Metamorphosis: From Teraflops to Gigawatts

I woke up in Singapore to a hyperscaler shedding roughly US$190 billion in market value in the pre-market.[1]

Half a world away in New York, the market was digesting the largest capital expenditure announcement in corporate history: two hundred billion dollars, under one name and one year.[2] To give you some perspective, this was enough to fund two-thirds of the entire NASA Apollo program (roughly US$260-320 billion in 2026 dollars, depending on inflation, with NASA's $25.8 billion nominal cost by CPI or by aerospace-sector indices).[3] Most of it primarily went towards data centers, chips, and power.[4]

The funniest thing was that Amazon wasn't even the story. It was the cherry on top of the cake that had been baking for two weeks.

Microsoft had opened up the sequence in late January, reporting US$37.5 billion of capex in a single quarter — a US$150 billion annual pace, up 66%. Furthermore, Microsoft's report signaled to investors that the spend would keep accelerating.[5] Its stock sold off.[6] Alphabet followed with US$175-185 billion, which was close to double its 2025 spend of US$91 billion. Sold off.[7] Meta guided its number to US$115-135 billion, up from US$72 billion.[8] Finally, Amazon walked in last, on a Thursday, and placed the largest capital expenditure number in corporate history on the table.[9] By the time I got up in Singapore on Friday morning, the verdict

---

[1] Amazon reported Q4 2025 after the close on Thursday 5 February 2026. Shares fell about 11% in after-hours trading and closed down roughly 8% on Friday 6 February. Market capitalisation going into the print was approximately US$2.4 trillion, so an 8% move is on the order of US$190 billion. CNBC, "Amazon stock falls 8% on $200 billion spending forecast, earnings miss," 6 February 2026, https://www.cnbc.com/2026/02/05/amazon-amzn-q4-earnings-report-2025.html. Across the nine-session slide that followed, Amazon shed more than US$450 billion, its worst run since 2006.

[2] https://www.sec.gov/Archives/edgar/data/1018724/000101872426000002/amzn-20251231xex991.htm

[3] Project Apollo cost $25.8 billion in nominal dollars across FY1960–1973 per NASA's congressional budget submissions (~$28 billion including Project Gemini and the robotic lunar precursor programs). The Planetary Society's Apollo cost study, which escalates each year's actual appropriation using the NASA New Start Index — an aerospace-specific deflator that tracks engineering and materials costs rather than consumer prices — puts this at roughly $257 billion in 2020 dollars. Carried forward to 2026 (CPI-U 258.8 → ~330, a 1.28x multiplier; aerospace escalation at 2.5–4% annually gives a similar 1.16–1.27x range), the full program lands at approximately $300–330 billion in 2026 dollars. A simple CPI adjustment from the program's spending-weighted midpoint (~1966) gives a lower bound near $260 billion. Amazon's 2026 capex guidance of ~$200 billion therefore equals roughly two-thirds of the entire Apollo program — per year.

[4] Jassy told analysts the 2026 spend would go "predominantly" to AWS — data centres, custom silicon and networking. Q4 2025 earnings call, 5 Feb 2026.

[5] https://www.microsoft.com/en-us/investor/events/fy-2026/earnings-fy-2026-q2

[6] Microsoft fell 7% in extended trading on 28 January 2026 and closed the following session down roughly 10%, erasing about US$357 billion of market value — its worst single day since March 2020.

[7] Alphabet, Q4 and fiscal year 2025 results, 4 February 2026: 2026 capital expenditure guidance of US$175–185 billion against full-year 2025 capex of US$91.4 billion. The stock fell 7.4% on heavy volume.
https://www.sec.gov/Archives/edgar/data/1652044/000165204426000012/googexhibit991q42025.htm

[8] https://investor.atmeta.com/investor-news/press-release-details/2026/Meta-Reports-Fourth-Quarter-and-Full-Year-2025-Results/default.aspx

[9] 5 February 2026 fell on a Thursday; Amazon reported after the close, so the market reaction landed on Friday 6 February.

was in and clear: all three out of four companies had a combined commitment of US$490-520 billion — and Microsoft hadn't even announced their full number yet.[10] It was only in April, when the answer turned out to be US$190 billion.[11] Bringing the four companies' total to a record of roughly US$700 billion, 70% more than a year that was itself an all-time record.[12]

The capital expenditure figure forced everyone to question if this was a bubble. But no one is asking the better question. What if this isn't an overpriced mania, but a mispriced metamorphosis? The largest software and technology companies that we've spent the past twenty years loving because they were asset-light and owned nothing but software and networks are being converted, quarter over quarter at a run rate approaching a trillion dollars a year, into infrastructure utilities. Capital intensity that belongs to railroads, debt that belongs to telecoms, and power contracts that could feed nations. An evolution is underway. The same caterpillar that built apps is becoming a butterfly that owns watts.

And the transition is silent. Not because it is hidden but simply because it is happening in the one phase of the current cycle that the market's traditional instruments were never built to read in real time. Traditional metrics like free cash flow, operating margin, depreciation — none of those are obsolete, and none of them are wrong. But they are harvesting metrics, and this is planting season.

Part I – The Right Metric

Okay, beginning with the dollar. Aggregate capital expenditure across the four largest cloud providers was roughly US$226 billion in 2024.[13] Then US$410 billion in 2025.[14] And now US$700 billion is guided in 2026. So take the ratio: the number compounds roughly at 1.7-1.8x per year, across three years and four firms that would prefer to spend less than each other.[15] For 2027, sell-side projections already exceed US$1 trillion, which isn't a bold forecast.[16] Why?

---

[10] Arithmetic: Alphabet US$175–185 billion + Meta US$115–135 billion + Amazon US$200 billion = US$490–520 billion.
[11] Microsoft, fiscal year 2026 third quarter results, 29 April 2026. CFO Amy Hood guided calendar-year 2026 capital expenditure to roughly US$190 billion, up about 61% from 2025 and some US$35 billion above the US$154.6 billion Visible Alpha consensus, of which approximately US$25 billion reflects higher component pricing. https://www.microsoft.com/en-us/investor/events/fy-2026/earnings-fy-2026-q3
[12] Arithmetic. 2026: Microsoft US$190bn + Alphabet US$180bn (midpoint) + Meta US$125bn (midpoint of the January guide) + Amazon US$200bn = US$695bn. 2025 actuals: Amazon US$131.8bn + Alphabet US$91.4bn + Meta US$72.2bn + Microsoft calendar-2025 of roughly US$118bn (FY25 Q3 US$21.4bn + FY25 Q4 US$24.2bn + FY26 Q1 US$34.9bn + FY26 Q2 US$37.5bn) ≈ US$413bn. Growth ≈ 68%.
[13] 2024 company filings: Amazon US$77.7bn, Microsoft approximately US$55.7bn on a calendar basis, Alphabet US$52.5bn, Meta US$39.2bn ≈ US$225bn. The Belfer Center puts the same aggregate at over US$200 billion, a 62% year-over-year increase from 2023. https://www.belfercenter.org/research-analysis/ai-data-centers-us-electric-grid
[14] See the preceding note for the component arithmetic (approximately US$413bn).
[15] 410 ÷ 226 = 1.81; 700 ÷ 410 = 1.71.
[16] Morgan Stanley's Ben Reitz put 2027 hyperscaler capital expenditure at approximately US$1.1 trillion, having raised the 2026 figure to US$805 billion, and noted that US$1.1 trillion would match the combined capex of every non-technology company in the S&P 500. Goldman Sachs' Ryan Hammond observed that consensus implies US$920 billion in 2027 and argued the base case is nearer US$1.1 trillion, with a bull case of US$1.4 trillion; Goldman's tracked series shows roughly US$1,018 billion for 2027E. https://finance.yahoo.com/sectors/technology/articles/goldman-says-consensus-2027-hyperscaler-140152065.html

There isn't a reason to bend that line. Now play it forward further and the annual figures start looking absurd: Goldman Sachs's baseline of US$7.6 trillion in cumulative investment across compute, data centers, and power between 2026 and 2031 already factors in growth rate deceleration.[17] No one modelling this line believes that it can run at 1.8x forever. If you plot these on a graph, you'll find yourself looking at something like a straight line. Let's talk about the demand that reflects this cumulative investment. A primary driver is that capacity is being monetized as quickly as it's being built, and that demand exceeds supply. In 2025, AWS added 3.9 GW of power capacity, expects to double that power capacity by the end of 2027, and is monetizing that capacity as fast as it is being installed whilst continuing to face capacity constraints that yield unserved demand (For scale: 7.8 GW of installed data center power capacity is enough to power the peak demand of Singapore, depending on usage assumptions).[18] Microsoft has energized nearly 1 GW in a single quarter and as of April 2026, we have Microsoft spending nearly US$200 billion in one year on aggressive expansion of their data centers and they still expect capacity to remain constrained at least through mid-2026, even as they bring GPU, CPU, and storage capacity online quicker.[19] And then there's Oracle, which is late to the game and is compensating with one of the most aggressive balance sheet campaigns in corporate history. Issuing debt, diluting shareholders, just to fund its transition to become a cloud infrastructure company.[20] Whether Oracle wins is besides the point. The point is that we have a fifty-year-old software company that is leveraging its balance sheets just to build data centers, giving you a rough idea of the demand on the other side (I will go into what the credit market thinks further down in this essay). Finally, we have neoclouds whose businesses have grown rapidly because hyperscalers cannot meet all GPU demand. CoreWeave, a company that barely existed five years ago, holds 3.5 GW under contract.[21]

