

[Full research report](#) | [Methods note](#)

Evidence classes in GPU pricing

Evidence about a GPU system's price comes in six classes, from specification to verified transaction, and mixing them misleads valuations and indexes.

Published by	Rillor (Rillor Corporation)
Published	9 October 2026
Pillar	Compute
Web version	rillor.com/insights/evidence-classes-in-gpu-pricing

Contents

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

Summary

- A price observation for a GPU system belongs to one of six evidence classes: specification, listing, quote, indicative price, reported sale or verified transaction. Each class establishes a different fact, and the classes should be recorded and calculated separately.
- In the public sources reviewed for this report, statements about GPU prices were listings, indicative prices and reported sales. Verified transactions require documents held by the parties to a sale, and none of the public sources reached that class.
- One provider's listings moved in opposite directions within seven months. AWS cut on-demand prices for H100-based P5 instances by 44% in June 2025 [5], then raised Capacity Block prices for H200-based instances by about 15% in January 2026 [6]. A rental series that mixes pricing types records changes that come from the mix.
- One buyer's purchase on one date appears in two documents with different GPU counts and different scope. The press release implies about \$54,200 per GPU and the annual report about \$49,400 per GPU, a gap of almost 10% [10, 11].
- Reported purchase figures of about \$45,900 to \$80,200 per GPU, from agreements that in two of three cases state that they include ancillary equipment [11], are a different grade from the chip-level indicative range of \$30,000 to \$40,000 stated in 2024 [8]. Comparing them as if they were the same item produces a false price trend.
- Financial benchmark standards rank transaction data first and ask administrators to publish the mix of input types behind each value [1, 2]. Accounting standards rank valuation inputs by how observable they are [3, 4]. Both give a GPU price series a tested design pattern.
- Book values of server fleets rest on useful-life estimates that large operators revise in opposite directions [12, 13], while GPU fleets back secured debt [14, 16]. A valuation used for lending should state which evidence classes it rests on.

Background

A GPU system is a configured product. An HGX 8-GPU server combines GPUs with CPUs, system memory, network adapters, storage and a chassis, and a rack-scale system such as the GB200 NVL72 adds switch trays, power shelves and liquid cooling. NVIDIA's chief executive made the point himself in March 2024: after giving a \$30,000 to \$40,000 range for a Blackwell GPU, he told CNBC that the cost also involves designing data centers and integrating into other companies' data centers [8].

Prices for these systems matter to several groups at once. Buyers compare quotes. Cloud operators set rental rates against hardware costs. Lenders and lessors size advances against GPU collateral: in April 2024 Lambda secured a GPU-backed financing vehicle of up to \$500 million [14], and in July 2026 Nebius entered a senior secured facility of about \$775 million backed by deployed GPU infrastructure and contracted cash flows [16]. Debt also funds fleet growth: in May 2024 CoreWeave signed a \$7.5 billion debt financing facility to develop its fleet of high-performance compute [15]. Accountants depreciate fleets over estimated useful lives. Index builders try to summarize all of it in a series.

Each of these users needs a number, and the public record offers many numbers that describe different things. Financial markets met the same problem in benchmark design. After enforcement actions over attempted manipulation of major interest rate benchmarks, IOSCO published its Principles for Financial Benchmarks in July 2013 [1]. The 19 principles cover governance, benchmark quality, methodology and accountability. Three of them speak directly to input data: Principle 7 on data sufficiency, Principle 8 on the hierarchy of data inputs and Principle 9 on transparency of each determination [1]. The EU Benchmarks Regulation, adopted on 8 June 2016, wrote a similar preference into law for benchmarks in its scope [2].

Accounting takes a parallel approach. IFRS 13, issued by the International Accounting Standards Board in May 2011, defines fair value as the price that would be received to sell an asset in an orderly transaction between market participants at the measurement date, an exit price [3]. Inputs to a fair value measurement fall into three levels: Level 1 inputs are unadjusted quoted prices in active markets for identical items, Level 2 inputs are other inputs that are directly or indirectly observable, and Level 3 inputs are unobservable inputs that still have to reflect the assumptions market participants would use [4].

