Rillor

EXPLAINER

What to check before buying a used GPU server

On Ampere and later GPUs, nvidia-smi cannot clear lifetime ECC counts. What to read and test on a used GPU server, and the warranty, drive and export checks.



Rillor

Contents

01 Summary

02 Background

03 Method

04 Findings

05 Implications

06 Limits of this analysis

07 Sources

Summary

A used GPU server carries its history in counters, logs and paperwork, and public documents from NVIDIA, server makers, NIST and the Bureau of Industry and Security (BIS) say which records persist and what each one shows.

- On Ampere and later NVIDIA GPUs, nvidia-smi keeps aggregate ECC (error-correcting code) counts as a lifetime counter and does not support clearing them. Remapped memory rows are recorded in the InfoROM, a small persistent store on each GPU [1, 2].
- NVIDIA's return (RMA) criterion for row remapping, which replaces a degrading memory row with one of the spare rows each DRAM bank holds [17], is a set row-remapping failure flag, validated by the NVIDIA Field Diagnostic tool. The flag is set by an uncorrectable-error remap attempt in a bank that already has eight such remaps or on a row already remapped, or once 512 uncorrectable-error remaps have occurred in total [3].
- Error history beyond the GPU memory counters sits in logs: Xid messages, the NVIDIA driver's error reports in the kernel log, Fabric Manager logs on NVSwitch systems and the management controller's own event log [4, 5, 6].
- In a 2024 study of two research clusters of A100 servers, GPUs caused 28.2% of the faulty "lemon nodes" behind repeated job failures, DIMMs 20.5% and PCIe 15.4%, and the authors found historic data necessary to find such nodes [7].
- On the Delta system's 1,056 A100 and H100 GPUs, with the H100s measured over 146 days, H100s had a 3.2 times lower mean time between memory errors per GPU than A100s and fewer errors in other critical hardware [8]. Those H100s sit in GH200 Superchips [8].
- Warranty transfer differs by maker. Supermicro's published warranty covers only the original purchaser, and Dell's system warranties may transfer, at Dell's sole discretion, when the transfer is recorded with Dell [9, 10]. NVIDIA's support entitlements are non-transferable, and the NVIDIA pages reviewed state no transfer rule for the hardware warranty itself [11].
- NIST SP 800-88 Rev. 2 names resale as a point where control of storage media changes, and its sample certificate records each drive's serial number, sanitization method, tools and verification [12].
- BIS guidance of 13 May 2025 names servers by export classification with no line between new and used units, and its red flags include an unknown delivery or installation address [13].

Background

This report covers GPU servers that have run in someone else's service, whatever condition grade a seller gives them. The records described here show what a unit has been through.

Large operators revise the useful life they assign to servers. Amazon's annual report for 2025 records that server useful lives moved from five to six years from 1 January 2024, and that a subset of servers and networking equipment moved from six to five years from 1 January 2025, citing the increased pace of technology development [14]. That is an accounting estimate, and the filing does not tie it to resale.

Studies of fleets under load show where faults appear. A 2024 technical report on training a 405-billion-parameter model on up to 16,384 H100 GPUs counted 419 unexpected job interruptions in a 54-day snapshot of pre-training [15]. About 78% were attributed to confirmed or suspected hardware issues, and GPU issues accounted for 58.7% [15]. Faulty GPUs caused 148 and GPU HBM3 memory 72, together about 52.5% of the 419 (computed for this report) [15]. Those figures come from a fleet in production and are not failure rates for used units.

This report is for buyers of secondary GPU servers, the engineers who inspect them on arrival and owners preparing units for resale.

Method

What was examined. The records a buyer can read or request for a used GPU server, from GPU identity to export screening, and what each one shows according to the body that defines it.

