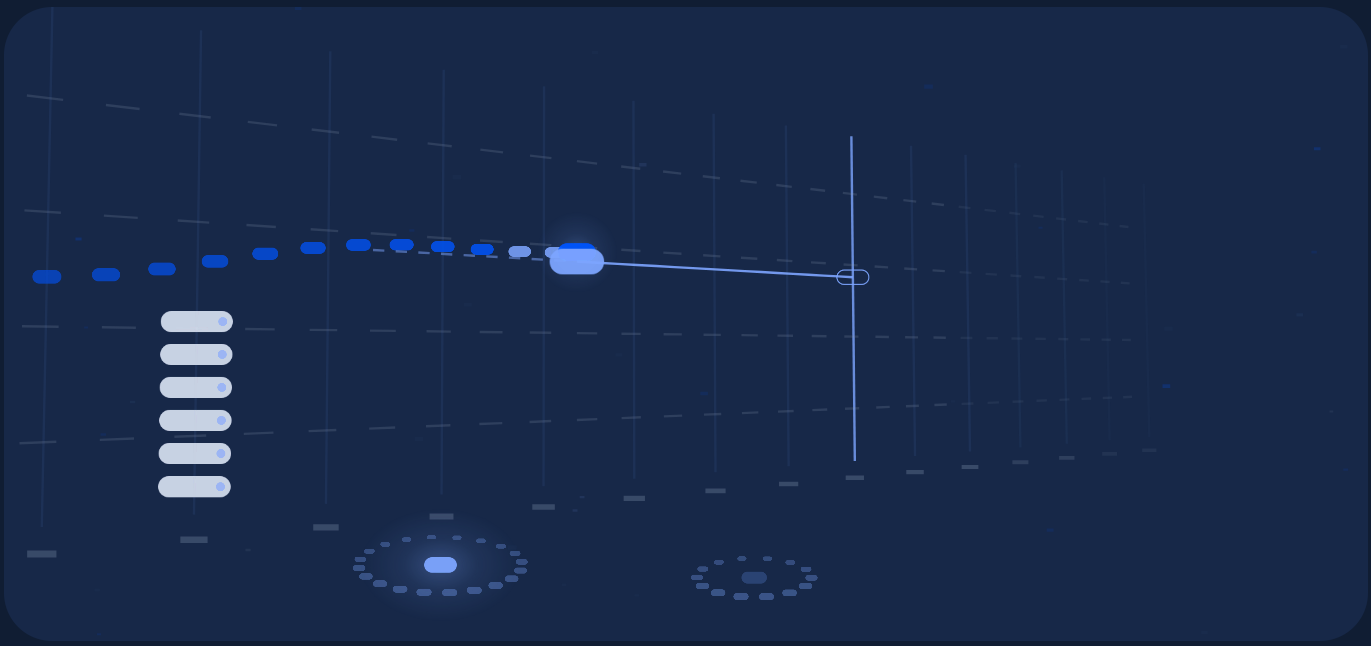
Rillor

EXPLAINER

How to buy GPUs in bulk: allocation, grades and delivery

Bulk GPU purchases are servers and racks bought through system makers and distributors. Allocation, the grade the order names and the delivery terms set what arrives and when.



Rillor

Contents

01 Summary

02 Background

03 Method

04 Findings

05 Implications

06 Limits of this analysis

07 Sources

Summary

Data-center GPUs are bought in bulk as configured servers and racks from system makers and distributors, and the written order decides what arrives and on what terms.

- NVIDIA's annual report for the year to 25 January 2026 names original equipment and device manufacturers, system integrators and distributors among its direct customers, and says its indirect customers primarily buy through system integrators and distributors [1]. Two direct customers accounted for 22% and 14% of its revenue that year [1].
- Allocation has a legal meaning in U.S. sales law. Under the uniform text of the Uniform Commercial Code (UCC), a seller whose capacity is partly cut by an event the contract assumed would not occur must allocate deliveries among its customers in a fair and reasonable manner and tell each buyer its estimated quota [2]. The buyer may then take that quota or end the unexecuted part of the contract [3].
- Bulk supply is ordered far ahead. NVIDIA says its own supply has had extended lead times of more than 12 months, and it has paid premiums and deposits to secure it [1]. CoreWeave, a large buyer of GPU infrastructure, orders on individual purchase orders with lead times its suppliers specify [4].
- The grade, the written description of the system an order names, decides what arrives. One GPU generation ships as an air-cooled 8U Supermicro server with eight B300 GPUs, as an NVIDIA DGX B300 with eight B300 GPUs on AC power supplies or a 54 V DC busbar, or as a fully liquid-cooled GB300 NVL72 rack of 72 Blackwell Ultra GPUs [5, 6, 7].
- The unit of purchase grows with the order: one 8-GPU server, a 4-node Supermicro rack at 60 kW, or an NVIDIA DGX SuperPOD scalable unit of 64 or 72 DGX B300 systems, depending on the power and network design [5, 8, 9]. NVIDIA says to contact it for information on DGX SuperPOD solutions of four scalable units or more [8].
- Delivery terms set when risk of loss passes, where and how the buyer inspects, and whether each tranche is accepted on its own [10, 11, 12]. Under NVIDIA's standard payment terms, payment is generally due shortly after delivery [1], and it has given some investment-grade buyers of large data-center builds terms of 90 days up to one year [13].
- Export screening asks who the buyer is and where the systems run. U.S. Bureau of Industry and Security (BIS) guidance of 13 May 2025 lists red flags such as an unknown installation address and a data center that cannot affirm its power and cooling [14]. Its guidance of 31 May 2026 confirms that a license is required to export advanced computing items to entities headquartered in Country Group D:5 or Macau, or whose ultimate parent is headquartered there, a requirement first introduced on 17 November 2023 [15].