None of these frameworks was written for GPU hardware. Together they show that separating evidence by type is established practice in price measurement, and they give a GPU price series a tested pattern to follow.

Data and method

Scope. This report reviews public statements made between March 2024 and July 2026 about the prices of NVIDIA data-center GPUs, GPU systems and GPU cloud capacity, together with the text of the benchmark and accounting standards cited above. It is a methods study built on worked examples. It does not estimate a market price.

Sources. Every statement used here was read in its original document: company filings with the U.S. Securities and Exchange Commission, company announcements, provider pricing pages, regulator and standard-setter texts, and reporting by established technology and business publications. Where a press report relays another party's figure, the record keeps both the publication and the original party.

Fields recorded. For each price statement we recorded:

- who made the statement and their role (maker, seller, buyer, analyst, publication);
- the date of the statement and the date the price refers to, where they differ;
- the item priced: GPU, server, rack, cloud instance-hour or capacity reservation;
- the scope: what the figure includes, such as servers, networking, storage, labor or cabling;
- the unit and the quantity;
- the commitment behind the figure: none, an offer to the public, an offer to a named buyer, a signed agreement or a completed sale;
- the evidence class.

Class assignment. The class follows from the commitment and the documents behind a statement:

If the statement is	It is recorded as
A description of contents with no price	Specification
A price a seller publishes to anyone	Listing
A price a seller offers a named buyer, with terms and an expiry	Quote
A figure from a seller, broker, analyst or executive with no commitment	Indicative price
A statement by a party or a publication that a sale or purchase agreement took place	Reported sale
A completed sale checked against invoice, payment and delivery records	Verified transaction

When a statement could belong to two classes, the record takes the class with the weaker commitment and stores the reason. A forwarded message that might be a quote or an estimate is recorded as an indicative price until the written quote is seen.

Derived figures. Per-GPU figures divide a reported total by the reported GPU count, with no adjustment for scope. Percentage changes compare the before and after prices in the same source. All figures are nominal U.S. dollars as reported.

Findings

1. The public record, classed

Statement	Date	Party	Item and scope	Class
P5 on-demand price cut of 44% [5]	Jun 2025	Provider	Cloud instance-hour, on-demand	Listing
Capacity Block price rise of about 15% for H200 instances in US East (Ohio) [6]	Jan 2026	Provider, via press	Cloud instance-hour, reserved for a future start	Listing
Blackwell GPU at \$30,000 to \$40,000 [8]	Mar 2024	Maker's chief executive	Single GPU, with a caveat that cost includes data center integration	Indicative price
GB200 NVL72 rack at about \$3 million [9]	May 2024	Bank analysts, via a journalist's post, via press	Full rack	Indicative price
About 2,400 B200 and B300 GPUs for about \$130 million [10]	Jul 2025	Buyer	GPUs with servers, storage, labor, racks and cabling	Reported sale
About 1,300 B200 and 1,200 B300 GPUs for about \$123.6 million [11]	Jul 2025	Buyer, in annual report	GPUs with ancillary equipment	Reported sale
4,200 B200 GPUs for about \$192.9 million [11]	Aug 2025	Buyer, in annual report	GPUs; scope not itemized	Reported sale
1,200 B300 GPUs for about \$71.4 million [11]	Aug 2025	Buyer, in annual report	GPUs with ancillary equipment	Reported sale
1,200 GB300 GPUs for about \$96.2 million [11]	Aug 2025	Buyer, in annual report	GPUs with ancillary equipment	Reported sale

Three of the six classes appear. The public record reviewed here contains no quotes, because quotes go to named buyers and are rarely published, and no verified transactions, because those depend on documents the parties hold. The three August 2025 agreements in the annual report were entered into after the fiscal

year end and describe agreements to acquire [11]. In this scheme they remain reported sales, and the record notes that the source describes signed agreements with delivery to follow.