Now, back to that straight line I mentioned earlier, the dollar line isn't going to run forever. If you extrapolate at 1.7-1.8x annual growth, you will eat the combined capex of the entire S&P

---

[17] Goldman Sachs Global Institute, "Tracking Trillions: The Assumptions Shaping the Scale of the AI Build-Out." The baseline model projects approximately US$7.6 trillion of capital between 2026 and 2031 across compute, data centers and power, with annual spend rising from US$765 billion in 2026 to US$1.6 trillion in 2031. Disclosed assumptions include a 1.2 PUE, US$15 million per MW for data centers and US$2,500 per kW for new power.
https://www.goldmansachs.com/insights/articles/tracking-trillions-the-assumptions-shaping-scale-of-the-ai-build-out

[18] https://www.aboutamazon.com/news/company-news/amazon-ceo-andy-jassy-2025-letter-to-shareholders;
Author's calculation. AWS's disclosed 3.9 GW of 2025 additions, doubled per company guidance, implies ~7.8 GW of installed data center power capacity by end-2027. Singapore's peak system electricity demand is approximately 7.5–8 GW (Energy Market Authority of Singapore, Singapore Energy Statistics — ema.gov.sg. The comparison is capacity-to-peak-demand: data centers run at high, near-continuous utilization, so on an energy (TWh) basis the fleet would consume more than a peak-demand comparison implies. The figure is illustrative of scale, not a claim of equivalence.

[19] https://www.microsoft.com/en-us/investor/events/fy-2026/earnings-fy-2026-q3

[20] Oracle raised US$43 billion of debt and US$5 billion of equity in fiscal 2026, with a further approximately US$40 billion planned for fiscal 2027 including a US$20 billion at-the-market equity programme. Total debt has passed US$125 billion, and S&P cut Oracle to BBB– in July 2026. Oracle, Q4 and fiscal year 2026 results, 10 June 2026.
https://investor.oracle.com/investor-news/news-details/2026/Oracle-Announces-Record-Q4-and-FY-2026-Results-Driven-by-Cloud-Infrastructure--Cloud-Applications/default.aspx

[21] CoreWeave, Q1 2026 results, 7 May 2026: over 3.5 GW of contracted power and over 1 GW of active power.
https://www.cnbc.com/2026/05/07/coreweave-crwv-q1-earnings-report-2026.html

500 within a few years, and it isn't difficult for anyone with a calculator to piece that together.[22] Lines like this don't break on financing. Money is relatively flexible in the sense that banks can always expand lending, bond markets can absorb an enormous amount of issuance, private equity and sovereign wealth funds can always supply capital, and the reasons can go on. Lines like this break on physics: land, transformers, and grid interconnection. Hence, using the dollar going forward will be the wrong unit of account.

Let's swap over to the watt now.

The infrastructure buildout has three tiers that are occasionally conflated and should not be. There is aspirational capacity, contracted capacity, and active energized capacity. Aspirational capacity is announced infrastructure plans and multi-year ambitions, which is where the press releases live. Contracted capacity reflects gigawatts attached to signatures, binding commitments such as power purchase agreements, take-or-pay compute contracts, and project financing already drawn. Finally, we have active energized capacity, essentially what is already physically drawing power into compute and generating revenue today.[23] The market knows how to judge the first tier, and completely ignore the second. The evidence in this essay revolves around contracted and active energized capacity.

Starting with energized capacity, as I mentioned above, Amazon added 3.9 GW of power capacity in 2025, roughly the output of three to four nuclear reactors, brought online in just one singular year by one company, and expects to double its total capacity by 2027.[24] Microsoft stood up more than 2 GW of capacity in 2025, added a further 1 GW in only the first quarter of 2026, and is on track to double its overall footprint over the next two years.[25] Contracted capacity is being accumulated in chunks that would have been national news a decade ago. CoreWeave and Nebius both have roughly 3.5 GW under contract.[26] Playing this forward, you get a buildout that stops looking like a corporate spending program and starts looking more like the electrification of a mid-sized industrial country.

If the extrapolation sounds unserious or if there is any doubt regarding the contracted and energized capacity, just take a look at what the infrastructure builders are doing with their own

---

[22] Morgan Stanley made the same observation from the other direction: a US$1.1 trillion 2027 hyperscaler figure would equal the combined capital expenditure of every non-technology company in the S&P 500.

[23] The taxonomy is the author's, but it maps directly onto Nebius's own reporting convention, which distinguishes contracted power, connected power and active power. Nebius Group N.V., Q4 2025 shareholder letter, Form 6-K. https://www.sec.gov/Archives/edgar/data/1513845/000110465926013946/tm266173d1_ex99-2.htm

[24] A typical US nuclear reactor is roughly 1 GW of nameplate capacity (US Energy Information Administration), so 3.9 GW is three to four reactors' worth of output.

[25] Microsoft added more than 2 GW across calendar 2025, roughly 1 GW in the December 2025 quarter and approximately another gigawatt in the March 2026 quarter, and has said it is on track to double its total data centre footprint within two years. Fiscal year 2026 second and third quarter earnings calls.

[26] CoreWeave, Q1 2026: over 3.5 GW contracted. Nebius Group N.V., Q1 2026 results, 13 May 2026: over 3.5 GW contracted, ahead of its prior 3 GW year-end target, with guidance raised to more than 4 GW by end-2026. https://nebius.com/newsroom/nebius-reports-first-quarter-2026-financial-results

money: buying the watts in dollars with irreversible, decade-scale commitments. Microsoft recently signed a twenty-year agreement with Constellation Energy to restart Three Mile Island.[27] Amazon secured up to 1.2 GW of carbon-free energy directly from the Comanche Peak Nuclear Power Plant and is building "Project Spectrum," which is a US$5 billion data center on adjacent land.[28] Meta agreed to a 20-year Power Purchase Agreement with Vistra (PPA) for 2.609 GW of carbon-free energy from its Perry, Davis-Besse, and Beaver Valley nuclear plants.[29] Alphabet couldn't seem to find a sufficient power contract, so it simply bought the power developer: Intersect Power for US$4.75 billion.[30] I hope this clears some doubt with regards to how seriously the infrastructure players are treating power capacity. You don't simply recommission an entire nuclear plant to serve a trend you expect to break.

One more measurement, because it helps to quantify the transformation that this essay is named for. Capital intensity, which is capex as a share of revenue, ran near 10% for these firms across the software era.[31] The entire asset-light business model rested on that number staying small, and dominating through intangible assets like software, network, and brand power. Now, this number runs between 45-57%.[32] For calibration, that is not the profile of a software buildout; that is the profile of an intense utility buildout: railroads in the 1880s, a telecom in the fiber buildout, or an electric utility company during a major generation and grid expansion. Income statements are starting to register it. For example, Alphabet's depreciation expense rose 38% from US$15.3 billion to US$21.1 billion in a single year.[33][34] Its depreciation is simply past capex arriving on the income statement with a delay. The metamorphosis in the title is not a metaphor; it's real, and it's visible in the filings to anyone reading the right numbers in the right units. These software and technology companies with their traditionally asset-light roots are pivoting away from that and into asset-heavy utility and infrastructure companies.

---

[27] Constellation Energy, 20 September 2024: a twenty-year power purchase agreement with Microsoft supporting the restart of Three Mile Island Unit 1 as the Crane Clean Energy Center, 835 MW.
https://www.constellationenergy.com/newsroom/2024/Constellation-to-Launch-Crane-Clean-Energy-Center-Restoring-Jobs-and-Carbon-Free-Power-to-The-Grid.html

[28] Vistra disclosed approximately 3.8 GW of nuclear under long-term power purchase agreements, including 1,200 MW over twenty years with Amazon Web Services at Comanche Peak (Q4 2025 results and earnings call, February 2026). Project Spectrum is an approximately US$5 billion AWS campus on some 435 acres adjacent to the plant, up to eighteen two-storey buildings, with construction expected to begin in 2027.
https://www.datacenterdynamics.com/en/news/amazon-looks-to-build-large-scale-data-center-campus-next-to-nuclear-plant-in-texas/

[29]https://investor.vistracorp.com/2026-01-09-Vistra-and-Meta-Announce-Agreements-to-Support-Nuclear-Plants-in-PJM-and-Add-New-Nuclear-Generation-to-the-Grid

[30]https://abc.xyz/investor/news/news-details/2025/Alphabet-Announces-Agreement-to-Acquire-Intersect-to-Advance-U-S--Energy-Innovation-2025-DVIuVDM9wW/default.aspx

[31] Software-era capital intensity, from company filings. Microsoft FY2015: US$5.9bn capex on US$93.6bn revenue = 6%. Alphabet 2015: US$9.9bn on US$75.0bn = 13%. Meta 2015: US$2.5bn on US$17.9bn = 14%. Mean ≈ 11%.

