

Compute market review: Q1 2026

Looking back at January to March 2026: GPU rental capacity sold out across generations, and one-year H100 contract prices rose almost 40% from their October low.

Published by	Rillor (Rillor Corporation)
Published	9 October 2026
Period covered	January to March 2026
Series	Compute market review
Pillar	Compute
Web version	rillor.com/insights/compute-market-review-q1-2026

Contents

[01](#) Summary[02](#) Background[03](#) Data and method[04](#) Findings[05](#) Implications[06](#) Limits of this analysis[07](#) Sources

Summary

Looking back at January to March 2026, the public record shows scarcity returning to the GPU rental market for every generation at once, and memory prices adding a new cost to capacity. The large platforms raised capital spending plans, and the capital markets supplied the funding.

- **Contract prices rose and capacity sold out.** SemiAnalysis's one-year H100 contract index rose from \$1.70 per GPU-hour in October 2025 to \$2.35 by March 2026, almost 40% [1]. It reported on-demand capacity "sold out across all GPU types" and H100 renewals at the rates signed two to three years earlier [1].
- **Posted indexes split.** Silicon Data's B200 index ended March 24.4% above its 1 January level, while its hyperscaler H100 index held between \$7.40 and \$7.52 per GPU-hour from 1 March to 20 April [3, 4].
- **Rubin entered production and Blackwell orders continued.** NVIDIA said Rubin was in full production in January, reported data center revenue of \$62.3 billion for its quarter ended 25 January, and announced a partnership with Meta covering millions of Blackwell and Rubin GPUs [5, 6].
- **Memory prices nearly doubled.** First-quarter conventional DRAM contract prices rose 93% to 98% from the prior quarter, and DRAM industry revenue reached \$97 billion [28]. Meta and Microsoft later cited component pricing when raising 2026 spending plans [25, 26].
- **Platforms borrowed; neoclouds refinanced.** Oracle, Alphabet and Amazon sold notes and equity-linked securities across the quarter, including a 100-year sterling bond [15, 16, 17, 18, 19]. CoreWeave closed an \$8.5 billion loan rated A3 by Moody's, priced at SOFR plus 2.25% for its floating tranche [14].
- **Spending plans reset higher.** For 2026, Alphabet guided \$175 billion to \$185 billion, Amazon about \$200 billion and Meta \$115 billion to \$135 billion, against 2025 figures of \$91.4 billion, \$131.8 billion and \$72.2 billion [23, 10, 21].

Background

AI compute reaches users in two ways. Buyers rent GPU capacity by the hour or by contract from hyperscale clouds and from neoclouds, the cloud providers built around GPU capacity. Or they own GPU systems, financed with equity, debt or customer prepayments. Rental prices, hardware prices, memory, power and financing terms are linked: what a GPU earns on rent supports what an owner can pay for it and what a lender will advance against it.

The quarter opened from a turning point. One-year H100 contract prices had bottomed at \$1.70 per GPU-hour in October 2025, according to SemiAnalysis [1]. Silicon Data's posted H100 index had risen 10% in four weeks, from \$2.00 on 9 December 2025 to \$2.20 on 6 January 2026 [2]. Fourth-quarter 2025 conventional DRAM contract prices had already risen 45% to 50% from the prior quarter [32]. NVIDIA's next platform, Vera Rubin, was due from partners in the second half of 2026 [5].

Data and method

Period. 1 January to 31 March 2026. Results reported in late April and May 2026 are used for what they state about January to March, and each is labeled with its fiscal period. NVIDIA's fiscal fourth quarter ended on 25 January 2026 and its fiscal first quarter on 26 April 2026; Microsoft's fiscal third quarter ended on 31 March 2026. Commentary published shortly after the quarter (SemiAnalysis on 2 April, Silicon Data on 27 April, TrendForce on 1 June) is used for what it states about this period.