Sources. NVIDIA documentation [1, 2, 3, 4, 5, 11, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26], Linux kernel and DMTF documentation [6, 27, 28], three operator studies [7, 8, 15], the published warranties of Supermicro and Dell [9, 10], sanitization and reuse standards from NIST, IEEE and Sustainable Electronics Recycling International (SERI) [12, 29, 30, 31], a CISA alert and an NVD record [32, 33], BIS guidance, the Know Your Customer supplement to the Export Administration Regulations (EAR) and the Consolidated Screening List [13, 34, 35], and one annual report [14].

How. NVIDIA's pages were read as published, with the memory error, Xid and debug guides last updated on 4 November 2026. Field names and Xid numbers are given as NVIDIA prints them. Studies were read in full text, and each figure comes from the passage or table cited. Figures computed for this report are labelled. Warranty and export texts are summarized as published, and nothing here is legal advice. All sources were viewed on 8 November 2026.

Findings

1. What a quote for a secondary system should state

A quote for a used GPU server prices one unit, so it has to describe that unit. SERI's R2v3 standard, in its Appendix C on test and repair, recognizes organizations that test, repair or refurbish equipment and "accurately communicate the level of functionality, cosmetic condition, and data sanitization status" [31]. It separates evaluation from testing. Appendix C certification is not required for "simple power-on or non-functional evaluations of electronic devices or components, such as visual inspections" [30], and evaluations such as power-on verification "are not a final indication of the device's level of functionality" [31].

R2v3 is one standard a seller may hold. Its distinction gives a buyer three questions for any quote: what level of function was tested, what physical condition the unit is in and what happened to the data on its drives.

2. Serials and configuration as received

nvidia-smi reports a serial number that "matches the serial number physically printed on each board" [1]. On a board with several GPUs, that serial "might be shared between multiple GPUs on the same board", while the GPU UUID is a globally unique, immutable identifier that corresponds to no physical label [1]. Listing both for every GPU ties each set of counters to one GPU and each GPU to a label a buyer can see.

The same output gives the VBIOS version, the board, GPU and FRU part numbers, the serial number of the chassis and each GPU's module ID within the node [1]. It reports the InfoROM image version, which "uniquely describes the exact version of the infoROM flashed on the board" [1]. A checksum validation field, available only through `--query-gpu inforom.checksum_validation`, reads valid, invalid or N/A [1].

The GPU and switch counts follow from the baseboard. NVIDIA's Fabric Manager guide gives an HGX H100 8-GPU baseboard eight GPUs and four NVSwitches, an HGX A100 eight GPUs and six NVSwitches, and HGX B200 and B300 eight GPUs and two NVSwitch ASICs [4]. NVIDIA's node checks start with the same items: confirm GPU count, PCIe link speed and VBIOS, and on NVSwitch systems confirm that Fabric Manager is running and the NVLink fabric topology is correct [23]. They also check that all network links are active, either simply or with a more robust check of expected link speed, firmware and PCIe link width [23].

PCIe fields need care. nvidia-smi gives the current and the maximum link generation and width, and the current values "may be reduced when the GPU is not in use" [1]. A reading at idle can therefore show a narrower link than the system supports, so the link is read with the GPU busy.

The management controller belongs in the configuration too. CVE-2023-25528, published in September 2023, was a vulnerability in a web server plugin of the NVIDIA DGX H100 baseboard management controller (BMC) in which an unauthenticated attacker may cause a stack overflow with a crafted network packet, scored 8.8 by NVIDIA and 9.8 by the National Vulnerability Database (NVD) under CVSS 3.1 [33]. BMC firmware has had published vulnerabilities, so its version belongs in the report.

3. Memory health counters NVIDIA documents

ECC counts. Volatile ECC counters cover errors since the last driver load. Aggregate counts "persist indefinitely and thus act as a lifetime counter", and nvidia-smi's documentation states that "Clearing aggregate counts is not supported on Ampere+" [1]. On Ampere and later GPUs, SRAM errors are split into parity and SEC-DED errors, with a flag for an exceeded threshold [1].