[32] 2026 guidance against trailing or current-year revenue. Microsoft: US$190bn ÷ approximately US$330bn CY2026E ≈ 57%. Meta: US$125–135bn ÷ approximately US$235bn 2026E ≈ 53–57%. Alphabet: US$180bn ÷ approximately US$400bn 2025 ≈ 45%. Amazon sits well outside this band at roughly 28%, because the great majority of its revenue is retail rather than cloud.

[33] Anat Ashkenazi, Alphabet Q4 2025 earnings call, 4 February 2026: depreciation increased by nearly US$6 billion, or 38%, from US$15.3 billion in 2024 to US$21.1 billion in 2025.
https://abc.xyz/investor/events/event-details/2026/2025-Q4-Earnings-Call-2026-Dr_C033hS6/default.aspx

[34] https://www.wsj.com/finance/stocks/big-tech-accounting-creates-a-blind-spot-in-the-ai-boom-e2d7d1eb

All this information that I've provided is disclosed figures, public contracts, and the markets have seen every number in this section. They see the capex line and see that spending must revert. It's not a cost curve; it's a supply curve that is chasing a demand curve that it has not once even caught.

Part II – The Demand Is Pre-Sold

What separates this buildout from other buildouts in the past doesn't come from size. It comes from capacity being bought before it even exists. And I've compiled the evidence, stacked in order of contractual weight.

Exhibit one: Oracle. Oracle has reached US$553 billion in RPO (Remaining Performance Obligations), up 325% year over year, which is revenue contractually committed but not yet delivered.[35] For clarification, RPO is not a forecast, pipeline, or management aspiration. It's an accounting category governed by ASC606 reserved for signed contracts.[36] Why I bring up Oracle's RPO is because there is half a trillion dollars of committed demand sitting on the books of a company that has yet to construct the buildings to serve it. The primary constraint disclosed in Oracle's filings is not sales but physical infrastructure constraints like concrete, transformers, and time.[37]

Exhibit two: Microsoft. Microsoft has disclosed a backlog of roughly US$80 billion in Azure demand that it is currently unable to serve.[38] Despite describing building at the quickest rate they've ever built, Microsoft is being forced to take computing power away from their own internal products to preserve capacity for their paying cloud customers.[39] Think about that for a bit. One of the most valuable companies on earth is rationing compute away from their own products that they have been selling for the past twenty years, because external customers have contracted more computing infrastructure than what currently exists today.

---

[35] Oracle, fiscal year 2026 third quarter results, 10 March 2026: remaining performance obligations of US$553 billion, up 325% year over year and up US$29 billion sequentially, with most of the increase attributable to large-scale AI contracts. https://www.oracle.com/news/announcement/q3fy26-earnings-release-2026-03-10/

[36] ASC 606-10-50-13 requires entities to disclose the aggregate transaction price allocated to performance obligations that are unsatisfied or partially unsatisfied at the reporting date. It is a measure of contracted consideration, not of pipeline or forecast.

[37] Oracle's Q3 FY2026 disclosure notes that for most of the new AI contracts the required equipment is funded upfront through customer prepayments or supplied directly by the customer — the company's own framing of the constraint as build-out rather than sales or financing.

[38] Amy Hood, Microsoft fiscal year 2026 second quarter, 28 January 2026: roughly US$80 billion of Azure orders Microsoft cannot serve until new capacity comes online. Total commercial remaining performance obligations reached US$625 billion, up about 99% year over year, rising to US$627 billion at Q3.

[39] Hood, fiscal year 2026 second quarter call: Microsoft must balance meeting growing Azure demand against expanding first-party AI usage across M365 Copilot and GitHub Copilot, against increased allocations to research and development teams, and against replacement of end-of-life equipment. https://www.microsoft.com/en-us/investor/events/fy-2026/earnings-fy-2026-q2

Exhibit three: Amazon. AWS sits at a revenue run rate of about US$142 billion and consistently reports that its capacity is monetized as quickly as it's being built.[40] The whopping 3.9 GW added in 2025 did not create any form of buffer for Amazon; it was all absorbed and used on arrival.

Exhibit four: CoreWeave. The biggest neocloud reports a revenue backlog of US$99.4 billion against 3.5 contracted GW.[41] I will go more in depth with CoreWeave's balance sheet later. But for now, know that their backlog is made of signed take-or-pay commitments from the largest technology companies in the world.[42]

There are softer signals that round out the exhibits. Meta, with a spending order of US$130 billion a year, reports that it frequently gets regular inbound offers for their computing power.[43] Offers so frequent that it created enough pressure for Meta to launch its own external compute business within months.[44] We also have evidence that requires no disclosure at all because it is structural. But the existence of neoclouds, like Nebius and CoreWeave, companies that rent GPU capacity at a premium, shouldn't even exist. In a market where hyperscaler supply met demand, these companies would not be as relevant as they are today. The existence of neoclouds reflects the fact that hyperscalers spending US$700 billion a year just isn't enough. The existence of the middleman is itself the demand signal.[45]

But a contract is only as good as the balance sheet behind it, and Oracle's half-a-trillion-dollar backlog is owed by a company that isn't even making a tenth of that in cumulative revenue.[46] Calling that "pre-sold" is a bit of a stretch. A commitment from a buyer who cannot yet pay is speculative capacity disguised as contracted capacity. If the entire order book looked like that, I would not be writing this. This is just a piece of the picture. By using who stands behind these contracted commitments, I've graded the demand into three tiers. Starting from the first tier that doesn't announce itself as demand, hyperscalers are buying from themselves. Microsoft is

[40] Andy Jassy, 2026 letter to shareholders, and the Q4 2025 earnings call: AWS reported 24% year-over-year growth on a US$142 billion revenue run rate, with capacity monetised as fast as it is installed and capacity constraints still yielding unserved demand.
[41] CoreWeave, Q1 2026, 7 May 2026: revenue backlog of US$99.4 billion, up nearly 50% sequentially and close to four times year over year; over 3.5 GW contracted and over 1 GW active. Remaining performance obligations rose to US$98.8 billion from US$14.7 billion a year earlier.
[42] Q1 2026 bookings exceeded US$40 billion, including a US$21 billion Meta agreement running to December 2032 and a new Anthropic relationship. Ten clients are now committed to at least US$1 billion each, and non-investment-grade AI customers fell to under 30% of backlog. Of the backlog, 36% is expected to be recognised within twenty-four months and 75% within four years.
[43] Mark Zuckerberg, Meta shareholder meeting, May 2026: outside companies approach Meta almost every week asking it either to stand up an API service or to sell them compute at a premium to what Meta paid for it. He added that no external deals had been struck because internal demand had absorbed the capacity.
[44] Zuckerberg announced Meta Compute as a top-level initiative on 12 January 2026, targeting tens of gigawatts this decade. Bloomberg reported on 1 July 2026 that Meta was building a cloud business to sell surplus capacity and model access externally; Meta shares rose approximately 8.8% while CoreWeave fell 10.8% and Nebius 12.4%.
[45] The structural point is visible in the contracts themselves: Nebius's US$17.3 billion supply agreement with Microsoft (September 2025) and US$27 billion agreement with Meta (March 2026), and CoreWeave's Microsoft, OpenAI, Meta and Anthropic book. These are hyperscalers renting third-party capacity.
[46] Oracle's fiscal year 2026 revenue was approximately US$69 billion (FY2025: US$57.4 billion; Q4 FY2026 revenue US$19.18 billion, up 21%), against fiscal year 2027 guidance of US$90 billion. Oracle Q4 FY2026 release, 10 June 2026.

rationing compute away from its own products. AWS is monetizing fresh capacity on arrival. This is demand in its purest form. No external counterparty, no customer default risk, and no question of whether the demand is real because the hyperscaler is self-aware of the capacity that it needs. The customer is the infrastructure owner. The second tier would be contracted demand to buyers who could clear the bill multiple times over: Meta being accountable for 21% of CoreWeave's backlog is a good example.[47] Essentially, a promise from someone who keeps their promises. Finally, we have our venture-stage labs, OpenAI, which is the loudest of them, signing for capacity that reflects their ambition for tomorrow over what is actually possible today. A demand signal, but it shouldn't be confused with demand security like the hyperscalers I mentioned above. It reflects their ambition but not what they can actually pay.[48]

Therefore, the honest version of the argument is narrower than the headline claim: the buildout is pre-sold to the extent that the buyers are real. The problem with the market is not that they are pricing the third tier skeptically. They should. The real problem is that they price the self-funded and investment-grade majority of the demand at tier three discounts. Not a single person has rewarded these companies for trying to build the capacity they need.