2. Listings: pricing type sets the direction

AWS announced its June 2025 reductions with a note on supply conditions: demand for GPU capacity had outpaced industry-wide supply, making GPUs a scarce resource [5]. The cuts applied to on-demand purchases from 1 June 2025 and to Savings Plan purchases after 4 June, measured against prices on 31 May 2025 [5].

Instance	GPU	On-demand reduction	3-year EC2 Instance Savings Plan reduction
P4d	A100	33%	25%
P4de	A100	33%	25%
P5	H100	44%	45%
P5en	H200	25%	26%

Source: AWS News Blog, 5 June 2025 [5].

Seven months later, Network World reported that AWS had raised prices for EC2 Capacity Blocks for ML by about 15% [6]. Capacity Blocks reserve GPU instances for a future start date, for up to six months, and can be booked up to eight weeks in advance [7]. In US East (Ohio), the effective hourly rate for a p5e.48xlarge instance with eight H200 GPUs rose from \$34.608 to \$39.799, and for a p5en.48xlarge from \$36.184 to \$41.612 [6]. Both changes are 15.0%. An AWS spokesperson said Capacity Block pricing is dynamic and that AWS's commitment not to raise prices on fixed pricing models such as On-Demand and Savings Plans was unchanged [6].

Both observations are listings from one provider about GPUs of the same generations. One pricing type fell and another rose. A rental series that pools pricing types would report a change caused by the mix of observations. A series that keeps on-demand, spot and reserved prices apart, and records the term of each reservation, reports what each pricing type did.

3. Indicative prices: the chain of custody

The two indicative figures in the table travelled differently.

The Blackwell range came from the maker's chief executive in a televised interview, and CNBC reported it the same day with his later clarification that the cost is not only about the chip [8]. The same article gave the H100's price as between \$25,000 and \$40,000 per chip, according to analyst estimates [8]. That comparison is itself a second-hand indicative range.

The \$3 million rack figure passed through three hands. Tom's Hardware attributed it to HSBC analysts, as cited in a social media post by a senior writer at Barron's, and described the figures as "alleged HSBC numbers" and "merely analyst estimates" [9]. The same set of estimates placed a GB200 NVL36 rack at about \$1.8 million [9].

An indicative price is useful when it is recorded with its full chain: the original party, each intermediary and the date at each step. Recorded as a bare number, it looks like market data and can enter a series as if someone had paid it.

4. Reported sales: one buyer, two figures

IREN Limited, a Nasdaq-listed company, announced on 3 July 2025 the purchase of about 2,400 NVIDIA Blackwell GPUs, split as about 1,300 B200 and 1,100 B300, for about \$130 million including fit-out costs [10]. A note to the release defines those costs as servers, storage, labor and ancillary equipment such as InfiniBand, racks and cabling [10]. The company's annual report for the year to 30 June 2025 states that on 3 July 2025 it entered an agreement to purchase about 1,300 B200 and 1,200 B300 GPUs with ancillary equipment for about \$123.6 million [11].

Document	GPUs	Scope	Total	Per GPU (derived)
Press release, 3 Jul 2025 [10]	about 2,400	Servers, storage, labor, racks, cabling	about \$130.0 million	about \$54,200
Annual report, filed 28 Aug 2025 [11]	about 2,500	GPUs with ancillary equipment	about \$123.6 million	about \$49,400

The per-GPU figure from the release is about 9.6% above the figure from the filing. The two documents differ in GPU count and in scope, and both use approximate figures. Neither has to be wrong. They describe different bundles, and a record that keeps the count, the scope and the source of each can carry both without contradiction.