Row remapping. From the Ampere architecture on, every DRAM bank has spare rows in hardware, and a faulty row is remapped to one of them [17]. A remap needs a GPU reset to take effect and persists for the life of the GPU [17]. Row remapping supports up to 512 remaps for the frame buffer, against a maximum of 64 retirements under the older page retirement scheme, and a remap for a correctable error can be replaced by one for an uncorrectable error when a bank's reserved rows run out [17]. NVIDIA lists row remapping on GA100, GA10x, AD10x, GH100 and GB10x GPUs, and GPU memory repair, which swaps in spare DRAM channels or L2 cache slices, on GB10x only, where support "is not guaranteed for all Blackwell GPUs" [16, 18].

The fields to read. nvidia-smi's row remapper section reports active and inactive remaps for correctable and for uncorrectable errors, a Pending field and a Remapping Failure Occurred field [1]. NVIDIA's statistics page adds that the count of remapped rows "is the number of entries recorded in the InfoROM, not the ones remapped in hardware", and that the counts and both flags are readable in-band through NVML and nvidia-smi and out-of-band through SMBPBI [2]. If a repair flag such as Channel Repair Pending is set, the documentation says to check the GPU's recovery action [1].

The bank histogram. nvidia-smi classes each bank's remap availability as Maximum, High, Partial, Low or None, and correctable remaps do not count toward it because uncorrectable remaps can evict them [1]. NVIDIA's examples show 640 banks at Maximum, and in a second case 635 banks with eight remaps left, three with seven and two with one [2]. Those are examples on one page, not a rule for every GPU, and the per-bank availability column shown in them "is not available through APIs or nvidia-smi" [2].

Repairs waiting for a reset. The row-remapping pending flag means the remap takes effect at the next GPU reset, and unaffected applications can keep running until then [19]. A buyer can ask the seller to reset the GPUs and read the counters again before the report is issued.

One fleet study, both halves. A study of the Delta system used 2.5 years of operational data, 11.7 million GPU hours, from 1,056 GPUs: 895 days and 9.6 million GPU hours on 448 A100s, and 146 days and 2.1 million GPU hours on 608 H100s [8]. It found H100 memory resilience worse than A100's: a 3.2 times lower per-GPU mean time between memory errors, and a per-GB mean time between errors for H100 HBM3 about 24% lower than for A100 HBM2e, about 8.5 million hours against 11.3 million [8]. Row remapping and related mechanisms still mitigated 92% of uncorrectable ECC memory errors on the H100s [8]. In other critical hardware the H100s did better, with only three cases of GPU system processor (GSP) errors and no NVLink errors in their 146 days [8]. The study's H100s sit in GH200 Superchips coupled to Grace CPUs, which its authors distinguish from discrete H100 GPUs [8].

4. NVIDIA's row-remapping RMA thresholds

NVIDIA's RMA Policy page states that the NVIDIA Field Diagnostic tool determines whether a GPU qualifies for return. For row remapping, the criterion is met when the row-remapping failure flag is set and validated by the field diagnostic [3]. Any of three events sets the flag [3]:

- a remap attempt for an uncorrectable error in a bank that already has eight rows remapped for uncorrectable errors
- a remap attempt for an uncorrectable error on a row already remapped, which can happen with fewer than eight remaps in the bank
- 512 remaps for uncorrectable errors having occurred in total

On Blackwell GPUs, a third remap attempt for an uncorrectable error in one bank invokes HBM channel repair if a spare channel is available [3]. The driver prints Xid 160 when a DRAM channel or L2 slice is marked for repair, and the repair needs a reboot to take effect [18]. SRAM has its own flag, set after more than four uncorrectable unique-count events within an address bank of parity-protected SRAM, or more than two in SEC-DED protected SRAM [3].

NVIDIA's triage guide calls the field diagnostic "the authoritative and comprehensive NVIDIA tool for determining the health of GPUs", says it is usually required before an RMA can be started and tells users to ask their system vendor when, if and how to run it [22].

A set failure flag, once the field diagnostic confirms it, marks a GPU that meets NVIDIA's documented return condition. Below the flag, no NVIDIA page reviewed here sets a count of remapped rows or ECC errors that makes a GPU acceptable or unacceptable. A buyer who wants a stricter limit writes it into the purchase terms.