Plot Part I's supply figure with this section's demand evidence on the same chart and you'll get a strange picture: the fastest capacity expansion in corporate history, building capacity flat out, constantly behind its order books. Every quarter the builders add reactors' worth of power, every quarter their backlog grows faster. Just think about it for a second.

As I said before, this is all publicly disclosed information. Everyone has seen this chart. Which begs the question this essay exists to answer: Why can't it read it?

Part III – Why the Market Cannot See It

It's all about instruments. For twenty years, technology companies and their valuation were tied to an asset-light software infrastructure: gross margins usually near 90%, marginal costs near

[47] US$21 billion Meta commitment ÷ US$99.4 billion backlog = 21.1%. CoreWeave Q1 2026 disclosure and the April 2026 Meta agreement.
[48] OpenAI's disclosed compute commitments include approximately US$300 billion with Oracle, US$250 billion with Microsoft and US$38 billion with Amazon Web Services — more than half a trillion dollars against an approximately US$25 billion annualised revenue run rate. HSBC Global Investment Research, November 2025.

zero.[49] The toolkit to calibrate these companies was simple. Price-to-earnings, free cash flow yield, operating margin. That toolkit was correct, and using that toolkit would reflect that capital spending was a deduction from the company's value. Apply this toolkit to a technology company deliberating becoming a utility company, and every dial flashes red. Just to be clear, the instruments are not broken. At the end of the day, it really is all about free cash flow. The valuation skeleton of a business really boils down to present and future free cash flow. Compare a dollar on a buyback to a dollar on substation land, and one of them would be generating electricity in 2055. A dollar on a buyback can be spent whenever; a dollar spent to secure land, substations, and power capacity. And they might be unable to buy the power they need in 2055 at anything remotely close to today's price. Negative free cash flow during a buildout isn't a loss; it is a rate of construction, the rate at which a company is converting capital into future productive capacity. Why are we punishing that? The market always overestimates technology in the short-term but strongly underestimates technology in the long-term.[50]

As I've said before, the unit is the watt, and in this phase, you can count what the accounts cannot yet see. Let's look at this from a technical standpoint.

A gigawatt is a billion watts, the output of one nuclear reactor.[51] When a data center company says "1 GW", it refers to the facility power at the meter. The GW doesn't purely reach just the chips. Some of the GW goes towards cooling, conversion, and distribution. This is what engineers express as PUE, or power usage effectiveness. PUE usually eats up around 20-30% of a GW. Engineers capture PUE primarily through a ratio: the ratio of total facility power to the power actually delivered to IT equipment. So a PUE of 1.2, for example, means that 1.0 MW is consumed by actual compute: servers and GPUs. And 0.2 MW goes towards cooling, conversion, distribution, and other infrastructure needs. Hence, a PUE of 1.3 means that 1.0 MW of compute power is delivered for every 1.3 MW drawn by the facility. So you could say that a 1 GW facility delivers roughly 800 MW of actual IT load.[52]

So fill it out. An H100-class draws about 700 W at the board;[53] wrap that in a server with CPUs, memory, networking, and storage, and you'll be looking at around 1.2 KW per GPU of IT load, assuming the facility has a PUE of 1.25. 1.2 KW x 1.25 would give you around 1.5 kW at the meter. So each H100-class GPU represents roughly 1.5 kW of total facility power demand, after PUE. Divide that number out: 1 GW of facility power hosts on the order of 650,000 to 700,000

[49] The reversal is already legible in the accounts: Microsoft's Q3 FY2026 gross margin of 67.6% was its narrowest since 2022, specifically because data-centre depreciation now runs through cost of revenue. CNBC, 29 April 2026.
[50] Amara's Law, attributed to Roy Amara of the Institute for the Future: we tend to overestimate the effect of a technology in the short run and underestimate it in the long run.
[51] US nuclear reactors average roughly 1 GW of nameplate capacity (US Energy Information Administration). For reference, Comanche Peak is 2.4 GW across two units and Beaver Valley 1.87 GW across two.
[52] Power usage effectiveness is defined by The Green Grid and standardised as ISO/IEC 30134-2: total facility energy divided by IT equipment energy. Goldman Sachs' AI build-out model assumes a 1.2 PUE for new hyperscale capacity. Arithmetic check: 1 GW ÷ 1.25 = 800 MW of IT load.
[53] Nvidia H100 SXM5 board thermal design power is 700 W (Nvidia H100 datasheet).

H100-class GPUs.[54] Blackwell-generation parts draw more per unit but deliver more per watt, making the basic order of magnitude the same.[55] So when Amazon says it added 3.9 GW in 2025, it's describing about 2.5 million top-end accelerators, activated and energized in a singular year, all by one company.[56]

Run the energy dimension. 1 GW running around the clock is exactly 8.76 terawatt-hours a year.[57] Amazon's 3.9 GW works out to 34 TWh annually. To give you an example from where I sit: Singapore, an entire developed country of about six million people, consumes about 58 TWh a year against a peak demand of 7.7 GW.[58] A single company's capacity addition equates to 60% of my country's annual electrical consumption, and half of its peak load.[59] A training cluster wants power at 90%+ uptime, and solar runs at 25% capacity factor.[60] So if you wanted to serve a 1.0 GW campus, you'd need 4.0 GW of panels alongside storage.[61] You can do this or simply sign a nuclear PPA. Now, looking at the financials, all-in cost for AI capacity runs about US$50 billion per GW, of which perhaps US$10-15 billion is shell, land, and power infrastructure with a lifespan of thirty years.[62] And the remaining US$35 billion is chips depreciating over the horizon of five to six years.[63] At that rate, the four hyperscalers get around 14 GW of capacity with their US$700 billion, which is consistent with what is being disclosed.[64] The dollars and the watts are telling the same story through different filings.

Rent a H100 at about US$2 an hour,[65] multiply that by the minimum of what 1 GW can host, which is 650,000 GPUs, times 8,760 hours. That's around US$11 billion a year per gigawatt at full utilization. Let's be conservative and call it US$8 billion at realistic loads.[66] Run that against US$50 billion of capex; that is a 6-year payback before operating costs, on chips with an

---

[54] Arithmetic: 1,000,000 kW ÷ 1.5 kW per GPU ≈ 667,000 GPUs.
[55] A GB200 NVL72 rack draws roughly 120 kW. Bernstein prices such a rack at about US$5.9 million — US$3.4 million of compute hardware and US$2.5 million of physical infrastructure — which holds the per-gigawatt order of magnitude. https://www.investing.com/news/stock-market-news/how-much-does-a-gw-of-data-center-capacity-actually-cost-4314046
[56] Arithmetic: 3,900,000 kW ÷ 1.5 kW = 2.6 million accelerators.
[57] 1 GW × 8,760 hours = 8.76 TWh. 3.9 GW × 8.76 = 34.2 TWh.
[58] Energy Market Authority of Singapore, Singapore Energy Statistics: total electricity consumption rose 4.0% to 58 TWh in 2024. https://www.ema.gov.sg/resources/singapore-energy-statistics/chapter3. Population 6.04 million (Singapore Department of Statistics, 2024). The most recent actualised peak system demand is 8,189 MW, recorded in July 2025; 7.7 GW corresponds to the 2023 peak.
[59] 34.2 ÷ 58 = 59%. 3.9 ÷ 7.7 = 51%.
[60] The US Energy Information Administration reports average capacity factors for utility-scale solar photovoltaics in the 23–25% range.
[61] 1.0 GW ÷ 0.25 = 4.0 GW of nameplate panels, before accounting for storage round-trip losses.
[62] Nvidia has put the all-in figure at US$50–60 billion per gigawatt (fiscal Q2 2026 earnings call). JLL puts fully built-out AI campuses at US$45–55 billion per GW; Bernstein models roughly US$35 billion; Epoch AI models US$38 billion of up-front capital expenditure for a 1 GW site. A "powered shell" — building, land and power infrastructure without chips — runs US$9–11 billion per gigawatt. https://epoch.ai/data-insights/ai-datacenter-cost-breakdown
[63] Epoch AI's model puts servers at about 60% of total cost of ownership, assuming a five-year IT equipment life against fourteen years for the facility; shortening IT life to three years raises annual cost from US$8.5bn to US$12bn. Disclosed server useful lives run five to six years across the hyperscaler Form 10-K filings.
[64] US$700 billion ÷ US$50 billion per GW = 14 GW.
[65] SemiAnalysis H100 one-year rental contract price index: US$1.70 per hour in October 2025, rising roughly 40% to US$2.35 per hour by March 2026, with on-demand capacity effectively sold out. Market median on-demand runs US$2.29–3.12 per hour. https://newsletter.semianalysis.com/p/the-great-gpu-shortage-rental-capacity
[66] Arithmetic: 650,000 × US$2 × 8,760 = US$11.4 billion. At roughly 70% realised utilisation, US$8.0 billion.

accounting lifespan of 5-6 years.[67] Do the maths, and you'll realize the spot prices are approximately breakeven with the depreciation schedule of the chips. But nothing in this industry is sold at spot. The take-or-pay contract is the primary business model; the chip is the melting ice cube, and the contract is the real asset that secures revenue.