Background

What a bulk GPU purchase is

Data-center GPUs reach buyers inside systems. NVIDIA's DGX B300 is a 10-rack-unit system with eight B300 GPUs of 288 GB each, two Intel Xeon Platinum 6776P processors and 2 TB of system memory by default, up to 4 TB, at a power consumption of 14.5 kW [6]. NVIDIA's GB300 NVL72 is a whole rack: a fully liquid-cooled platform of 72 Blackwell Ultra GPUs and 36 Grace CPUs [7]. Supermicro sells its air-cooled HGX B300 servers as clusters of 4, 8 and 32 nodes [5]. A bulk GPU purchase is therefore a purchase of configured systems, and Supermicro says server prices "can vary widely depending on the configuration", naming CPUs, GPUs, SSDs, cooling systems and memory [16]. Retail graphics cards bought in quantity are outside this report's scope.

Orders, deposits and supply in 2026

Dell said on 1 September 2026 that in its quarter to 31 July 2026 it booked $60.9 billion in orders and ended with a $95 billion backlog in the servers it reports as AI-Optimized Servers, and it raised its full-year guidance for that line to $74.0 billion from $60.0 billion [17]. Buyers are paying earlier. NVIDIA held $2.8 billion of customer advances at 26 July 2026, against $160 million at 25 January 2026 [13]. Supermicro's customer deposits stood at $702.0 million at 30 September 2026, against $314.7 million at 30 June 2026 [16].

Supply is committed too. NVIDIA's supply commitments rose from $119 billion to $279 billion in the quarter to 26 July 2026, primarily for memory and manufacturing facilities [13, 18]. Supermicro said supply constraints for memory, storage, GPUs and CPUs affected the timing of some of its deliveries in its fiscal year to 30 June 2026 [19], and Dell cites substantial inflation in memory component costs [20].

Who this report is for

This report is for procurement leads, infrastructure teams and finance staff planning purchases of GPU servers and racks. It answers three questions a bulk buyer meets at once: who sells, how short supply is shared and what the contract must say.

Method

What was examined. Sales channels, allocation, grades and units of purchase, delivery and acceptance, payment and deposits, warranty and condition records, and export screening for data-center GPU systems.

Sources. Filings with the U.S. Securities and Exchange Commission by NVIDIA [1, 13, 18], Super Micro Computer [16, 19], CoreWeave [4, 21, 36], Dell Technologies [17, 20] and SANGRIX, formerly Bit Origin [22, 23, 37]. Article 2 of the UCC as published by the Legal Information Institute at Cornell Law School [2, 3, 10, 11, 12, 29, 30, 31, 32, 33], with the International Chamber of Commerce (ICC) on Incoterms® [24], UNCITRAL on the sales convention [25] and Wex on escrow [26]. NVIDIA and Supermicro product documentation [5, 6, 7, 8, 9, 28, 34, 35], BIS guidance [14, 15] and the Consolidated Screening List [27].

How. Each source was opened on the date its entry gives, 8 or 14 November 2026, and each quoted passage found in its text. Filing dates were checked on EDGAR index pages. Figures marked "derived" are arithmetic on a source's own numbers. Statements about future periods carry the date they were made. Public companies appear only as publishers of their filings and documentation, and the seller named in the SANGRIX releases is left unnamed.

Legal content. The UCC text quoted is the uniform text. Each state enacts it with variations, contracts can change many of its defaults, and a cross-border sale may fall under the United Nations Convention on Contracts for the International Sale of Goods (CISG) [25]. This report is general information, not legal advice.