Sources. Company press releases and SEC filings (Forms 8-K with their exhibits); earnings releases and earnings call webcast pages; NVIDIA's event announcements; grid operator releases; published commentary from Silicon Data and SemiAnalysis on GPU rental prices; TrendForce's published memory price research; and one business publication's report of an earnings call. Every source appears in the numbered list at the end. We opened each source and took figures as stated.

Measures. Seven groups:

1. Rental prices: contract price surveys and posted rental indexes by GPU and tier.
2. Availability: statements about sold-out capacity, lead times and renewals.
3. Accelerator supply: NVIDIA results and platform status, and volume agreements with other chip suppliers.
4. Memory: DRAM contract price forecasts and outcomes, and companies' statements about component costs.
5. Financing: notes, equity programs, equity investments and secured loans.
6. Capital spending: reported and guided capital expenditures of Alphabet, Amazon, Meta and Microsoft.
7. Power: grid operator forecasts and decisions, and statements about capacity constraints.

Treatment of figures. We keep surveyed contract prices, posted rental rates and reported deal values apart, and we keep hyperscaler and neocloud tiers apart where the source does. Index levels are reported as their publishers stated them. Capital expenditure measures differ by company: Alphabet and Amazon report purchases of property and equipment; Meta reports capital expenditures including principal payments on finance leases; Microsoft reports capital expenditures including finance leases. Where we add or subtract reported figures, the text says so.

Relation to the Rillor Compute Index. The Rillor Compute Index methodology v1.0 defines four series: RCI Rental, RCI Availability, RCI Hardware and RCI Rent-to-Value. This review keeps the same separation between rental, availability, hardware and financing questions. It reports third-party figures only and contains no RCI values.

Findings

1. Contract prices rose and capacity sold out

Measure	Value	Source
One-year H100 contract, October 2025 low	\$1.70 per GPU-hour	[1]
One-year H100 contract, late January 2026	Above \$2 per GPU-hour	[1]
Change, end of January to mid-to-late February	Up 15% to 20%	[1]
One-year H100 contract, March 2026	\$2.35 per GPU-hour, almost 40% above the October low	[1]
B200 spot on AWS (p6-b200)	Customers "fighting to pay \$14/hr/GPU"	[1]
Blackwell lead times	Into June and July 2026	[1]

SemiAnalysis launched its one-year H100 rental price index on 2 April 2026. It is built from monthly surveys of more than 100 market participants, including neoclouds, buyers and sellers of compute, reported as a 25th to 75th percentile range, and checked against transaction data and a few transactions SemiAnalysis arranged itself [\[1\]](#). It targets the contract market, which SemiAnalysis says carries most rental volume on terms of six months or longer and which posted prices often follow with a lag [\[1\]](#).

The same report gave direct availability evidence. It said "On-Demand GPU rental capacity is sold out across all GPU types", that by March it was "increasingly impossible to find any H100s, H200s or B200 rental capacity for any term", and that when it asked for 8 nodes (64 GPUs) of H100 or H200, "half the providers we asked were completely sold out" [\[1\]](#). Some large neoclouds no longer sold single nodes, and some renters were subletting capacity [\[1\]](#). Lead times for new Blackwell deployments ran into June and July and, market-wide, "all capacity coming online until August to September 2026 has already been booked" [\[1\]](#).

Renewal terms changed with the market. SemiAnalysis reported that "H100s are getting renewed at the exact same rate they were signed at 2-3 years ago", and that some H100 contracts were being renewed for four years through 2028 [\[1\]](#). It reported that smaller AI developers were increasingly contracting for more than four years and agreeing to prepayments above 20%, which it described as previously atypical for such terms [\[1\]](#). It described long-term offtakes of 50 MW, 100 MW or more, with five years the most common tenor, and said neoclouds use such contracts to arrange matching debt, locking in "a teens project IRR in most cases" [\[1\]](#).

SemiAnalysis tied the demand to multi-agent workloads running at high concurrency, growing use of open models, media generation and fundraising by AI developers [1]. These are its attributions. This review does not test them.