If I cross-check that number with a neocloud publishing both numbers: CoreWeave's US$99.4 billion backlog coupled with 3.5 contracted GW is around US$28 billion per GW; divide that by the chip shelf life, and you'll get around US$5.5-6 billion a year on a five-year term schedule.[68] Around the same vicinity as the number I calculated above. Hence, we have US$50 billion per GW in, US$25-30 billion per contracted GW back, and a chip that dies in six years. But now, bring that forward because the finance story becomes a physical one. 14 GW of new AI load in a single year is about 123Twh run continuously, which is more electricity than the Netherlands consumes for everything.[69] Added in twelve months, by four single companies, in a buildout that is only just beginning. American electricity demand grew at less than 0.6% a year across the first two decades of this century, and now that system is demanding 15-25% more electricity by the year 2030, with peak demand rising from about 760 GW towards 850-930 GW.[70]
Furthermore, EPRI (Electric Power Research Institute) projects American data centers to consume 9-17% of national generated power by 2030, up from 4.4% in 2023, which is roughly a 60% increase from the number it published two years prior.[71] It's a mirror image of Part I, where every skeptic's ceiling became the ground floor for the year ahead. And at the top of that range, one country's data centers would draw more electricity than France consumes as a whole.[72] Right now, the electricity for this does not even exist, and money cannot make it appear.

To make it clear, aggregate generation is not the problem. The United States will add a record 86 GW of utility-scale capacity in 2026, which on the surface looks like more than enough to cover the 14 GW of AI load.[73] But wait, solar is 51% of that number, and battery storage is 28%, whilst

[67] US$50 billion ÷ US$8 billion = 6.25 years.
[68] US$99.4 billion ÷ 3.5 GW = US$28.4 billion per GW; ÷ five years = US$5.7 billion per GW per year. CoreWeave's own duration disclosure supports the five-year framing: 75% of backlog is expected to be recognised within four years, with 25% extending beyond forty-eight months.
[69] 14 GW × 8.76 = 122.6 TWh. Netherlands total electricity consumption runs roughly 110–115 TWh a year (International Energy Agency; Eurostat).
[70] RaboResearch, "The US scrambles to meet surging power demand by 2030," July 2026: for years demand grew less than half a percent annually; by 2030 total consumption could climb as much as 20% from roughly 4,300 TWh, with system peak rising from around 760 GW today to between 850 GW and 930 GW.
https://www.rabobank.com/knowledge/d011512857-the-us-scrambles-to-meet-surging-power-demand-by-2030
[71] Electric Power Research Institute, "Powering Intelligence 2026," 26 February 2026: US data centers could consume 9–17% of national electricity generation by 2030, more than double the current 4–5%, with the new estimates running 60% higher than EPRI's 2024 figures. https://powering-intelligence.epri.com/executive-summary.html. The 4.4% baseline for 2023 is from Lawrence Berkeley National Laboratory, 2024 United States Data Center Energy Usage Report — 176 TWh against roughly 4,000 TWh of national consumption.
[72] EPRI's high scenario puts US data-centre consumption at roughly 790 TWh by 2030. France's total annual electricity consumption runs approximately 445–460 TWh (Réseau de Transport d'Électricité, Bilan électrique).
[73] US Energy Information Administration, "New U.S. electric generating capacity expected to reach a record high in 2026," 20 February 2026: 86 GW of planned utility-scale additions, against 53 GW added in 2025 — itself the largest single-year installation since 2002. https://www.eia.gov/todayinenergy/detail.php?id=67205

net new gas capacity comes under 4 GW, and new nuclear comes to zero.[74] I mentioned earlier in this article that a training cluster wants power at 90%+ uptime, and solar runs at 25% capacity factor. Hence, the relevant figure is the 4 GW of newly dispatchable capacity, against an AI buildout wanting about three times more than that. Essentially, the argument is that the added 86 GW of capacity additions cannot and shouldn't be read as 86 GW of firm, 24/7 supply.

Which is why turbine order books are important to watch. GE Vernova, Siemens Energy, and Mitsubishi supply more than 70% of global production, and their order books now hold a figure higher than 170 GW.[75] Combined cycle lead times have extended from three and a half years to five, meaning power infrastructure needs to be planned further in advance, and a data center developer can't simply say I need 1 GW of power next year.[76] Northern Virginia's interconnection queue now runs to seven years.[77] Roughly one third of data centers scheduled to open this year may be delayed or never built.[78] And the binding constraint isn't capital or demand; it's firm power at a specific node, on a schedule that is measured in years.

I hope this illustrates what is genuinely scarce here, and it isn't the GPU. Nvidia will sell that to anyone with the money. The scarce asset here is energized interconnection: a megawatt contracted, built, connected, flowing to the data center, and operationally running. When supply is constrained by the speed of turbine manufacturing and new capacity takes years to arrive, the value of existing energized interconnection and active capacity is determined by the replacement cost of the marginal new unit, not what it's worth historically. That's scarcity rent.

This brings me to a point that I've been meaning to express for a while. Notice which direction the risk runs. The worst-case scenario treats expensive electricity as a cost problem for hyperscalers. But Microsoft holds Three Mile Island on a twenty-year PPA. Amazon has Comanche Peak, and Google bought Intersect Power outright. When power does get dearer, which it will, the incumbent's input cost barely moves while the marginal replacement cost of its

---

[74] Same EIA release: solar 51% (43.4 GW), battery storage 28% (24 GW), wind 14% (11.8 GW). Planned natural gas additions total 6.3 GW gross — 3.3 GW combined-cycle and 2.8 GW combustion turbine — which nets down materially against scheduled gas retirements. No new nuclear capacity is scheduled to enter service in 2026.

[75] GE Vernova disclosed gas power equipment backlog and slot reservation agreements of 116 GW as at Q2 2026, up from 100 GW at Q1, guiding to at least 125 GW by year-end 2026 against a total company backlog of US$176 billion. https://www.sec.gov/Archives/edgar/data/0001996810/000199681026000147/gevpressrelease2q26.htm. Siemens Energy's backlog reached €146 billion at end-2025 with a book-to-bill of 1.82, and Mitsubishi is sold out into 2028. The three together are generally put at 70–75% of global heavy-duty gas turbine supply.

[76] Lead times for a combined-cycle gas plant have moved out sharply; reported ranges run from three and a half to five years, and in some accounts from two or three years out to five to seven. Bloomberg, "Siemens Energy, Mitsubishi Struggle to Keep Up With AI-Driven Demand For Gas Turbines," October 2025. https://www.bloomberg.com/features/2025-bottlenecks-gas-turbines/

[77] JLL, 2026 Global Data Center Outlook: the average wait for a 100 MW connection in Northern Virginia is seven years. Dominion Energy has roughly 70 GW in its large-load interconnection queue and connects on the order of ten large-load customers a year; as of its February 2026 State Corporation Commission filing, about 25 GW had projected connection dates through end-2031 and 45 GW remained under study with no date.

[78] Sightline Climate, April 2026: between 30% and 50% of large data centers scheduled to open in 2026 will be delayed or cancelled, with only about 5 GW of the roughly 16 GW announced pipeline actually under construction and 11 GW showing no sign of building. Wood Mackenzie reached a comparable conclusion, finding only about a third of the 241 GW pipeline under active development.

asset base climbs.[79] So the point that I'm trying to make is that these companies, who are already energizing themselves, are creating a moat for themselves. And somehow no one can see it.

And then the buildouts end, which is the part that almost nobody is modelling. Only the chips recur. The shell, substation, interconnect, US$10-15 billion is not bought twice. Growth will flatten, capex will fall towards the refresh cycle whilst US$5-6 billion per contracted GW will continuously arrive towards an asset base already paid for. What was once a capital sink will begin to generate substantial free cash flow, making free cash flow invert. And the instrument that signals avoid today will print something entirely different.

The metric was never the problem; the phase was. "Ignore the P/E, look at this instead" is the signature move of every mania. In 1999, investors pointed to website eyeballs, an impressive number but one with no guaranteed cash flow behind it.[80] A contracted megawatt is different. If Meta signed a contract for that power, it represented a real customer and a claim on future revenue. In this case, capacity is an indicator of cash flow, but never a substitute. Buffett has described this market as "a church with a casino attached."[81] The casino is definitely real, but so is the church. Buffett also underwrote BNSF railway on enterprise value per mile of track.[82] Valuing an asset based on contracted capacity is not a new trick. It's an old trick in finance.

Part IV — The Casualties, and the Rebuttals

A thesis that only predicts winners is a narrative. A real thesis identifies who gets hurt in the process, names them, and tests itself against the strongest case for the other side. Here, both questions lead to the same place: the counterparty.