Findings

1. Who sells GPUs in bulk

NVIDIA's annual report for the year to 25 January 2026 says: "Some of our direct customers include original equipment manufacturers, or OEMs, original device manufacturers, or ODMs, system integrators and distributors which we partner with to help bring our products to market" [1]. Its list of direct customers also includes add-in board makers, cloud service providers and model makers, and some direct customers may use third-party system integrators to complete their build [1]. Indirect customers, including neocloud builders, enterprises and public sector entities, buy through direct customers: "Indirect customers primarily purchase our products through system integrators and distributors" [1]. Many buy "through multiple OEMs, ODMs, system integrators, distributors, and other channel partners" [1]. In fiscal 2026, sales to one direct customer were 22% of NVIDIA's total revenue and sales to another 14% [1].

One large buyer downstream uses several sources at once. CoreWeave maintains relationships with semiconductor manufacturers, OEMs, ODMs and software providers to source its infrastructure [21]. Three suppliers accounted for 23%, 20% and 17% of its purchases in 2025, and all of the GPUs in its infrastructure are NVIDIA GPUs [4]. System builders are concentrated as well: one supplier accounted for 63.1% of Supermicro's total purchases in its fiscal year to 30 June 2026, and its largest supplier accounted for 64.4% in fiscal 2025 and 65.4% in fiscal 2024 [19].

Can you buy GPUs directly from NVIDIA? The annual report counts cloud service providers and model makers among direct customers [1]. NVIDIA's DGX SuperPOD reference architectures for DGX B300 both say: "Contact NVIDIA for information regarding DGX SuperPOD solutions of four scalable units or more" [8, 9]. Four scalable units are 256 DGX B300 systems in the busbar design [8], or 288 in the AC design (derived, 4 x 72 from [9]). For indirect customers, the annual report names system integrators and distributors as the main route [1].

2. What allocation means in filings and in sales law

In the filings. NVIDIA reports broad requests for information from competition regulators in the European Union, the United States, the United Kingdom, China and South Korea about its sales of GPUs and "our efforts to allocate supply" [1, 13]. The filings do not say how supply is allocated. Supermicro names one factor: "In the event of shortages, some of our larger competitors may have greater abilities to obtain materials and key components due to their larger purchasing power" [19]. It often buys components on a purchase order basis, without long-term supply contracts for all of them [19].

CoreWeave describes the buyer's side. Its suppliers fill its needs "on the basis of individual purchase orders, which we often place with lead times specified by the specific supplier", and it has no control over those lead times [4]. It has a limited number of long-term supplier contracts that guarantee capacity, so once a purchase order is complete its suppliers are generally not obligated to keep supplying it, and prices can rise

on short notice [4]. Among its risks it lists "asymmetry between component availability and contractual performance obligations, including where specified components are required" [4]. NVIDIA warns that publicly announced purchase intentions "are often non-binding and may not result in committed volumes" [1].

In sales law. Section 2-615 of the UCC excuses a seller's delay or non-delivery when performance has been made impracticable by a contingency whose non-occurrence was a basic assumption of the contract, or by good-faith compliance with a governmental regulation or order [2]. Where such causes affect only part of the seller's capacity, the seller "must allocate production and deliveries among his customers", may include regular customers not then under contract and its own requirements, and may allocate "in any manner which is fair and reasonable" [2]. It must notify the buyer seasonably and, where allocation is required, state "the estimated quota thus made available for the buyer" [2].

Section 2-616 gives the buyer the choice. By written notification the buyer may terminate the unexecuted portion of the contract or modify it by agreeing to take the available quota [3]. If the buyer does not modify the contract within a reasonable time not exceeding thirty days, the contract lapses for the deliveries affected [3]. Both sections give way where the seller has assumed a greater obligation [2, 3]. They set a legal rule for an excused shortfall and describe no chipmaker's or system builder's practice.

Buying a place in the queue. On 29 June 2026, Bit Origin, now SANGRIX, announced it had acquired sixteen NVIDIA B300 servers "that have already been purchased by the seller and are currently expected to be delivered during the third quarter of 2026" [22]. The price was US$1 million in cash and US$10 million in equity [22], so the filing gives no cash price per server. On 16 October 2026, SANGRIX said the 16 servers, 128 B300 GPUs in all, were scheduled for delivery in the week of 3 November 2026, updating the third-quarter expectation, with timing subject to the manufacturer's logistics schedule [23]. Its filings through 14 November 2026 report no delivery [37].

3. Lead times and product transitions

NVIDIA says it has experienced, and may continue to experience, "extended lead times of more than 12 months", and that to secure supply it has paid premiums, provided deposits and entered into long-term supply agreements and capacity commitments [1]. In periods of growth it may place non-cancellable inventory orders for some components ahead of its historical lead times [1]. On 26 August 2026 it said inventory had risen to $31.6 billion from $25.8 billion "as we prepare for the introduction of Vera Rubin in the third quarter" of its fiscal 2027 [18]. Dell says the scale of AI opportunities, the varying stages of customer readiness and frequent component transitions make the timing of demand and shipments for its AI-optimized server offerings non-linear [20]. It also leaves cancelable contracts with no substantive termination penalty out of its remaining performance obligations, about $132 billion company-wide at 31 July 2026 [20], so a backlog and a firm obligation are different measures.