5. Error history in system logs

Xid messages. An Xid is an error report from the NVIDIA driver, printed to the operating system's kernel log or event log. It can point to hardware, NVIDIA software or a user's application, and its meaning is consistent across driver versions [5]. On Linux, Xids go to the kernel log buffer and, depending on the distribution, to files such as /var/log/messages or /var/log/syslog, and NVIDIA suggests searching for "NVRM: Xid" [20].

Xid	Meaning in NVIDIA's documentation
48	Uncorrectable (double-bit) ECC error, with Xid 171 for DRAM and Xid 172 for SRAM [2, 21]
63	A row-remapping entry recorded to the InfoROM [2]
64	A failure to record a row-remapping entry to the InfoROM [2]

Xid	Meaning in NVIDIA's documentation
74	NVLink error, at the link itself or the device at its remote end [21]
79	GPU has fallen off the bus, often a hardware failure on the PCIe link [21]
92	High single-bit ECC error rate [21]
94 and 95	Contained and uncontained ECC errors [2, 21]

NVIDIA's triage guide ties Xids 48, 63, 64, 92, 94 and 95 to GPU memory errors and sets out responses [22]. If Xid 48 is followed by Xid 63 or 64, it says to drain the node and reset the GPUs, and on an A100 Xid 64 calls for an immediate reboot [22]. Xid 79 means drain the node and report the issue, and Xid 92 means run the field diagnostic [22]. On NVSwitch systems, switch errors appear as SXids, which apply only to Hopper and earlier generations [5], and Fabric Manager writes its own log, by default /var/log/fabricmanager.log [4].

What to ask for. When a GPU issue is reported, NVIDIA's guide asks for the operating system and driver details, the output of nvidia-bug-report.sh, Fabric Manager logs for HGX systems and DCGM diagnostic logs [22]. Because Xids are written to the operating system's logs [20], a buyer can ask for logs captured before the operating system was reinstalled.

Host logs. Linux prints PCIe Advanced Error Reporting (AER) messages to the console, provided the firmware grants the operating system AER control [27]. The kernel's EDAC module harvests correctable and uncorrectable memory errors, and its documentation says correctable errors "can but must not necessarily be a predictor" of uncorrectable ones [28]. DMTF's Redfish LogEntry schema covers the legacy IPMI System Event Log (SEL) among its log formats [6], and the BMC that keeps it operates separately from the operating system [32].

History finds lemon nodes. A 2024 study of two research clusters of A100 servers analyzed 11 months of data, 4 million jobs and over 150 million A100 GPU hours, on DGX-based nodes of eight A100 80GB GPUs [7]. It defines lemon nodes as servers that cause repeated job failures but escape existing health checks and repair workflows, and it found them with signals such as the number of unique Xids a node experienced, its repair tickets and the number of times it was taken out of availability [7]. It identified 40 lemon nodes with more than 85% accuracy, 1.2% of one cluster's footprint and 1.7% of the other's, and reports that failures of large jobs on 512 or more GPUs fell from 14% to 4% [7]. Their root causes were spread across the server [7]:

Root cause	Share of lemon nodes
GPU	28.2%

Root cause	Share of lemon nodes
DIMM	20.5%
PCIe	15.4%
EUD (the paper's label)	10.3%
NIC	7.7%
BIOS	7.7%
Power supply	5.1%
CPU	2.6%
Optics	2.6%

DIMMs, PCIe and NICs together account for 43.6%, more than the GPUs' share (computed for this report from Table II) [7]. The authors state that historic data is necessary to find defective nodes [7]. Repair tickets and out-of-service counts are the history to request, and the host's memory, PCIe path and network cards are part of what is bought.