Starting in the credit market with Oracle's five-year credit default swaps trading near 198 basis points, the highest on record for the company, whilst the same company reports the strongest booking growth in its history. Record demand on one axis, but record default insurance on the other.[83]

---

[79] The mechanism is that power purchase agreements signed at 2024–26 prices fix the incumbent's input cost for two decades, while replacement cost for a new entrant rises with turbine, interconnection and construction scarcity.

[80] On the non-financial metrics of the 1999–2000 period — "eyeballs," page views, registered users — see the Securities and Exchange Commission's cautionary guidance on internet-company disclosure of that era and the subsequent Financial Accounting Standards Board discussion of non-GAAP performance measures.

[81] Warren Buffett, CNBC interview at the Berkshire Hathaway annual meeting, 2 May 2026, describing markets as a church with a casino attached and adding that people have never been in a more gambling mood. https://fortune.com/2026/05/02/warren-buffett-investing-gambling-mood-one-day-options-prediction-markets-betting-berkshire-hathaway/. He returned to the same framing in a CNBC interview published 15 July 2026.

[82] Berkshire Hathaway acquired the remaining approximately 77.4% of Burlington Northern Santa Fe in February 2010 at a total enterprise value of roughly US$44 billion including assumed debt, against approximately 32,000 route miles — on the order of US$1.4 million per route mile. Berkshire Hathaway 2009 annual letter to shareholders and the BNSF merger proxy.

[83] Oracle's five-year credit default swap closed at 198.18 basis points on 27 March 2026 — the highest on record in data going back to end-2008, surpassing the financial-crisis peak — and set a further record of 198.23 bp on 17 July 2026. Oracle holds roughly US$120 billion of bonds in the Bloomberg US high-grade corporate index, the largest non-bank issuer, and has the most

Concentration. The US$553 billion backlog is committed by essentially three customers, with OpenAI's commitment being near US$300 billion.[84] An RPO is really only as good as its counterparty, and in this case. OpenAI's 2026 annualized revenue sits at US$25 billion, while contracted infrastructure obligations run into the hundreds of billions, if not trillions.[85] This backlog is a credit exposure wearing a growth costume, and the credit market has definitely taken notice of this.

Followed with timing. Oracle spends now with revenue that arrives later. Free cash flow ran to the negative of US$24.7 billion, and the company has raised on the order of US$45-50 billion across 2026;[86] project loans run to US$18 billion in New Mexico and US$38 billion across Texas and Wisconsin.[87] Completion dates have reportedly slipped from 2027 towards 2028,[88] and banks providing financing have begun buying credit default swaps on their own client.[89] Standard risk management, and also the credit market quietly stating its estimate of the tail.

And now the circular financing. OpenAI's revenue grew from US$3.7 billion to US$13.1 billion in the span of two years, which is extraordinary; however, that's against infrastructure commitments reported near US$1.4 trillion;[90] HSBC modeled a funding gap near US$207 billion even on generous assumptions.[91] Nvidia has pledged up to US$100 billion into OpenAI, which uses that capital to buy Nvidia chips.[92] Nvidia also holds equity in CoreWeave and has agreed to

---

liquid investment-grade CDS, averaging over US$830 million of weekly trades. ICE Data Services via Bloomberg. https://www.bloomberg.com/news/articles/2026-03-27/oracle-s-credit-risk-measure-nears-record-high-on-ai-debt-fears

[84] Oracle's US$300 billion, five-year compute agreement with OpenAI, announced September 2025, covering 4.5 GW of capacity.

[85] OpenAI reached a roughly US$25 billion annualised revenue run rate by the end of February 2026, up from US$21.4 billion at the close of 2025 (The Information; Sacra). Disclosed compute commitments: approximately US$300bn with Oracle, US$250bn with Microsoft and US$38bn with AWS. Sam Altman has put total infrastructure commitments at roughly US$1.4 trillion.

[86] Oracle's trailing-twelve-month free cash flow stood at negative US$24.7 billion as of Q3 FY2026, reported 10 March 2026, as capital expenditure rose from US$21.2 billion in FY2025 towards a guided US$50 billion. Across fiscal 2026 Oracle raised US$43 billion of debt and US$5 billion of equity — US$48 billion. https://finance.yahoo.com/news/oracle-blows-investors-away-22-234802787.html

[87] Roughly twenty banks arranged an approximately US$18 billion project finance loan for a New Mexico data centre campus with Oracle as anchor tenant; a further approximately US$38 billion facility was being syndicated for Vantage Data Centers sites in Texas and Wisconsin, structured through special purpose vehicles. Financial Times, November 2025.

[88] Bloomberg reported in December 2025 that Oracle had pushed completion dates for several OpenAI-related data centres from 2027 to 2028, citing labour and material shortages. Oracle publicly denied any delay to sites required to meet contractual commitments. Reuters, 12 December 2025. https://finance.yahoo.com/news/oracle-pushes-back-several-data-161438968.html

[89] Banks arranging Oracle-linked AI financings increased credit default swap protection on Oracle through late 2025 and 2026; Oracle became the market's preferred instrument for hedging AI-related credit exposure. ICE Data Services volumes; Bloomberg reporting from November 2025 onward.

[90] OpenAI booked US$3.7 billion of revenue in 2024 and US$13.07 billion in 2025 (leaked financials reported by Fortune, June 2026), against approximately US$1.4 trillion of announced infrastructure and compute commitments.

[91] HSBC Global Investment Research (Nicolas Cote-Colisson), November 2025: OpenAI would need approximately US$207 billion of new financing by 2030. The model assumes approximately US$792 billion of cloud and AI infrastructure spend between late 2025 and 2030, approximately US$1.4 trillion by 2033, annual revenue of about US$213 billion by 2030 and three billion regular users — that is, the gap persists on generous revenue assumptions. https://www.theregister.com/2025/11/26/openai_funding_gap_hsbc/

[92] Nvidia's letter of intent with OpenAI, announced 22 September 2025, contemplated investment of up to US$100 billion deployed progressively against 10 GW of Nvidia systems.

absorb whatever excess capacity it might have, whilst also selling CoreWeave its chips.[93] About US$92 billion of partner debt has been raised across the ecosystem.[94] It's simple. Investors fund OpenAI, OpenAI signs a contract with Oracle, Oracle borrows against the contract that they have, the backlog appraises both companies, raises both stocks, and higher stocks raise more capital. Every link does make sense and is rational. But the chain as a whole requires capital markets to keep the faith until cash generation catches up. The neoclouds run the same exact structure but with higher leverage and lower quality. CoreWeave's quarterly interest expense doubled to US$536 million, nearly half its adjusted EBITDA, against US$24.9 billion of debt and a net loss widening to US$740 million, with its top two customers at roughly 65% of revenue.[95] Nebius carries around US$10 billion of convertible debt accreting to 120% of principal at maturity.[96] Just to be clear, these are not utilities with rate bases. These are leveraged companies with concentrated customer books. They are financed at a spread, coupled with racing chip depreciation. All of this sets up the rebuttal that actually threatens this essay, and it comes in two parts.

Part one: Dark fiber. The oversupply in the 2000s was a timing failure, not a demand failure. The demand turned out to be real, and eventually all the excess capacity was consumed. The problem was that the fiber built in 1999 was not needed and was not used until 2010, and the companies that built it went bust by 2002.[97] If efficiency keeps getting better, and compute demand for each unit of intelligence decreases, today's gigawatts could become tomorrow's dark fiber.

My rebuttal to this. Fiber was speculative and empty; it earned absolutely nothing while it waited for buyers who did not exist. However, this capacity is take-or-pay, pre-sold before it is even energized. The question is whether the customers who signed the contract are financially intact and capable of paying, or whether there are even customers at all. This is a much more narrow and specific risk. The problem with dark fiber was not merely due to there being too much capacity, but instead there was too much uncontracted capacity that was built. Efficiency is the

[93] Nvidia agreed to purchase up to US$6.3 billion of any CoreWeave capacity left unsold through April 2032 (CoreWeave disclosure, September 2025), holds an equity stake, and closed a further US$2 billion Class A common stock investment in March 2026 while remaining CoreWeave's principal chip supplier.

[94] Financial Times tally, November 2025: approximately US$96 billion — US$30 billion already borrowed by SoftBank, Oracle and CoreWeave, US$28 billion by Blue Owl Capital and Crusoe, and US$38 billion under negotiation for Oracle and Vantage. Separately, Bank of America counted US$121 billion of new debt issued in 2025 by Amazon, Google, Meta, Microsoft and Oracle, roughly four times their five-year average.
https://finance.yahoo.com/news/openai-partners-carrying-96-billion-120300663.html

[95] CoreWeave Q1 2026 Form 8-K, 7 May 2026: net interest expense of US$536 million (against US$264 million a year earlier), adjusted EBITDA of US$1.157 billion (a 56% margin), total current and non-current debt of US$24.859 billion, and a GAAP net loss of US$740 million against US$315 million a year earlier. US$536m ÷ US$1,157m = 46%. Customer concentration is disclosed in the accompanying Form 10-Q; Microsoft alone accounted for 62% of revenue in 2024.