Week counts and the supply stages behind them are in Rillor's report on GPU system lead times (https://rillor.com/insights/gpu-system-lead-times-2026).

4. Writing the grade

A grade is the written description of what the buyer receives. On Rillor's forward agreements page, a system grade is a reference grade plus each listed difference in CPU, memory, networking and certification, and air-cooled and liquid-cooled builds are separate grades. Vendor documentation shows how far builds of one GPU generation differ.

Build	GPUs	Cooling and power	Source
Supermicro SYS-822GS-NB3RT, HGX B300	8 B300 at 1,100 W TDP each, up to 2.3 TB HBM3e	Air-cooled, 8U, 15 kW maximum node draw at full load	[5]
NVIDIA DGX B300	8 B300 with 288 GB each	AC power supplies at 200 to 240 V AC, or a 54 V DC busbar, 10 RU, 14.5 kW	[6]
NVIDIA GB300 NVL72	72 Blackwell Ultra GPUs with 36 Grace CPUs	Fully liquid-cooled rack	[7]

Table 1. One GPU generation in three builds, from the makers' documentation.

Power design. NVIDIA's busbar SuperPOD design for DGX B300 uses MGX racks at more than 50 kW each, and a data center without DC busbar can still build a SuperPOD with PDU and AC powered EIA racks [8]. The AC design fits four DGX B300 in a rack at about 56 kW, and NVIDIA notes that some data centers may need fewer per rack [9]. Supermicro gives 60 kW per 4-node rack for its air-cooled clusters [5]. A grade that names the GPU alone leaves cooling, power input and rack density open.

CPU, memory and price. The DGX B300 ships with 2 TB of system memory by default and takes up to 4 TB [6], so "DGX B300" alone leaves a choice open. Supermicro says configuration drives price [16], and it has at times accepted orders with component pricing protection [19].

Certification. NVIDIA-Certified Systems pass a suite of functional, security and performance tests, including single-GPU, multi-GPU and multi-node tests [28]. A grade can list certification as a difference. What the DGX, HGX and MGX names change in a quote is set out in DGX, HGX and MGX explained (https://rillor.com/insights/dgx-hgx-mgx-explained).

5. The unit of purchase

Unit	Contents	Power or cooling	Source
One server	8 GPUs, as a DGX B300 or a Supermicro 8U HGX B300 system	14.5 kW, or 15 kW maximum at full load	[5, 6]

Unit	Contents	Power or cooling	Source
Supermicro clusters	4, 8 or 32 nodes, with 32, 64 or 256 GPUs	60 kW per 4-node rack	[5]
NVIDIA GB300 NVL72	One rack of 72 GPUs and 36 CPUs	Fully liquid-cooled	[7]
SuperPOD scalable unit, busbar	64 DGX B300, 512 GPUs, four per rack	Above 50 kW per rack	[8]
SuperPOD scalable unit, AC	72 DGX B300, 576 GPUs, four per rack	About 56 kW per rack	[9]

Table 2. Units of purchase in vendor documentation.

At four systems per rack, a busbar scalable unit fills 16 compute racks and an AC scalable unit 18 (derived from [8, 9]), before network and management equipment. Supermicro's datasheet states U.S. production capacity "exceeding 5,000 racks per month" and claims "Complete L12 system and cluster-level validation before shipment" [5], both Supermicro's own statements. Its annual report recognizes rack installation and integration services revenue over time [19].

The unit in law. The UCC defines a "commercial unit" as a unit that commercial usage treats as a single whole for purposes of sale, whose division materially impairs its value, and it may be a single machine, a set of articles or any unit treated as a whole in the relevant market [29]. On an improper delivery the buyer may accept any commercial unit and reject the rest [30], and accepting part of a commercial unit accepts the whole unit [31]. A "lot" is "a parcel or a single article which is the subject matter of a separate sale or delivery" [29]. Goods not both existing and identified are "future" goods, and a present sale of them operates as a contract to sell [29].

6. Delivery terms

Weight and site. A packaged DGX B300 weighs 207.6 kg with AC power supplies and 162.6 kg in the busbar version, in a package of 1065 x 711 x 1178 mm [6]. A busbar scalable unit of 64 systems is about 10.4 tonnes of packaged systems, and an AC unit of 72 about 14.9 tonnes (derived from [6, 8, 9]), before racks and network equipment. Under the UCC's tender rule the seller holds conforming goods at the buyer's disposition with the notice needed to take delivery, and unless otherwise agreed "the buyer must furnish facilities reasonably suited to the receipt of the goods" [32].

