new information will make possible.

An initial measure could be the **commitment–learning gap**: the interval between the moment a material share of expenditure becomes unrecoverable and the first date on which enough evidence exists to test the assumption that justified it. Not every long gap is an error. Some factories, roads, and networks must be built before use. But a long gap prompts a question: can some uncertainty be transferred into an option contract, a design stage, a conditional purchase agreement, or a customer commitment made before capacity expands?

At each gate, three levels should be examined without reducing one to another:

1. **The asset:** can the equipment be sold, reused, or adapted, and what is the actual rather than merely accounting loss?
2. **The contract and data:** is there a clear right to terminate, export, transfer licenses, and recover files, interfaces, and documentation? Has exercising those rights been tested?
3. **The institution:** is there a team capable of operating the alternative, or has the vendor come to own the system's knowledge even without owning its code?

The ledger's hard test is to force an answer to this question: **what will we do if our assumptions about demand or technology prove wrong?** If the only answer is “we will need to persuade the market,” there may be no alternative plan—only a compensatory marketing plan.

When an institution can identify a clean stop gate, known exit invoices, and a tested migration path, it has acquired an invisible asset alongside its physical ones: the ability not to continue in the wrong direction.

The test may also refute the value of all this caution. If portable architecture adds substantial cost throughout a project's life, the probability of change is low, and a specialized service brings large benefits, deliberately accepting technical dependence may be justified. British guidance recognizes this paradox: perfect independence is neither free nor always best. The important thing is for dependence to be a **measured choice**, not a fact discovered on the day the contract ends.13

A mature strategy does not keep every door open forever. Construction, research, and expansion require many doors to close eventually. But strategy must know which door closes first, which right is worth buying so another remains open a little longer, and which door will disappear if hesitation continues.

The difference between a hurried project and an ambitious one is not visible in the size of an engineering plan, the contract price, or the length of a press release. It appears when new information contradicts the original plan: can the institution change its mind without declaring the world wrong because it has already signed?

The project's greatest cost may not be the money we spent. It may be **the later decision that spending prevented us from making**.

References

1. Intel, Fiscal Year 2025 Form 10-K, US Securities and Exchange Commission, 2026.
2. Intel, Ohio site manager's cover letter and annual onshoring incentive report, February 27, 2025.
3. US Department of Defense Office of Inspector General, Report on the JEDI Cloud Procurement, 2020.
4. US Government Accountability Office, Statement on Oracle's JEDI bid protest, 2018.
5. US Department of Defense, Future of the JEDI Cloud Contract, 2021.
6. US Department of Defense, Award of JWCC contracts, 2022; Office of the Chief Information Officer, memorandum on rationalizing cloud use, 2023.
7. Robert S. Pindyck, Irreversibility, Uncertainty, and Investment, NBER Working Paper 3307, 1990.
8. UK National Audit Office, The National Programme for IT in the NHS: an update on the delivery of detailed care records systems, 2011.
9. UK National Audit Office, Review of the final benefits statement for programmes previously managed under the National Programme for IT in the NHS; The use of digital technology in the NHS, 2020.
10. International Energy Agency, Key Questions on Energy and AI: Executive Summary, 2026.
11. NVIDIA, Inside NVIDIA Rubin GPU Architecture, 2026. This source describes manufacturer specifications and claims.
12. AWS, Free data transfer out to internet when moving out of AWS, 2024, with later updates.
13. UK Government, Managing technical lock-in in the cloud.
14. UK Government, Digital, Data and Technology Playbook, 2023.
15. US Government, FAR 39.103: Modular Contracting; UK Government, Guidelines for AI procurement.
16. Strategic Delay in R&D Competition, *Review of Economic Studies*, 2002; Option Exercise Games, *Review of Financial Studies*, 2002.