[96] Nebius convertible notes: approximately US$1 billion (June 2025), US$3.1 billion (September 2025) and US$4.34 billion (March 2026), giving roughly US$8.5–10 billion of original principal outstanding; non-current debt more than doubled to US$8.43 billion at 31 March 2026. The 2031 and 2033 notes carry an accretion schedule under which the accreted principal amount reaches 120% of original principal at maturity.
https://www.sec.gov/Archives/edgar/data/1513845/000110465926029863/tm268409d3_ex99-1.htm

[97] Global Crossing filed for Chapter 11 in January 2002; WorldCom followed in July 2002, then the largest bankruptcy in United States history. Contemporary estimates put lit fibre at low single-digit percentages of installed strand mileage.

only risk, and so far any collapse in inference costs has led to expanded consumption, not less compute demand.[98]

Part two: Counterparty collapsing. What if a customer at the center of the order book cannot pay? A client like OpenAI that has provided backlog worth billions of dollars. But let's set up a clear distinction. Counterparty risk is not demand risk; they are two separate things. If OpenAI defaults, it'll be a credit event for Oracle, not a demand event for the world of compute. The gigawatts and computing power don't disappear; they are physically there. Almost like an empty apartment in a booming city. A tenant defaulting in a full city is a landlord's problem; in an empty city it is everyone's. Part II reflects that the city is full. When Microsoft rations compute away from its own products, and AWS has an unserved demand at a US$142 billion run rate, repossessed capacity finds buyers. In this case, Oracle's capacity will easily be re-let.[99]

All this only holds while scarcity persists – and it won't forever. If this thesis is correct, the buildout eventually catches up with demand, capacity becomes abundant, and returns begin to normalize. The failure case is different: hyperscaler capacity surpluses flood the rental market even before the demand arrives, collapsing counterparty and demand risk into one.[100] This thesis is falsified by continuous utilization demands across energized capacity, backlog cancellations over renegotiations, GPU rental prices falling below the cost of capital and staying there, or a continuous lapse of counterparties walking back commitments the way OpenAI walked back in February.[101] None of which is observable currently. None of this is really bearish as well depending on how you read it. Railroads went bankrupt in waves while rails stayed and carried an entire economy that compounded over a century. WorldCom died but fiber carried the internet.[102] The buildout can be real, necessary, and even civilization changing but there will always be several entities that are destroyed by the attempt.

Risk asymmetry highly favors the builders. For a hyperscaler, overbuilding is a survivable accounting event. Just write down capacity, compress returns, and send a little apology letter to shareholders. To which most of them can revert back to their business models that the cash to fund this entire buildout in the first place. Underbuilding, however, is a death sentence. It means

---

[98] The Jevons dynamic applied to inference: each order-of-magnitude fall in cost per token has to date been met with more than proportionate growth in tokens consumed. Alphabet disclosed that its Gemini app processes 22 billion API tokens per minute as of Q2 2026, up from 16 billion the prior quarter.

[99] When the Abilene expansion lapsed, Meta was reported to be in line to lease the unbuilt capacity from developer Crusoe, with Nvidia placing a US$150 million deposit to broker the arrangement. Bloomberg via The Register, 7 March 2026. https://www.theregister.com/2026/03/07/openai_oracle_dc/

[100] Meta's 1 July 2026 disclosure that it would sell surplus compute is a live test of exactly this failure case: CoreWeave fell 10.8% and Nebius 12.4% in a single session, and the Philadelphia Semiconductor Index dropped more than 10% across two sessions.

[101] Oracle and OpenAI ended talks on a roughly 600 MW expansion of the Abilene campus beyond its committed 1.2 GW, after financing negotiations failed and OpenAI's capacity forecasts shifted. The underlying 4.5 GW / US$300 billion agreement remained in place. Bloomberg, reported 6–7 March 2026.

[102] Penn Central's 1970 filing was the largest United States bankruptcy of its time; WorldCom's in 2002 surpassed it. In both cases the physical asset survived the capital structure.

losing the future of infrastructure that compute and demand will grow upon indefinitely. Every CEO in this essay can do that math. In fact, all of them have. Hence, US$700 billion.[103]

Zoom all the way out.

Strip away the tickers, and the filings describe the fourth great network buildout of the industrial era. Railroads for the movement of goods, electrification for power, fiber for information. Each consumed a scandalous share of national capital.[104] Each was called a bubble mid-construction, sometimes correctly as measured by the builders' equity and never correctly as measured by the infrastructure's use. And each forced finance to change its instruments: railroad accounting invented depreciation, utility regulation invented the rate base.[105] The instruments always change last, after the steel is already in the ground.

Part V — What the World Looks Like If This Is Right

Take a step back.

But this time the substrate is computation, and the difference from every predecessor is that it is being built pre-sold, by the most profitable companies in history, off order books visible in audited filings. If Part I's “line” continues even in decelerated form, the defining corporate species of 2030 is something markets have never priced: a utility's asset base, a railroad's capital intensity, software margins on the services running above the metal. There is no comp for that, and no historical multiple to reach for. I know the market will construct one, because it always does, and the repricing has visibly begun at the edges, in credit desks running EV-per-megawatt math and in "contracted gigawatts" migrating quietly into sell-side models.[106] The market is not stupid. It is mid-realization. This essay is simply the argument for finishing the thought.

The metamorphosis is silent because it speaks in watts, backlogs, and depreciation schedules, a language the ticker does not display. But the filings are public, the contracts are signed, the reactors are restarting, and the staircase keeps adding steps. The instruments will catch up. They

[103] Combined 2025 operating cash flow across the four ran on the order of US$550 billion (Alphabet, Amazon, Meta and Microsoft Form 10-K filings) — the internal funding base against which a write-down of deployed capacity would be absorbed.
[104] Railroads absorbed roughly 15–20% of United States gross capital formation in peak decades of the nineteenth century. See Robert Fogel, Railroads and American Economic Growth (Baltimore: Johns Hopkins Press, 1964), and Alfred D. Chandler Jr., The Visible Hand (Cambridge, MA: Harvard University Press, 1977).
[105] Depreciation accounting was formalised through United States railroad practice and the Interstate Commerce Commission’s uniform systems of accounts from 1907. The utility rate base derives from Smyth v. Ames, 169 U.S. 466 (1898), and was recast on an “end result” basis in Federal Power Commission v. Hope Natural Gas Co., 320 U.S. 591 (1944).
[106] Enterprise value per contracted megawatt and “contracted gigawatts” now appear routinely in sell-side and credit work on CoreWeave, Nebius and Oracle. Goldman Sachs’ own AI capital expenditure model is constructed bottom-up from gigawatts, dollars per kW for new power and dollars per MW for data centres rather than from revenue multiples.

always do, one buildout too late for the builders and right on time for whoever read the gigawatts first.

My closing note to you is this:

I believe at the end of this, common intelligence will exist. Common intelligence will be a utility that flows through every household, airport, hospital, factory, university, bank, power grid, government, military and so on. Just like how water, electricity, and internet flows through each of those.

Common intelligence will be embedded in global infrastructure. It will simply become so ubiquitous that we will stop thinking about it as AI and instead infrastructure. The significance of this isn't that we achieve AGI or Superintelligence. It is that common intelligence becomes a fundamental input to economic activity and output. Just like railroads did, just like electricity did, and just like the internet did. Electricity and the internet transformed the world because it did not conform to a single machine, it was an input to every machine.

I believe that intelligence will follow a similar path. And when intelligence becomes a utility, the infrastructure required to produce, transmit, and deliver it becomes equally fundamental. Compute, power, networks, memory, and data centers cease to be merely technology infrastructure; they become the physical foundation of an economy of common intelligence. Thank you for reading.

This work is dedicated to my mentor, Mr. Robert Ludwig, whose guidance, generosity, and belief in me have shaped this work and my thinking.

Thank you to Dr. Anthony Watson of the University of Cambridge and The Kinkaid School for his continuous support with this paper and with my goals, and for being a formative figure in my life.

Thank you to everyone who has supported and helped me throughout this process. Special thanks to Mr. Arvind Khattar for his invaluable guidance and encouragement.

And to my friends, Martin Lim, Jonathan Quek, and Skye Flecker — thank you for your support, conversations, and friendship along the way.

Works Cited

I. Company Filings, Earnings Materials and Corporate Announcements

1. Alphabet Inc. “Alphabet Announces Agreement to Acquire Intersect to Advance U.S. Energy Innovation.” Investor news release, 22 December 2025. https://abc.xyz/investor/news/

2. Alphabet Inc. “Alphabet Announces Fourth Quarter and Fiscal Year 2025 Results.” Form 8-K, Exhibit 99.1, 4 February 2026. https://www.sec.gov/Archives/edgar/data/1652044/000165204426000012/googexhibit991q4 2025.htm

3. Alphabet Inc. Fourth Quarter 2025 Earnings Call. Transcript and prepared remarks of Anat Ashkenazi, 4 February 2026. https://abc.xyz/investor/events/

4. Amazon.com, Inc. “Amazon.com Announces Fourth Quarter Results.” Form 8-K, Exhibit 99.1, 5 February 2026. https://www.sec.gov/Archives/edgar/data/1018724/000101872426000002/amzn-20251231x ex991.htm

5. Constellation Energy Corporation. “Constellation to Launch Crane Clean Energy Center, Restoring Jobs and Carbon-Free Power to the Grid.” Press release, 20 September 2024.