Risk of loss. In a shipment contract the risk passes when the goods are duly delivered to the carrier, in a destination contract when they are duly tendered there, and otherwise on the buyer's receipt if the seller is a merchant, all subject to contrary agreement [10].

Inspection and acceptance. The buyer may inspect before payment or acceptance at any reasonable place and time, after arrival when goods are shipped [11]. Inspection costs fall on the buyer but may be recovered from the seller if the goods do not conform and are rejected, and a place or method the parties fix is presumed exclusive [11]. Acceptance occurs when the buyer, after a reasonable opportunity to inspect, signifies that the goods conform or fails to reject them effectively [31]. SANGRIX said on 16 October 2026 that inspection, racking and testing after delivery were expected to take about one more week, with production deployment after commissioning and acceptance [23].

Tranches. An installment contract requires or authorizes delivery in separate lots to be separately accepted [12]. The buyer may reject a non-conforming installment that substantially impairs its value and cannot be cured, and a default that substantially impairs the whole contract is a breach of the whole [12].

Cross-border sales. The ICC's Incoterms® rules are eleven three-letter trade terms that clarify the tasks, costs and risks of delivering goods, and the 2020 edition took effect on 1 January 2020 [24]. The CISG, adopted on 11 April 1980 and in force since 1 January 1988, applies between parties whose places of business are in different Contracting States, and may also apply through private international law or by the parties' choice [25].

Readiness. Dell ties shipment timing in part to the varying stages of customer readiness [20]. A delivery date needs a site, power and cooling ready to receive the systems.

7. Payment, deposits and financing

Payment terms. NVIDIA's standard terms make payment generally due shortly after delivery [1]. For some investment-grade customer purchases it has provided longer terms of 90 days up to one year to assist with large data-center builds, depending on size [13]. On 26 August 2026 it put accounts receivable at $63.1 billion and days sales outstanding at 60, up from 45 the quarter before, because of extended payment terms on large, multi-quarter agreements with certain investment-grade customers [18]. Apart from contractually limited return rights for certain stocking distributors, it does not allow returns, although it may approve them based on the facts [1].

Advances and deposits. NVIDIA's deferred revenue additions in the first half of fiscal 2027 included $15.6 billion of customer advances, against $7.5 billion in the same half a year earlier [13]. Supermicro's customer deposits more than doubled in one quarter, to $702.0 million from $314.7 million [16]. Under UCC section 2-718, liquidated damages are allowed only at an amount reasonable in light of the anticipated or actual harm, and an unreasonably large amount is void as a penalty [33]. Where the seller justifiably withholds delivery because of the buyer's breach, the buyer may recover what it paid beyond the liquidated damages or, without such a term, beyond "twenty per cent of the value of the total performance for which the buyer is obligated under the contract or $500, whichever is smaller" [33]. On any order worth more than $2,500, that default is $500 (derived from [33]). The buyer's restitution is subject to offset to the extent the seller establishes a right to damages under other provisions of Article 2, and the value of any benefits the buyer received under the contract [33].

Escrow. Escrow places money or other assets with a neutral third party, the escrow agent, who releases them only when the conditions in the escrow agreement are satisfied [26].

OEM financing. CoreWeave financed equipment through agreements with original equipment manufacturers, secured on that equipment, generally for one to three years, with an aggregate notional balance of $4.8 billion at 31 December 2025 against $1.3 billion a year earlier [21]. The outstanding balance was $7.4 billion at 30 September 2026, against $3.8 billion at 31 December 2025 [36].

8. Warranty and condition

Warranty. Supermicro describes warranties typically ranging from one to three years, with extension options, for products sold by its direct sales team and approved indirect channel partners [19]. Its published limited warranty lists GPU systems at three years of labor, one year of parts and one year of advance cross-ship, from the date of Supermicro's invoice [34]. Those terms apply in the absence of special terms in the invoice, cover only the original purchaser and exclude third-party components [34]. A return needs the invoice date and number, the model name, the serial number and the reason [34]. These are one OEM's published defaults. Rillor's GPU systems page says each quote states the warranty that applies and who provides it, together with the condition grade (new, surplus, refurbished or used) and the testing performed.

Condition records. NVIDIA's nvidia-smi documentation says aggregate error counts "persist indefinitely and thus act as a lifetime counter", and that clearing them is not supported on Ampere and later GPUs [35]. Several GPUs on one board may share a board serial, so it recommends the UUID or PCI bus ID [35]. For used or refurbished lots, these give a buyer a unit-by-unit record. On Rillor's GPU systems page, the per-unit report lists serial numbers for the system and each GPU, the configuration as received, firmware versions, the burn-in and diagnostic results, physical condition and the condition grade.

9. Export screening

This section is general information as of 8 November 2026.