5. Grade and scope: figures that look comparable

The same annual report lists three further purchase agreements. The B300 and GB300 agreements are stated "with ancillary equipment", and the B200 agreement carries no itemized scope [11]. Dividing by GPU count gives a figure per GPU:

Item	Reported total	GPUs	Per GPU (derived)	Class
B200 purchase agreement [11]	about \$192.9 million	4,200	about \$45,900	Reported sale
B300 purchase agreement [11]	about \$71.4 million	1,200	about \$59,500	Reported sale
GB300 purchase agreement [11]	about \$96.2 million	1,200	about \$80,200	Reported sale
GB200 NVL72 rack estimate [9]	about \$3 million	72	about \$41,700	Indicative price
Blackwell GPU, maker's range [8]	\$30,000 to \$40,000	1	\$30,000 to \$40,000	Indicative price

Read as one column, the table suggests that GPU prices rose between 2024 and 2025. That reading is an artifact. The 2024 figures are indicative prices for a chip and for a rack. The 2025 figures are reported purchase agreements for three different products, two of them stated to include ancillary equipment. A comparison needs a common grade: the same system, configuration, condition, delivery terms and scope.

6. Contract terms are a separate kind of evidence

Financing agreements contain numbers that look like prices and record something else. The annual report describes a 36-month lease covering the full purchase price of 1,000 B200 and 1,200 B300 GPUs and certain ancillary equipment, with an initial purchase cost of about \$101.8 million and fixed monthly payments of \$2.8 million after delivery, \$100.2 million in aggregate [11]. At the end of the term the company may buy the GPUs at the lower of their prevailing fair market value and 18% of the initial purchase cost [11]. A separate 24-month lease covers the full purchase price of the 1,200 GB300 GPUs, with fixed monthly payments of \$4.4 million, \$106.4 million in aggregate, and a purchase option at \$1 [11].

The 18% figure, about \$18.3 million on these terms, is a negotiated ceiling on a buyout price. It records what two parties agreed, and it is evidence about the financing. It does not establish what a three-year-old B200 system sells for. The aggregate payments on the GB300 lease come to about 110.6% of the purchase price over 24 months, which reflects the cost of financing. Neither figure belongs in a hardware price series. Both belong in a record of contract terms, linked to the reported sale they finance.

7. Book values rest on estimates

Large operators depreciate servers over useful lives that they set and revise. In its annual report for 2024, Amazon stated that, effective 1 January 2025, it was changing the useful lives of a subset of its servers and networking equipment from six years to five, and attributed the change to an increased pace of technology development, particularly in artificial intelligence and machine learning [12]. It estimated the change would reduce 2025 operating income by about \$0.7 billion, and it recorded about \$920 million of accelerated depreciation and related charges in the fourth quarter of 2024 for the early retirement of certain servers and networking equipment [12].

Meta's annual report for 2024 describes a change in the opposite direction. After an assessment completed in January 2025, Meta extended the estimated useful life of certain servers and network assets to 5.5 years from fiscal year 2025, and expected the change to reduce 2025 depreciation expense by about \$2.9 billion [13].

These are management estimates made for accounting purposes and disclosed in annual reports filed with the SEC. They are not observations of what used hardware sells for. A valuation of GPU collateral that relies on book values rests on an estimate, and it should say so.

8. How the six classes map onto established hierarchies

Evidence class	IOSCO Principle 8 input [1]	EU Benchmarks Regulation term [2]	Closest IFRS 13 analogue [3, 4]
Verified transaction	Reported or observed concluded arm's-length transactions	Transaction data, between unaffiliated counterparties	Observable input; Level 1 only for identical items quoted in an active market
Reported sale	Reported or observed concluded arm's-length transactions	Transaction data, where verifiable	Observable input for similar items (Level 2)
Quote	Firm (executable) bids and offers	Committed quotes	Observable input for similar items (Level 2)
Listing	Offers, firm only where executable	Quotes	Observable input for similar items (Level 2)
Indicative price	Other market information or expert judgment	Estimated prices	Unobservable input (Level 3)
Specification	Not a price input	Not input data	Defines the item measured

The mapping is ours, and the standards do not mention GPUs. It shows that each class has a counterpart in frameworks that regulators and auditors already use. IOSCO adds two points that matter for a thin market. First, the hierarchy is not a rigid order: a confirmed bid or offer can carry more meaning than an outlier transaction, and in a low-liquidity market non-transactional data such as bids and offers can predominate [1]. Second, Principle 9 asks for publication, with each determination, of the size and liquidity of the market assessed and indicative percentages of each type of market data considered [1]. The EU Regulation requires input data to be verifiable and asks administrators to publish guidelines on the types of input data, their priority and the use of expert judgment [2].