6. Diagnostics to run or require

NVIDIA's Data Center GPU Manager (DCGM) diagnostics are meant to "assess cluster readiness levels before a workload is deployed" [24]. They check pending page retirements, the PCIe and NVLink interfaces, frame buffer and memory, compute engines, PCIe replay counters, power and thermal stress and memory bandwidth [24]. NVIDIA states that they are not designed to provide comprehensive hardware diagnostics, replace the field diagnosis tools or facilitate any RMA process [24]. Run times below are NVIDIA's measurements on Hopper systems [24]:

Level	Tests that start at this level	4-GPU system	8-GPU system
r1	software	under 2.5 seconds	under 2.5 seconds
r2	pcie, memory, memory_bandwidth	under 2.5 minutes	under 10.5 minutes

Level	Tests that start at this level	4-GPU system	8-GPU system
r3	diagnostic, targeted_stress, targeted_power, nvbandwidth, nccl_tests	under 10 minutes	under 35 minutes
r4	memtest, pulse	under 45 minutes	under 2.25 hours

Logs go to /var/log/nvidia-dcgm/nvvs.log [24]. NVIDIA's triage guide recommends level 3 for fully stressing a system [22], and its Best Practices page uses level 2 for a quick check before a job [23]. A used unit's report names the level run and carries its logs. Across nodes, the triage guide uses all_reduce_perf to establish network performance between groups of nodes, expects pairwise tests to give the same performance and advises recording network results before production [22]. The Best Practices page lists key metrics to collect as time series on GPU systems, in-band with the DCGM Prometheus plugin or out-of-band with IPMI, among them the power draw of each power supply, GPU power, temperature, memory temperature and clocks, node fan speed and inlet temperature [23].

7. Warranty and who stands behind it

Supermicro's published defaults. Supermicro's warranty runs from the date of its invoice, excludes "products, components, or software supplied by third parties" and "applies only to the original purchaser of the Product" [9]. For GPU systems it lists three years of labor, one year of parts and one year of advance cross-ship, terms that apply "In the absence of special warranty terms set forth in a corresponding invoice" [9]. An RMA needs the invoice date, invoice number, product model name, serial number and reason for return [9]. Repair work by a party other than Supermicro, or removal or alteration of the original identification markings, voids the warranty [9].

Dell's rule. Dell's limited hardware warranties on systems may be transferred, at Dell's sole discretion, if the current owner transfers ownership and records the transfer with Dell, while the warranty on Dell-certified memory may not be transferred [10]. Warranty support applies only when the product is located in the country where Dell originally sold the system, and for commercial customers Dell may void the warranty if third-party products not provided by Dell are installed [10].

NVIDIA's terms. Any user can check an NVIDIA product's warranty on NVIDIA's Warranty Check form [25]. NVIDIA Enterprise Support covers designated NVIDIA branded solutions bought from NVIDIA or a reseller it has authorized, and apart from the limited standard warranty on NVIDIA hardware, support needs a purchased contract and an active entitlement [11]. Entitlements "are non-transferable, are non-assignable and may be terminated when the product is transferred to another party" [11]. These pages state no transfer rule for the hardware warranty itself, so a buyer checks the serial number with NVIDIA instead of assuming either way. Once NVIDIA announces a product's End of Sales, its typical service obligation runs five years "unless noted otherwise", and End of Service Life is the last day a service contract is honored [26].

Three questions follow for a used unit. Whose warranty applies: the maker's under its transfer rule, the seller's own or a third party's? What does it exclude? When does it end, and where does the product sit in its maker's lifecycle?

8. Drives and prior-owner data

NIST SP 800-88 Rev. 2, the Guidelines for Media Sanitization published in September 2025, supersedes Rev. 1 of December 2014 [12]. It defines sanitization as a process that "renders access to target data on the media infeasible for a given level of effort" and names three methods, clear, purge and destroy, with purge to be used instead of clear when possible [12]. Logical purge techniques vary by media type, and NIST points to IEEE 2883 for them [12]. IEEE 2883-2022 specifies methods for sanitizing logical and physical storage [29].