Red flags. BIS guidance of 13 May 2025 on preventing the diversion of advanced computing chips covers ECCNs 3A090.a and 4A090.a and .z items, such as servers classified as ECCN 5A992.z [14]. It lists eleven red flags and seven due diligence actions (counted on the document) [14]. The red flags include a residential address with no other place of use where the quantity is inconsistent with personal use, an unknown installation address, consignees such as freight forwarders that do not normally need large quantities of such chips, and a data center that cannot affirm its power, cooling and space [14].

Due diligence. The guidance asks exporters to seek an end-user certification covering all parties and the intended end use, and a written attestation from the data center that the end user may operate there and that it has the infrastructure for the servers [14]. It also suggests on-site visits at the data center, and says data centers of 10 megawatts or more merit additional scrutiny [14].

Headquarters. BIS guidance of 31 May 2026 says a license is required to export advanced computing items to entities headquartered in Country Group D:5 or Macau, or with an ultimate parent headquartered there, wherever the entities are located, a requirement first introduced on 17 November 2023 [15]. It recalls that BIS announced in May 2025 it would not enforce the AI Diffusion Rule's new compliance requirements [15]. Bona fide data-center operators acting consistently with the regulations need not stop the ongoing use, storage, disposal or servicing of advanced computing items because of the guidance, until further notice from BIS [15].

Screening lists. The Consolidated Screening List names parties subject to U.S. export restrictions, and a possible match calls for additional due diligence before proceeding [27]. Its tools update every day at 5:00 AM Eastern time [27].

Rillor's export notice says Rillor screens every buyer, consignee and end user and requires end-use statements, and that orders may be refused or delayed for export review.

10. A buyer's checklist

1. Know which channel the seller sits in: direct customer, OEM, ODM, integrator or distributor (Finding 1).
2. Ask what the contract does if supply is cut: notice, quota and the right to terminate (Finding 2).
3. Treat announced volumes and backlog as different from firm orders (Findings 2 and 3).
4. Write the grade as a reference system plus each difference in CPU, memory, networking, certification, cooling and power (Finding 4).
5. Name the unit of sale and acceptance (Finding 5).
6. Set when risk passes, the inspection place and method, and whether tranches are accepted separately (Finding 6).
7. Plan the site for weight, rack power and cooling before the delivery date (Findings 5 and 6).
8. Pair any deposit with a liquidated damages term and an escrow decision (Finding 7).
9. Ask who holds the warranty and request per-GPU records for used units (Finding 8).
0. Prepare end-user, end-use and data-center information for export screening (Finding 9).

Implications

For procurement leads. The filings show allocation without saying how it works [1, 4, 19]. The contract names the grade, the unit, the schedule and the shortfall terms. The UCC's quota rules give way to a greater obligation the seller takes on [2, 3], so a buyer who wants more writes it in.

For infrastructure teams. A DGX B300 scalable unit is 16 or 18 compute racks at more than 50 kW or about 56 kW each, and 10 to 15 tonnes of packaged systems (derived from [6, 8, 9]). Air-cooled, busbar and liquid-cooled builds need different racks and plant [5, 6, 7].

For finance teams. Deposits and advances rose sharply in 2026 [13, 16], and long payment terms reached some investment-grade buyers [13, 18]. Without a liquidated damages term, the default rule lets a seller keep the smaller of 20% of the total price or $500 of a deposit after a buyer's breach, plus any damages it proves [33].

On Rillor's forward agreements page, a forward purchase agreement fixes the grade, quantity and price of GPU systems now, with physical delivery in a named month. The buyer pays a deposit to an independent escrow agent at signing, and the seller posts a performance bond.

Limits of this analysis

- No source reviewed states how any supplier decides allocation or what share each buyer receives.
- No current lead time in weeks was re-verified for this report.
- No source reviewed states volume discounts, minimum order quantities, bulk prices per GPU or typical deposit percentages. The SANGRIX consideration was US$10 million in equity and US$1 million in cash [22].
- The UCC text is the uniform text, enacted with state variations. Individual Incoterms® rules and CISG articles are not described.
- Warranty terms are Supermicro's published defaults only.
- Supermicro's capacity and validation statements were not tested.
- Export content reflects BIS guidance of 13 May 2025 and 31 May 2026 [14, 15]. Later rules were not searched for.
- SANGRIX's delivery is known only as the expectation of 16 October 2026 [23], and its filings through 14 November 2026 report no delivery [37].