2. Posted indexes: B200 moved, hyperscaler H100 did not

Silicon Data series	Period	Reading	Source
B200 index	1 January	4.40	[3]
B200 index	11 January (year-to-date low)	4.29	[3]
B200 index	February range	4.41 to 4.71	[3]
B200 index	25 March (year-to-date high)	6.11	[3]
B200 index	30 March	5.48, up 24.4% from 1 January	[3]
B200, all sources	March mean and median	\$5.09 and \$4.96 per GPU-hour	[3]
H100 hyperscaler	March	\$7.43 to \$7.52 per GPU-hour	[3]
H100 neocloud	March	\$2.43 to \$2.63 per GPU-hour, about 8% higher over the month	[3]
H100 hyperscaler	1 March to 20 April	\$7.40 to \$7.52; unchanged on 39 of 50 days	[4]

Silicon Data's B200 index was nearly flat in January and February and rose in March: it averaged 4.66 in the first half of the month and 5.53 in the second, crossing 5.0 on 15 March and 6.0 on 23 March [3]. Its March commentary named four drivers: HBM3e contract price increases passing through to cloud rates after a lag of four to six weeks, NVIDIA's GTC conference, a run of frontier model releases that absorbed idle B200 capacity, and agentic inference demand that is persistent and sensitive to latency [3]. It described the late-March move as "B200-specific", with no spillover into H100 demand [3].

The H100 tiers behaved differently from each other. The hyperscaler series moved about 1% in March with a coefficient of variation of 0.5%, while the neocloud series rose about 8% [3]. Across March and April, Silicon Data placed the hyperscaler tier at about three times the neocloud tier [4]. Its April commentary

attributed the hyperscaler tier's stability to enterprise reservations and switching costs [4].

Read together, the two sources show where price discovery happened. Contract prices and neocloud rates rose; hyperscaler list prices for the same GPU did not move. A buyer reading only hyperscaler list prices in this quarter would have seen no change.

3. Rubin entered production and Blackwell orders continued

Date	Event	Source
5 January 2026	NVIDIA says Rubin is "in full production"; Rubin-based products available from partners in the second half of 2026; AWS, Google Cloud, Microsoft, Oracle Cloud Infrastructure, CoreWeave, Lambda, Nebius and Nscale among the first cloud providers to deploy Vera Rubin-based instances in 2026	[5]
24 February 2026	AMD issues Meta a warrant for up to 160 million shares, vesting as Meta buys up to 6 GW-equivalent of AMD Instinct GPUs	[9]
25 February 2026	NVIDIA reports data center revenue of \$62.3 billion for the quarter ended 25 January, up 75%; outlook of \$78.0 billion for the next quarter	[6]
16 to 19 March 2026	NVIDIA GTC: members of its NVIDIA Cloud Partner program have deployed more than 1 million GPUs and more than 1.7 GW of capacity; Azure powers up Vera Rubin NVL72 systems	[7]
20 May 2026	NVIDIA reports data center revenue of \$75.2 billion for the quarter ended 26 April, up 92%	[8]

NVIDIA's fiscal 2026, which ended on 25 January, brought revenue of \$215.9 billion and data center revenue of \$193.7 billion, up 68% [6]. The same release listed a multiyear partnership with Meta covering "millions of NVIDIA Blackwell and Rubin GPUs", CoreWeave's buildout of "more than 5 gigawatts of AI factories by 2030" and a "non-exclusive licensing agreement" with Groq, with a \$13.0 billion Groq line in investing cash flows [6]. Its outlook assumed no data center compute revenue from China [6]. NVIDIA claimed Rubin would deliver up to a 10x reduction in inference token cost compared with Blackwell [5, 6]; that is the vendor's claim.

At GTC, NVIDIA's live coverage reported that its chief executive "now sees at least \$1 trillion in revenue from 2025 through 2027" [7]. Members of its NVIDIA Cloud Partner program had deployed a cumulative total of more than 1 million GPUs and more than 1.7 GW of AI capacity, against 400,000 GPUs and 550 MW at the previous year's conference [7]. AWS said it would deploy more than 1 million NVIDIA GPUs beginning in 2026, and Microsoft said Azure had powered up Vera Rubin NVL72 systems, "which will roll out globally over the coming months" [7].