NIST names the moments that matter in a sale: reuse includes refurbishment, and media control changes when storage is returned from a lease, donated, resold for reuse outside the organization or sent to a recycling facility [12]. It describes validating the result, since a drive "may appear fully functional" while part of it is no longer reachable through its interface [12]. Its sample Certificate of Sanitization records, per item, the vendor, media type, model and serial number, the method, technique and tools with their version, verification and validation, and the destination, with a signature and date [12]. A buyer can ask the seller for one such record per drive. SERI's R2v3 Appendix B sets requirements for logical sanitization and for an enhanced level of physical sanitization [30].

The BMC holds state of its own. CISA and NSA describe BMCs as trusted components that operate separately from the operating system and firmware to allow remote management "even when the system is shut down", and note that hardened credentials, firmware updates and network segmentation options are often overlooked [32]. On arrival, a buyer can reset the BMC's users and credentials, update its firmware and record the version before and after [32, 33].

9. Export rules on resale

This finding is general information about published rules, not legal advice.

BIS's Industry Guidance to Prevent Diversion of Advanced Computing Integrated Circuits is dated 13 May 2025, and the bis.gov file was regenerated on 15 October 2026 [13]. Its first footnote counts among the covered items export control classification numbers (ECCNs) 3A090.a and 4A090.a and .z items in Categories 3, 4 and 5, "such as servers classified as ECCN 5A992.z" [13]. These items have been subject to BIS export restrictions since October 2022 [13]. The guidance names servers by classification and draws no line between new and used units [13].

Among its 11 new transactional and behavioral red flags are [13]:

- a residential address with no other location for the chips, where the quantity is inconsistent with personal use
- an unknown ultimate delivery or installation address

- a customer co-located with, or at an address similar to, a party on the Consolidated Screening List
- a data center that does not or cannot affirm it has the infrastructure, such as power, cooling or physical space

Its due diligence actions include an end-user certification with detailed information on all proposed transaction parties, under § 748.5 of the EAR, and a written attestation from the data center that the end user is authorized to operate there and that it has the needed infrastructure [13]. In-country transfers, defined as a change in end use or end user, may need a license under the EAR's catch-all controls when there is knowledge that the transferee will use the items to train models for a weapons of mass destruction or military-intelligence end use or user, for or on behalf of parties headquartered in Country Group D:5 (including China) or Macau, and transferors "may not self-blind" [13]. Its recommended practices include learning whether a customer plans to resell and, if so, the next customer's name and address [13].

The EAR's Know Your Customer supplement lists red flags that fit a used sale: a customer declining installation and testing when included in the sales price or normally requested, equipment configurations incompatible with the stated destination, such as 120 volts in a country with 220 volts, a customer reluctant to offer end-use information and a freight forwarding firm listed as the final destination [34]. When red flags are raised there is a duty to check out the suspicious circumstances, and the supplement states that it "does not change or interpret the EAR" [34]. The Consolidated Screening List combines export screening lists of the Departments of Commerce, State and the Treasury, with search tools updated every day at 5:00 a.m. Eastern time, and its users must check the Federal Register before taking further action [35].

These rules explain the questions a seller that screens its sales asks: end user, end use, installation address and data center readiness.

10. A buyer's checklist

Stage	Check	Finding
Before the quote	The level of function tested, the physical condition and each drive's sanitization status [31]	1
Before the quote	The serial number and UUID of every GPU, the chassis serial and the configuration as received [1]	2
Before the quote	Which warranty applies, who provides it, when it ends and whether a transfer was recorded, with NVIDIA warranties checked by serial number [9, 10, 25]	7
Before shipment	Aggregate ECC counts, remapped-row counts, the pending and failure flags and the bank histogram, read after a GPU reset [1, 2, 19]	3