6. CoreWeave, Inc. First Quarter 2026 Results. Form 8-K and accompanying Form 10-Q, 7 May 2026.

7. GE Vernova Inc. Second Quarter 2026 Results. Form 8-K, 22 July 2026. https://www.sec.gov/Archives/edgar/data/0001996810/000199681026000147/gevpressrele ase2q26.htm

8. Jassy, Andy. “2026 Letter to Shareholders.” Amazon.com, Inc., April 2026. https://www.aboutamazon.com/news/company-news/amazon-ceo-andy-jassy-2025-letter-to -shareholders

9. Meta Platforms, Inc. “Meta Reports Fourth Quarter and Full Year 2025 Results.” Press release, 28 January 2026. https://investor.atmeta.com/investor-news/

10. Microsoft Corporation. Fiscal Year 2026 Second Quarter Earnings. Press release and conference call transcript, 28 January 2026. https://www.microsoft.com/en-us/investor/events/fy-2026/earnings-fy-2026-q2

11. Microsoft Corporation. Fiscal Year 2026 Third Quarter Earnings. Press release and conference call transcript, 29 April 2026. https://www.microsoft.com/en-us/investor/events/fy-2026/earnings-fy-2026-q3

12. Nebius Group N.V. Fourth Quarter 2025 Shareholder Letter. Form 6-K, February 2026. https://www.sec.gov/Archives/edgar/data/1513845/

13. Nebius Group N.V. “Nebius Group Announces Pricing of Upsized Private Offering of $4.0 Billion of Convertible Senior Notes.” Form 6-K, 17 March 2026.

14. Nebius Group N.V. “Nebius Reports First Quarter 2026 Financial Results.”

13 May 2026. https://nebius.com/newsroom/

15. NVIDIA Corporation. NVIDIA H100 Tensor Core GPU Datasheet. Santa Clara, CA: NVIDIA, 2023.

16. Oracle Corporation. “Oracle Announces Fiscal Year 2026 Third Quarter Financial Results.” 10 March 2026. https://www.oracle.com/news/announcement/q3fy26-earnings-release-2026-03-10/

17. Oracle Corporation. “Oracle Announces Record Q4 and FY 2026 Results.” 10 June 2026. https://investor.oracle.com/investor-news/

18. Vistra Corp. “Vistra and Meta Announce Agreements to Support Nuclear Plants in PJM and Add New Nuclear Generation to the Grid.” Press release, 9 January 2026. https://investor.vistracorp.com/

19. Vistra Corp. Fourth Quarter and Full-Year 2025 Results. Earnings release and conference call, February 2026.

II. Research, Institutional and Government Sources

20. Belfer Center for Science and International Affairs. “AI, Data Centers, and the U.S. Electric Grid: A Watershed Moment.” Harvard Kennedy School, February 2026.

21. Electric Power Research Institute. Powering Intelligence 2026. Palo Alto, CA: EPRI, 26 February 2026. https://powering-intelligence.epri.com/

22. Energy Market Authority of Singapore. Electricity Demand and Supply Outlook 2025. Singapore: EMA, November 2025.

23. Energy Market Authority of Singapore. Singapore Energy Statistics. Singapore: EMA, 2025. https://www.ema.gov.sg/resources/singapore-energy-statistics

24. Epoch AI. “Total Cost of Ownership of a One-Gigawatt AI Data Center.” Data Insights, 19 May 2026. https://epoch.ai/data-insights/ai-datacenter-cost-breakdown

25. Financial Accounting Standards Board. Accounting Standards Codification 606-10-50-13, Revenue from Contracts with Customers: Disclosure of Remaining Performance Obligations.

26. Goldman Sachs Global Institute. “Tracking Trillions: The Assumptions Shaping the Scale of the AI Build Out.” 2026. https://www.goldmansachs.com/insights/articles/tracking-trillions-the-assumptions-shaping-scale-of-the-ai build-out

27. Goldman Sachs Research. “Private Markets Are Expected to Have a Growing Role in Data Center Financing.” 2026.

28. Goldman Sachs Research (Ryan Hammond). “Consensus 2027 Hyperscaler Capex Estimates Are Too Conservative.” June 2026.

29. HSBC Global Investment Research (Nicolas Cote-Colisson). OpenAI financing outlook. November 2025.

30. International Organization for Standardization. ISO/IEC 30134-2: Information Technology — Data Centres — Key Performance Indicators — Part 2: Power Usage Effectiveness (PUE). Geneva: ISO, 2016.

31. JLL. 2026 Global Data Center Outlook. Chicago: Jones Lang LaSalle, January 2026.

32. Lawrence Berkeley National Laboratory. 2024 United States Data Center Energy Usage Report. Berkeley, CA: LBNL, December 2024.

33. Morgan Stanley Research (Ben Reitz). Hyperscaler capital expenditure model revision. May 2026.

34. RaboResearch. “The US Scrambles to Meet Surging Power Demand by 2030.” Rabobank, July 2026. https://www.rabobank.com/knowledge/d011512857-the-us-scrambles-to-meet-surging-power-demand-by-2030

35. SemiAnalysis. “The Great GPU Shortage — Rental Capacity: Launching Our H100 One-Year Rental Price Index.” April 2026. https://newsletter.semianalysis.com/p/the-great-gpu-shortage-rental-capacity

36. Sightline Climate. United States data center pipeline analysis. April 2026.

37. U.S. Energy Information Administration. “New U.S. Electric Generating Capacity Expected to Reach a Record High in 2026.” Today in Energy, 20 February 2026. https://www.eia.gov/todayinenergy/detail.php?id=67205

38. Wood Mackenzie. United States data center development pipeline. Q1 2026.

III. Press and Periodicals

39. Bloomberg. “Data Centers Face Seven-Year Wait for Power Hookups in Virginia.” 29 August 2024.

40. Bloomberg. “Siemens Energy, Mitsubishi Struggle to Keep Up With AI-Driven Demand for Gas Turbines.” October 2025.

41. Bloomberg. “Oracle’s Credit Risk Measure Nears Record High on AI-Debt Fears.” 27 March 2026.

42. CNBC. “Amazon Stock Falls 8% on $200 Billion Spending Forecast, Earnings Miss.” 5 February 2026. https://www.cnbc.com/2026/02/05/amazon-amzn-q4-earnings-report-2025.html

43. CNBC. “Microsoft Calls for $190 Billion in 2026 Capital Spending on Soaring Memory Prices.” 29 April 2026.

44. CNBC. “CoreWeave Q1 Earnings Report 2026.” 7 May 2026.

45. Financial Times. Analysis of OpenAI partner debt issuance. November 2025.

46. Fortune. “Oracle Blows Investors Away With 22% Hyper Growth — But Cash Flow Crunches to Negative $24.7 Billion.” 10 March 2026.

47. Fortune. “OpenAI’s Partners Are Carrying $96 Billion in Debt.” 1 December 2025.

48. Fortune. “Warren Buffett Says Markets Are Like a Church With a Casino Attached.” 2 May 2026. https://fortune.com/2026/05/02/warren-buffett-investing-gambling-mood-one-day-options-prediction-markets betting-berkshire-hathaway/

49. Reuters. “Oracle Denies Report on OpenAI Data Center Delays.” 12 December 2025.

50. The Register. “HSBC Spies $207B Crater in OpenAI’s Expansion Goals.” 26 November 2025. https://www.theregister.com/2025/11/26/openai_funding_gap_hsbc/

51. The Register. “OpenAI and Oracle Reportedly Abandon TX Stargate Expansion.” 7 March 2026. https://www.theregister.com/2026/03/07/openai_oracle_dc/

52. The Wall Street Journal. “Big Tech Accounting Creates a Blind Spot in the AI Boom.” 2026.

IV. Books, Legal Authorities and Historical Sources

53. Chandler, Alfred D., Jr. The Visible Hand: The Managerial Revolution in American Business. Cambridge, MA: Harvard University Press, 1977.

54. Federal Power Commission v. Hope Natural Gas Co., 320 U.S. 591 (1944).

55. Fogel, Robert William. Railroads and American Economic Growth: Essays in Econometric History. Baltimore: Johns Hopkins Press, 1964.

56. Interstate Commerce Commission. Uniform System of Accounts for Steam Roads. Washington, DC: Government Printing Office, 1907.

57. The Planetary Society. “How Much Did the Apollo Program Cost?” Pasadena, CA. https://www.planetary.org/space-policy/cost-of-apollo

58. Smyth v. Ames, 169 U.S. 466 (1898).