The quarter ended 26 April, which overlaps February to April, brought data center revenue of \$75.2 billion, with networking revenue of \$14.8 billion, up 199% from a year earlier [8].

Other accelerator suppliers signed volume as well. AMD's warrant to Meta covers up to 160 million shares at \$0.01, with a first tranche vesting on shipment of the first 1 GW-equivalent, share-price thresholds rising to \$600 for the final tranche, and expiry on 23 February 2031 [9]. Amazon said Trainium2 was fully subscribed with 1.4 million chips landed, that nearly all Trainium3 supply was expected to be committed by mid-2026, and that Trainium4 was expected to start delivering in 2027 [10]. In its first-quarter release, Amazon reported landing more than 2.1 million AI chips in the past 12 months, more than half of them Trainium, and said Anthropic would secure up to 5 GW of current and future Trainium generations [11].

4. Memory prices nearly doubled

Date	Statement	Source
2 February 2026	First-quarter conventional DRAM contract price forecast raised to 90% to 95%, from 55% to 60%; server DRAM up about 90%; NAND flash forecast raised to 55% to 60%	[27]
26 February 2026	Fourth-quarter 2025 conventional DRAM contract prices rose 45% to 50%	[32]
2 April 2026	LPDDR5 and DDR5 contract prices "tracking toward ~4x and ~5x year-on-year increases" in the first quarter	[1]
1 June 2026	First-quarter conventional DRAM contract prices rose 93% to 98%; DRAM industry revenue \$97 billion, up 81%	[28]

TrendForce attributed the increase to cloud data centers shifting deployments from AI-specific servers toward general-purpose servers as AI applications move from training to inference, and to cloud buyers' willingness to accept higher prices, which pushed other buyers to follow [28]. In February it described "intense competition among buyers for the limited supply" [27].

The cost reached capacity plans directly. SemiAnalysis reported that server makers' repricing led some operators to "slow-roll or abandon deployments" [1]. Silicon Data linked part of the March B200 rise to HBM3e price increases passed through to cloud rates [3]. In January, Microsoft said "rising memory prices would impact capital expenditures" and that the effect on cloud gross margins would build gradually "as these assets depreciate over six years" [22]. When reporting first-quarter results in April, Meta raised its 2026 capital expenditure range by \$10 billion, citing "our expectations for higher component pricing this year", and Microsoft said its calendar 2026 plan of roughly \$190 billion included about \$25 billion "from the impact of higher component pricing" [25, 26].

5. Platforms borrowed; neoclouds refinanced

Date	Issuer	Instrument	Amount and terms	Source
26 January	CoreWeave	Equity investment by NVIDIA	\$2 billion of Class A shares at \$87.20 per share	[12]
2 to 4 February	Oracle	Senior notes in eight tranches, due 2029 to 2066	\$25 billion	[15]
2 February	Oracle	At-the-market common stock program	Up to \$20 billion	[15]
2 to 5 February	Oracle	6.50% Series D mandatory convertible preferred stock	100 million depositary shares, each a 1/2,000th interest in a \$100,000 preferred share	[16]
13 February	Alphabet	U.S. dollar and sterling senior notes	\$20 billion and £5.5 billion, including £1 billion of 6.125% notes due 2126	[17]
27 February	Amazon and OpenAI	Equity investment and compute agreement	Amazon to invest \$50 billion in OpenAI (\$15 billion initially); AWS agreement expanded by \$100 billion over eight years	[20]
13 March	Amazon	U.S. dollar notes in eleven tranches, due 2028 to 2076	About \$37 billion	[18]
16 March	Amazon	Euro notes in eight tranches, due 2028 to 2064	About €14.5 billion	[19]
31 March	CoreWeave	Non-recourse delayed draw term loan (DDTL 4.0), maturing March 2032	\$8.5 billion; floating tranche at SOFR plus 2.25%, fixed tranche at about 5.9%; rated A3 by Moody's and A (low) by DBRS	[14]

The Oracle preferred shares carry a liquidation preference of about \$5 billion in total (our arithmetic: 100 million depositary shares at \$50 each) and convert into common stock in 2029 [\[16\]](#). Oracle's \$25 billion of notes and its equity program could together raise up to \$45 billion (our arithmetic), and the stated uses of the notes include capital expenditures [\[15\]](#).