Sources

1. NVIDIA Corporation Annual Report on Form 10-K for the fiscal year ended 25 January 2026. U.S. Securities and Exchange Commission, filed 25 February 2026, viewed 8 November 2026. https://www.sec.gov/Archives/edgar/data/1045810/000104581026000021/nvda-20260125.htm

2. U.C.C. § 2-615, Excuse by Failure of Presupposed Conditions. Legal Information Institute, Cornell Law School, viewed 8 November 2026. https://www.law.cornell.edu/ucc/2/2-615

3. U.C.C. § 2-616, Procedure on Notice Claiming Excuse. Legal Information Institute, Cornell Law School, viewed 8 November 2026. https://www.law.cornell.edu/ucc/2/2-616

4. CoreWeave, Inc. Quarterly Report on Form 10-Q for the quarter ended 30 June 2026. U.S. Securities and Exchange Commission, filed 12 August 2026, viewed 8 November 2026. https://www.sec.gov/Archives/edgar/data/1769628/000176962826000366/crwv-20260630.htm

5. Supermicro NVIDIA AI Factories HGX B300 datasheet. Super Micro Computer, Inc., 2026, viewed 8 November 2026. https://www.supermicro.com/datasheet/Datasheet_Supermicro_NVIDIA_AI_Factories_HGX_B300.pdf

6. Introduction to NVIDIA DGX B300 System, NVIDIA DGX B300 System User Guide. NVIDIA, last updated 10 November 2026, viewed 14 November 2026. https://docs.nvidia.com/dgx/dgxb300-user-guide/introduction-to-dgxb300.html

7. NVIDIA GB300 NVL72. NVIDIA, viewed 8 November 2026. https://www.nvidia.com/en-us/data-center/gb300-nvl72/

8. DGX SuperPOD Architecture, NVIDIA SuperPOD DGX B300 Systems, Spectrum-4 Ethernet and DC Busbar Power Reference Architecture. NVIDIA, last updated 19 November 2025, viewed 8 November 2026. https://docs.nvidia.com/dgx-superpod/reference-architecture/scalable-infrastructure-b300/latest/dgx-superpod-architecture.html

9. DGX SuperPOD Architecture, NVIDIA SuperPOD with DGX B300 Systems, NVIDIA Quantum-X800 InfiniBand switching and AC Power Reference Architecture. NVIDIA, last updated 2 September 2026, viewed 8 November 2026. https://docs.nvidia.com/dgx-superpod/reference-architecture/scalable-infrastructure-b300-xdr/latest/dgx-superpod-architecture.html

10. U.C.C. § 2-509, Risk of Loss in the Absence of Breach. Legal Information Institute, Cornell Law School, viewed 8 November 2026. https://www.law.cornell.edu/ucc/2/2-509

11. U.C.C. § 2-513, Buyer's Right to Inspection of Goods. Legal Information Institute, Cornell Law School, viewed 8 November 2026. https://www.law.cornell.edu/ucc/2/2-513

12. U.C.C. § 2-612, "Installment Contract", Breach. Legal Information Institute, Cornell Law School, viewed 8 November 2026. https://www.law.cornell.edu/ucc/2/2-612

13. NVIDIA Corporation Quarterly Report on Form 10-Q for the quarter ended 26 July 2026. U.S. Securities and Exchange Commission, filed 26 August 2026, viewed 8 November 2026. https://www.sec.gov/Archives/edgar/data/1045810/000104581026000075/nvda-20260726.htm

14. Industry Guidance to Prevent Diversion of Advanced Computing Integrated Circuits. U.S. Bureau of Industry and Security, 13 May 2025, viewed 8 November 2026. https://www.bis.gov/media/documents/ai-counter-diversion-industry-guidance-may-13-2025.pdf

15. Guidance Regarding Enforcement of License Requirements for Advanced Computing Items for Entities Headquartered in Country Group D:5 and Macau. U.S. Bureau of Industry and Security, 31 May 2026, viewed 8 November 2026. https://www.bis.gov/media/documents/bis-guidance-may-31-2026.pdf

16. Super Micro Computer, Inc. Quarterly Report on Form 10-Q for the quarter ended 30 September 2026. U.S. Securities and Exchange Commission, filed 6 November 2026, viewed 8 November 2026. https://www.sec.gov/Archives/edgar/data/1375365/000137536526000027/smci-20260930.htm

17. Dell Technologies Delivers Second Quarter Fiscal 2027 Financial Results (Exhibit 99.1 to Form 8-K). Dell Technologies, 1 September 2026, viewed 8 November 2026. https://www.sec.gov/Archives/edgar/data/1571996/000157199626000039/exhibit991earnings8kq2fy27.htm

18. CFO Commentary on Second Quarter Fiscal 2027 Results (Exhibit 99.2 to Form 8-K). NVIDIA, 26 August 2026, viewed 8 November 2026. https://www.sec.gov/Archives/edgar/data/1045810/000104581026000073/q2fy27cfocommentary.htm

19. Super Micro Computer, Inc. Annual Report on Form 10-K for the fiscal year ended 30 June 2026. U.S. Securities and Exchange Commission, filed 31 August 2026, viewed 8 November 2026. https://www.sec.gov/Archives/edgar/data/1375365/000137536526000022/smci-20260630.htm