Implications

For buyers. Compare a quote with other quotes and reported sales for the same grade: the same base system, the same listed differences in CPU, memory, networking and certification, the same cooling, condition, warranty and delivery terms. A listing for the same GPU tells you what a seller asks. A per-GPU figure from a rack purchase tells you what one buyer paid for a bundle. Neither is a direct check on a quote until the grade matches.

For lenders and lessors. Ask which classes a collateral valuation rests on. A value built from listings measures asking prices. A value built from book values measures an accounting estimate [12, 13]. A value built from reported and verified sales of the same grade measures what buyers paid. Contract terms such as buyout options [11] tell you what parties agreed in a financing, and they belong in the credit file next to the price evidence, kept apart from it. When a facility is backed by both hardware and contracted cash flows, as in the July 2026 facility described by Nebius [16], the price evidence for the hardware is one input among several.

For index builders. Publish the hierarchy, record the class of every observation and report the mix behind each value, as IOSCO Principle 9 asks of financial benchmarks [1]. Keep pricing types apart in rental series: the AWS example shows on-demand and reserved prices moving in opposite directions within months [5, 6]. Keep an administrator's own quotes and agreements out of the series it publishes.

For researchers. Store scope as data. A per-GPU figure without a record of what it includes cannot be compared with anything. Store the chain of custody for every indicative price, including each intermediary. Treat differences between a press release and a filing as information about scope, and record both.

In the Rillor Compute Index. The index methodology v1.0 defines four series: RCI Rental, RCI Availability, RCI Hardware and RCI Rent-to-Value [17]. In RCI Hardware, each observation of a published listing, quote or reported sale of a GPU system is tagged with its evidence class, and the classes stay separate in calculation [17]. A series publishes values only when it meets the methodology's data rules [17].

Limits of this analysis

- The examples are a small set of public statements chosen to illustrate each class. They are not a sample of the market, and no price level or trend should be read from them.
- Reported sales are the parties' own statements. This report did not check them against invoices, payment or delivery records.
- Per-GPU figures are simple divisions of reported totals. They do not separate GPU, server, networking or labor costs, and they inherit the approximations in the source documents.
- The mapping of evidence classes to IOSCO, EU and IFRS categories is an analytical comparison. Those frameworks do not address GPU hardware, and the mapping carries no regulatory or accounting status.
- The accounting changes described here apply to the subsets of assets named in each filing. They are not statements about any specific GPU model.
- The review covers NVIDIA data-center products and U.S. dollar figures. Other makers, currencies and regions were outside its scope.
- Nothing in this report is investment, accounting or legal advice.