Amazon's agreement with OpenAI combined capital and compute. Amazon committed to invest \$50 billion, starting with \$15 billion and the remaining \$35 billion "in the coming months when certain conditions are met" [\[20\]](#). OpenAI committed to consume about 2 GW of Trainium capacity through AWS,

spanning Trainium3 and Trainium4, as part of an expansion of their \$38 billion agreement by \$100 billion over eight years [20]. Amazon later said the Trainium commitment begins ramping in 2027 [11].

CoreWeave's two transactions show the neocloud financing model maturing. NVIDIA's \$2 billion investment came with a plan to "leverage NVIDIA's financial strength to accelerate CoreWeave's procurement of land, power and shell" toward more than 5 GW of AI factories by 2030 [12]. The DDTL 4.0 facility is secured by substantially all assets of one subsidiary; CoreWeave described it as financing secured by HPC infrastructure and an associated customer contract [14]. It was anchored by Blackstone Credit & Insurance, with MUFG and Morgan Stanley as co-structuring agents [14]. Its floating tranche priced at SOFR plus 2.25%; CoreWeave's DDTL 3.0 facility, closed in the third quarter of 2025, had priced at SOFR plus 4% [14, 31].

CoreWeave's year-end figures, reported on 26 February, showed a revenue backlog of \$66.8 billion, more than 850 MW of active power and about 3.1 GW of contracted power [13].

6. Spending plans reset higher

Company	Measure	2025 actual	2026 guidance given in January and February	Quarter to March 2026	Same quarter a year earlier
Alphabet	Purchases of property and equipment	\$91.4 billion	\$175 to \$185 billion	\$35.7 billion	\$17.2 billion
Amazon	Purchases of property and equipment	\$131.8 billion	About \$200 billion	\$44.2 billion	\$25.0 billion
Meta	Capital expenditures including finance lease principal	\$72.2 billion	\$115 to \$135 billion	\$19.8 billion	Not used here
Microsoft	Capital expenditures including finance leases	Fiscal year ends June	Lower in the March quarter than in December	\$31.9 billion	Not used here

Sources: Alphabet [23, 24]; Amazon [10, 11]; Meta [21, 25]; Microsoft [22, 26].

The 2026 guidance from Alphabet, Amazon and Meta, given between 28 January and 5 February, totals \$490 billion to \$520 billion, against \$295.4 billion for the same three companies in 2025 (both our arithmetic) [23, 10, 21]. In April, Microsoft added a calendar 2026 figure of roughly \$190 billion and Meta raised its range to \$125 billion to \$145 billion [26, 25].

The companies described demand ahead of supply. Microsoft reported commercial remaining performance obligations of \$625 billion at 31 December, up 110%, with about 45% from OpenAI, and said "customer demand continues to exceed our supply" [22]. By 31 March, the balance was \$627 billion [26]. Alphabet reported Google Cloud backlog of \$240 billion at the end of 2025 and "over \$460 billion" at the end of March 2026, and said roughly 60% of its capital spending goes to servers [23, 24]. Meta said its 2026 expense growth would be driven mainly by infrastructure costs, including third-party cloud spending [21].

7. Power stayed binding

Date	Source	Statement
14 January 2026	PJM	Summer peak forecast to reach about 222,000 MW in 2036, an increase of nearly 66,000 MW over ten years; near-term forecasts reduced by stricter vetting of data center and large-load adjustments
16 January 2026	PJM Board	Plan to integrate large loads, including bring-your-own-generation and connect-and-manage options with earlier curtailment, and an immediate backstop generation procurement
4 February 2026	Alphabet	Chief executive expects to go through 2026 "in a supply constrained way", citing power, land and supply chain
29 April 2026	Microsoft	"we expect to remain constrained at least through 2026"

Sources: [29, 30, 23, 26].