20. Dell Technologies Inc. Quarterly Report on Form 10-Q for the quarter ended 31 July 2026. U.S. Securities and Exchange Commission, filed 8 September 2026, viewed 8 November 2026. https://www.sec.gov/Archives/edgar/data/1571996/000157199626000046/dell-20260731.htm

21. CoreWeave, Inc. Annual Report on Form 10-K for the year ended 31 December 2025. U.S. Securities and Exchange Commission, filed 2 March 2026, viewed 8 November 2026. https://www.sec.gov/Archives/edgar/data/1769628/000176962826000104/crwv-20251231.htm

22. Bit Origin Acquires NVIDIA Blackwell B300 AI Infrastructure with Contracted Customer Deployment in Malaysia (Exhibit 99.1 to Form 6-K). Bit Origin Ltd, now SANGRIX Inc., 29 June 2026, viewed 8 November 2026. https://www.sec.gov/Archives/edgar/data/1735556/000110465926078558/tm2619272d2_ex99-1.htm

23. SANGRIX Announces November 2026 Delivery Schedule for 128-GPU NVIDIA Blackwell B300 Project in Malaysia (Exhibit 99.1 to Form 6-K). SANGRIX Inc., 16 October 2026, viewed 8 November 2026. https://www.sec.gov/Archives/edgar/data/1735556/000110465926116630/tm2627763d1_ex99-1.htm

24. Incoterms® rules. International Chamber of Commerce, viewed 8 November 2026. https://iccwbo.org/business-solutions/incoterms-rules/

25. United Nations Convention on Contracts for the International Sale of Goods (Vienna, 1980) (CISG). United Nations Commission on International Trade Law, viewed 8 November 2026. https://uncitral.un.org/en/texts/salegoods/conventions/sale_of_goods/cisg

26. Escrow. Wex, Legal Information Institute, Cornell Law School, viewed 8 November 2026. https://www.law.cornell.edu/wex/escrow

27. Consolidated Screening List. International Trade Administration, U.S. Department of Commerce, viewed 8 November 2026. https://www.trade.gov/consolidated-screening-list

28. NVIDIA-Certified Systems. NVIDIA, viewed 8 November 2026. https://www.nvidia.com/en-us/data-center/products/certified-systems/

29. U.C.C. § 2-105, Definitions: Transferability, "Goods", "Future" Goods, "Lot", "Commercial Unit". Legal Information Institute, Cornell Law School, viewed 8 November 2026. https://www.law.cornell.edu/ucc/2/2-105

30. U.C.C. § 2-601, Buyer's Rights on Improper Delivery. Legal Information Institute, Cornell Law School, viewed 8 November 2026. https://www.law.cornell.edu/ucc/2/2-601

31. U.C.C. § 2-606, What Constitutes Acceptance of Goods. Legal Information Institute, Cornell Law School, viewed 8 November 2026. https://www.law.cornell.edu/ucc/2/2-606

32. U.C.C. § 2-503, Manner of Seller's Tender of Delivery. Legal Information Institute, Cornell Law School, viewed 8 November 2026. https://www.law.cornell.edu/ucc/2/2-503

33. U.C.C. § 2-718, Liquidation or Limitation of Damages, Deposits. Legal Information Institute, Cornell Law School, viewed 8 November 2026. https://www.law.cornell.edu/ucc/2/2-718

34. Supermicro Limited Warranty. Super Micro Computer, Inc., viewed 8 November 2026. https://www.supermicro.com/en/support/warranty

35. NVIDIA System Management Interface (nvidia-smi) documentation. NVIDIA, viewed 8 November 2026. https://docs.nvidia.com/deploy/nvidia-smi/index.html

36. CoreWeave, Inc. Quarterly Report on Form 10-Q for the quarter ended 30 September 2026. U.S. Securities and Exchange Commission, filed 12 November 2026, viewed 14 November 2026. https://www.sec.gov/Archives/edgar/data/1769628/000176962826000465/crwv-20260930.htm

37. SANGRIX Inc. filings on Form 6-K, EDGAR company filing index. U.S. Securities and Exchange Commission, viewed 14 November 2026. https://www.sec.gov/cgi-bin/browse-edgar?action=getcompany&CIK=0001735556&type=6-K

About this report

Published by Rillor on 15 November 2026. Superintelligence (SI) tools assist research and drafting, and every figure cites its source. Rillor builds and secures SI systems, supplies the data they learn from and delivers the infrastructure they run on. We apply SI to financial and compute markets.

Cite as: Rillor. How to buy GPUs in bulk: allocation, grades and delivery. 15 November 2026. https://rillor.com/insights/how-to-buy-gpus-in-bulk

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.

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

Work with Rillor

Send the systems or capacity you need, the quantity, and the delivery month and location. Contact us at rillor.com/contact?type=forward or contact@rillor.com. More on forward agreements: rillor.com/markets/forward-agreements