Sources

1. Principles for Financial Benchmarks: Final Report (FR07/13). International Organization of Securities Commissions, July 2013. <https://www.iosco.org/library/pubdocs/pdf/ioscopd415.pdf>
2. Regulation (EU) 2016/1011 of the European Parliament and of the Council of 8 June 2016 on indices used as benchmarks in financial instruments and financial contracts or to measure the performance of investment funds. EUR-Lex, 8 June 2016. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R1011>
3. IFRS 13 Fair Value Measurement. IFRS Foundation, retrieved 9 October 2026. <https://www.ifrs.org/issued-standards/list-of-standards/ifrs-13-fair-value-measurement/>
4. 8.3.2 Fair value hierarchy, in Applying IFRS for the real estate industry. PwC Viewpoint, retrieved 9 October 2026. https://viewpoint.pwc.com/dt/gx/en/pwc/industry/industry_INT/industry_INT/real_estate__1_INT/Applying-IFRS-for-the-real-est/Disclosures/IFRS-13-disclosures/8-3-2-Fair-value-hierarchy.html
5. Announcing up to 45% price reduction for Amazon EC2 NVIDIA GPU-accelerated instances (Channy Yun). AWS News Blog, 5 June 2025. <https://aws.amazon.com/blogs/aws/announcing-up-to-45-price-reduction-for-amazon-ec2-nvidia-gpu-accelerated-instances/>
6. AWS hikes prices for EC2 Capacity Blocks amid soaring GPU demand (Nidhi Singal). Network World, 6 January 2026. <https://www.networkworld.com/article/4113150/aws-hikes-prices-for-ec2-capacity-blocks-amid-soaring-gpu-demand.html>
7. Amazon EC2 Capacity Blocks for ML. Amazon Web Services, retrieved 9 October 2026. <https://aws.amazon.com/ec2/capacityblocks/>
8. Nvidia's latest AI chip will cost more than \$30,000, CEO says (Kif Leswing). CNBC, 19 March 2024. <https://www.cnbc.com/2024/03/19/nvidias-blackwell-ai-chip-will-cost-more-than-30000-ceo-says.html>
9. Nvidia's next-gen Blackwell AI Superchips could cost up to \$70,000: fully-equipped server racks reportedly range up to \$3,000,000 or more (Anton Shilov). Tom's Hardware, 14 May 2024. <https://www.tomshardware.com/pc-components/gpus/nvidias-next-gen-blackwell-ai-gpus-to-cost-up-to-dollar70000-fully-equipped-servers-range-up-to-dollar3000000-report>
10. IREN Expands AI Cloud with 2.4k NVIDIA Blackwell GPUs, Exhibit 99.1 to Form 8-K. IREN Limited, 3 July 2025. https://www.sec.gov/Archives/edgar/data/1878848/000114036125024747/ef20051566_ex99-1.htm
11. IREN Limited, Annual Report on Form 10-K for the fiscal year ended June 30, 2025. U.S. Securities and Exchange Commission, filed 28 August 2025. <https://www.sec.gov/Archives/edgar/data/1878848/000187884825000063/iren-20250630.htm>
12. Amazon.com, Inc., Annual Report on Form 10-K for the fiscal year ended December 31, 2024. U.S. Securities and Exchange Commission, filed 7 February 2025. <https://www.sec.gov/Archives/edgar/data/1018724/000101872425000004/amzn-20241231.htm>

- L3. Meta Platforms, Inc., Annual Report on Form 10-K for the fiscal year ended December 31, 2024. U.S. Securities and Exchange Commission, filed 30 January 2025. <https://www.sec.gov/Archives/edgar/data/1326801/000132680125000017/meta-20241231.htm>
- L4. Greenberg Traurig Represents Lambda in Securing \$500M GPU-Backed Facility to Expand Cloud for AI. Greenberg Traurig, 12 April 2024. <https://www.gtlaw.com/en/news/2024/04/press-releases/greenberg-traurig-represents-lambda-in-securing-500m-gpubacked-facility-to-expand-cloud-for-ai>
- L5. CoreWeave Secures \$7.5 Billion Debt Financing Facility led by Blackstone and Magnetar. Blackstone, 17 May 2024. <https://www.blackstone.com/news/press/coreweave-secures-7-5-billion-debt-financing-facility-led-by-blackstone-and-magnetar/>
- L6. Nebius raises \$775 million in first secured debt financing to accelerate global buildout, Exhibit 99.1 to Form 6-K. Nebius Group N.V., 17 July 2026. https://www.sec.gov/Archives/edgar/data/0001513845/000110465926084452/tm2620683d1_ex99-1.htm
- L7. Rillor Compute Index, Methodology v1.0. Rillor Corporation, 9 October 2026. <https://rillor.com/markets/compute-index>

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. Evidence classes in GPU pricing. 9 October 2026.

<https://rillor.com/insights/evidence-classes-in-gpu-pricing>

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.