PJM's January forecast put average summer peak growth over the next ten years at 3.6% a year, against 0.3% in its 2021 forecast [29]. It also lowered several near-term years compared with the 2025 report, after requiring firm commitments for near-term large loads and derating longer-term projects [29]. The board's plan of 16 January listed an accelerated interconnection track for state-sponsored generation and a review of how PJM's markets support investment, alongside the large-load options [30]. Its chair said: "This is not a yes/no to data centers" [30].

Implications

For buyers of compute. The 2025 pattern reversed. Older GPUs became harder to find, renewals held at their original prices, and the newest systems were booked months ahead [1]. Hyperscaler list prices did not reflect any of this [4]. A buyer setting budgets for 2026 capacity needed contract-market evidence by tier and term; list prices were a poor guide in this quarter.

For builders and operators. Memory became a cost line large enough to change capital plans at large buyers [25, 26] and to delay deployments at some operators [1]. Server bills of materials priced at late-2025 memory costs understated what capacity cost to build in early 2026. SemiAnalysis also reported AI developers accepting prepayments above 20% on contracts of more than four years, a source of funding for that cost [1].

For lenders. CoreWeave's DDTL 4.0 facility is a reference point: a non-recourse loan secured by one subsidiary's assets, described by CoreWeave as secured by HPC infrastructure and an associated customer contract, rated investment grade and priced well inside the company's 2025 facilities [14, 31]. At the platform level, the quarter's bond sales extended maturities to 2066, 2076 and 2126 [15, 18, 17].

For researchers. This quarter is a test case for price indexes. Contract surveys, neocloud indexes and hyperscaler list prices for the same GPU moved differently, and a study using only one of them would describe a different market. Memory and component prices also became a measurable input to GPU system costs.

Limits of this analysis

- This review uses public sources only. It contains no transaction-level price data and no data collected by Rillor.
- SemiAnalysis's contract index and Silicon Data's posted indexes describe different markets. Index levels are reported as published, and we did not recompute them. Silicon Data's B200 readings are index levels; its commentary reports the same series in dollars per GPU-hour.
- SemiAnalysis's statements about sold-out capacity come from its own survey of providers and are not independently verified here.
- NVIDIA's performance and cost claims for Rubin are the vendor's claims.
- Deal values that combine investment and compute purchases are not comparable with plain capacity leases.
- Totals of guidance ranges and the Oracle liquidation preference are our arithmetic from reported figures.
- Fiscal periods do not match the calendar quarter. NVIDIA's and Microsoft's periods are labeled where used. Some figures were reported after the quarter ended and are used for what they state about January to March.
- One statement from Alphabet's earnings call comes from a business publication's report of the call [\[23\]](#).
- This is a review of public information. It is not investment advice and makes no forecast of prices.

Sources

1. Daniel Nishball, Jordan Nanos, Cheang Kang Wen and others, The Great GPU Shortage: Rental Capacity, Launching our H100 1 Year Rental Price Index. SemiAnalysis, 2 April 2026. <https://newsletter.semianalysis.com/p/the-great-gpu-shortage-rental-capacity>
2. Carmen Li, H100 Price Spike: Understanding the 10% Surge in GPU Rental Costs. Silicon Data, 9 January 2026. <https://www.silicondata.com/blog/h100-price-spike>
3. Carmen Li, B200 Index Price March 2026 Update: What the Charts, Price Bands Are Telling Us. Silicon Data, 30 March 2026. <https://www.silicondata.com/blog/b200-rental-price-march-2026-update>
4. Carmen Li, H100 Hyperscaler Index, April 2026: A Benchmark in Flat Mode. Silicon Data, 27 April 2026. <https://www.silicondata.com/blog/h100-hyperscaler-index-april-2026>
5. NVIDIA Kicks Off the Next Generation of AI With Rubin: Six New Chips, One Incredible AI Supercomputer. NVIDIA Corporation, 5 January 2026. <https://nvidianews.nvidia.com/news/rubin-platform-ai-supercomputer>
6. NVIDIA Announces Financial Results for Fourth Quarter and Fiscal 2026. NVIDIA Corporation, 25 February 2026. <https://nvidianews.nvidia.com/news/nvidia-announces-financial-results-for-fourth-quarter-and-fiscal-2026>
7. NVIDIA GTC 2026: Live Updates on What's Next in AI. NVIDIA Corporation (blog), 19 March 2026. <https://blogs.nvidia.com/blog/gtc-2026-news/>
8. NVIDIA Announces Financial Results for First Quarter Fiscal 2027. NVIDIA Corporation, 20 May 2026. <https://nvidianews.nvidia.com/news/nvidia-announces-financial-results-for-first-quarter-fiscal-2027>
9. Current Report on Form 8-K (warrant issued to Meta Platforms, event of 23 February 2026). Advanced Micro Devices, Inc., 24 February 2026. <https://www.sec.gov/Archives/edgar/data/2488/00000024882600045/amd-20260223.htm>
10. Amazon.com Announces Fourth Quarter Results. Amazon.com, Inc., 5 February 2026. <https://ir.aboutamazon.com/news-release/news-release-details/2026/Amazon-com-Announces-Fourth-Quarter-Results/default.aspx>
11. Amazon.com Announces First Quarter Results (Exhibit 99.1 to Form 8-K). Amazon.com, Inc., 29 April 2026. <https://www.sec.gov/Archives/edgar/data/1018724/000101872426000012/amzn-20260331xex991.htm>
12. Press release on NVIDIA's investment in CoreWeave and expanded partnership (Exhibit 99.1 to Form 8-K). CoreWeave, Inc., 26 January 2026. https://www.sec.gov/Archives/edgar/data/1769628/000176962826000044/ex991pressrelease_final.htm
13. CoreWeave Reports Strong Fourth Quarter and Fiscal Year 2025 Results (Exhibit 99.1 to Form 8-K). CoreWeave, Inc., 26 February 2026. <https://www.sec.gov/Archives/edgar/data/1769628/000176962826000094/coreweave4q25earningspress.htm>

14. Press release on the closing of the DDTL 4.0 facility (Exhibit 99.1 to Form 8-K). CoreWeave, Inc., 31 March 2026. <https://www.sec.gov/Archives/edgar/data/1769628/000176962826000129/ex991.htm>
15. Current Report on Form 8-K (equity distribution agreement and issuance of \$25 billion of notes). Oracle Corporation, 4 February 2026. <https://www.sec.gov/Archives/edgar/data/1341439/000119312526037840/d23732d8k.htm>
16. Current Report on Form 8-K (6.50% Series D mandatory convertible preferred stock). Oracle Corporation, 5 February 2026. <https://www.sec.gov/Archives/edgar/data/1341439/000119312526039344/d63638d8k.htm>
17. Current Report on Form 8-K (U.S. dollar and sterling senior notes offerings). Alphabet Inc., 13 February 2026. <https://www.sec.gov/Archives/edgar/data/1652044/000119312526051423/d946885d8k.htm>
18. Current Report on Form 8-K (U.S. dollar notes offering). Amazon.com, Inc., 13 March 2026. https://www.sec.gov/Archives/edgar/data/1018724/000110465926027729/tm266670d8_8k.htm
19. Current Report on Form 8-K (euro notes offering). Amazon.com, Inc., 16 March 2026. https://www.sec.gov/Archives/edgar/data/1018724/000110465926028556/tm266670d9_8k.htm
20. Press release on the Amazon and OpenAI strategic partnership (Exhibit 99.1 to Form 8-K). Amazon.com, Inc., 27 February 2026. https://www.sec.gov/Archives/edgar/data/1018724/000110465926021050/tm267374d1_ex99-1.htm
21. Meta Reports Fourth Quarter and Full Year 2025 Results. Meta Platforms, Inc., 28 January 2026. <https://investor.atmeta.com/investor-news/press-release-details/2026/Meta-Reports-Fourth-Quarter-and-Full-Year-2025-Results/default.aspx>
22. Microsoft Fiscal Year 2026 Second Quarter Earnings Conference Call. Microsoft Corporation, 28 January 2026. <https://www.microsoft.com/en-us/investor/events/fy-2026/earnings-fy-2026-q2>
23. Amanda Gerut, Alphabet plans to double capex spending to a possible \$185 billion, but it's keeping CEO Sundar Pichai up at night. Fortune, 4 February 2026. <https://fortune.com/2026/02/04/alphabet-google-ai-spending-supply-constraints/>
24. Alphabet Announces First Quarter 2026 Results (Exhibit 99.1 to Form 8-K). Alphabet Inc., 29 April 2026. <https://www.sec.gov/Archives/edgar/data/1652044/000165204426000043/googexhibit991q12026.htm>
25. Meta Reports First Quarter 2026 Results (Exhibit 99.1 to Form 8-K). Meta Platforms, Inc., 29 April 2026. <https://www.sec.gov/Archives/edgar/data/1326801/000162828026028364/meta-03312026xexhibit991.htm>
26. Microsoft Fiscal Year 2026 Third Quarter Earnings Conference Call. Microsoft Corporation, 29 April 2026. <https://www.microsoft.com/en-us/investor/events/fy-2026/earnings-fy-2026-q3>
27. Memory Price Outlook for 1Q26 Sharply Upgraded; QoQ Increases of All Product Categories to Hit Record Highs, Says TrendForce. TrendForce, 2 February 2026. <https://www.trendforce.com/presscenter/news/20260202-12911.html>

- 28. Rapid Contract Price Surge Drives 1Q26 DRAM Industry Up 81% QoQ, Says TrendForce. TrendForce, 1 June 2026. <https://www.trendforce.com/presscenter/news/20260601-13070.html>
- 29. PJM's Updated 20-Year Forecast Continues To See Significant Long-Term Load Growth. PJM Interconnection, 14 January 2026. <https://insidelines.pjm.com/pjms-updated-20-year-forecast-continues-to-see-significant-long-term-load-growth/>
- 30. PJM Board Outlines Plans To Integrate Large Loads Reliably. PJM Interconnection, 16 January 2026. <https://insidelines.pjm.com/pjm-board-outlines-plans-to-integrate-large-loads-reliably/>
- 31. CoreWeave Reports Strong Third Quarter 2025 Results (Exhibit 99.1 to Form 8-K). CoreWeave, Inc., 10 November 2025. <https://www.sec.gov/Archives/edgar/data/1769628/000176962825000059/coreweave3q25earningspress.htm>
- 32. Price Rally Drives 4Q25 DRAM Revenue Up 29.4%; Samsung Regains No. 1 Market Share, Says TrendForce. TrendForce, 26 February 2026. <https://www.trendforce.com/presscenter/news/20260226-12937.html>

About this report

Published by Rillor on 9 October 2026. AI tools assist research and drafting, and every figure cites its source. Rillor builds agentic AI systems, datasets, market research and compute services.

Cite as: Rillor. Compute market review: Q1 2026. 9 October 2026.
<https://rillor.com/insights/compute-market-review-q1-2026>

Notices

Rillor is not a registered investment adviser or commodity trading advisor and does not provide investment or trading advice. Research, forecasts, data and software described on this site are for research and engineering use. Nothing here is an offer or recommendation to buy or sell any security, commodity interest or digital asset. Past or simulated results do not indicate future results.

Rillor Compute Index series are for research and planning use. They are not offered for settlement of any contract.

NVIDIA, AMD and related product names are trademarks of their owners and are used only to identify products. No affiliation or endorsement is implied.