Stage	Check	Finding
Before shipment	Whether any failure flag is set and whether the field diagnostic was run [3, 22]	4
Before shipment	Bug-report output, Fabric Manager and DCGM logs from before any reinstall, the BMC event log and the repair history [6, 7, 20, 22]	5
Before shipment	One sanitization record per drive [12]	8
Before shipment	End user, end use and installation address for screening [13, 34]	9
On arrival	GPU and NVSwitch counts, VBIOS, InfoROM image and checksum, and PCIe links under load, against the report [1, 4, 23]	2
On arrival	Memory counters read again and compared with the seller's figures [1, 2]	3
On arrival	DCGM diagnostics at a stated level, with logs kept [24]	6
On arrival	BMC users and credentials reset and firmware updated [32, 33]	8
Before production	Network tests between nodes, with results recorded [22]	6

Implications

For buyers. The records can be named in a request: serials and UUIDs, firmware versions, the row remapper fields and failure flag, the logs NVIDIA asks for when a GPU issue is reported, the warranty with its provider and end date, and one sanitization record per drive [1, 3, 9, 12, 22]. Any limit below NVIDIA's failure flag is a term the buyer sets.

For owners preparing units for resale. The same records come from a system's time in service: logs from before the operating system is reinstalled, repair tickets, the original invoice details a maker's RMA process asks for and a sanitization record for each drive [7, 9, 12, 20].

For engineers inspecting a delivery. Read PCIe links under load, treat remapped-row counts as InfoROM entries and reset the GPUs before reading counters with a pending remap [1, 2, 19]. DCGM checks readiness, and the field diagnostic decides a return [22, 24]. The Delta results give memory counters extra weight on Hopper units, with that study's GH200 caveat [8].

For every party to a sale. The BIS guidance draws no line between new and used units, so screening questions come with a used sale as with a new one [13, 34].

Limits of this analysis

- This report gathered no prices for used GPU servers.
- No source reviewed defines condition grades such as surplus, refurbished or used, and this report proposes none.
- The failure studies measured fleets in service [7, 8, 15], and no source reviewed gives failure rates for used or refurbished GPU servers. The H100 results come from GH200 Superchips, not discrete HGX H100 boards, and cover 146 days against 895 for the A100s [8].
- The 2024 training report's table prints percentages that do not equal its counts divided by 419, so this report uses the counts and the report's prose figures [15].
- No source reviewed sets a burn-in duration or pass criteria for a used unit. DCGM's run times are NVIDIA's measurements on Hopper systems [24].
- Ways to detect a reflashed VBIOS, an altered InfoROM or a swapped baseboard are outside this report, beyond the version and checksum fields nvidia-smi reports [1].
- Warranty terms come from two makers' published pages, and an invoice can set special terms [9, 10]. Other makers' terms were not reviewed.
- Export content is general information from BIS guidance, the EAR supplement and the screening list page [13, 34, 35]. It is not legal advice.

Sources

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

2. User Visible Statistics. NVIDIA GPU Memory Error Management. NVIDIA, last updated 4 November 2026, viewed 8 November 2026. https://docs.nvidia.com/deploy/a100-gpu-mem-error-mgmt/user-visible-statistics.html

3. RMA Policy (RMA policy thresholds for row remapping). NVIDIA GPU Memory Error Management. NVIDIA, last updated 4 November 2026, viewed 8 November 2026. https://docs.nvidia.com/deploy/a100-gpu-mem-error-mgmt/rma-policy-thresholds-for-row-remapping.html

4. Fabric Manager User Guide. NVIDIA, last updated 7 October 2026, viewed 8 November 2026. https://docs.nvidia.com/datacenter/tesla/fabric-manager-user-guide/index.html

5. Xid Errors: Introduction. NVIDIA Xid Errors. NVIDIA, last updated 4 November 2026, viewed 8 November 2026. https://docs.nvidia.com/deploy/xid-errors/introduction.html

6. LogEntry schema, version 1.20.0. Redfish, DMTF, viewed 8 November 2026. https://redfish.dmtf.org/schemas/v1/LogEntry.v1_20_0.json

7. Revisiting Reliability in Large-Scale Machine Learning Research Clusters. arXiv 2410.21680 v2. arXiv, 6 February 2025, viewed 8 November 2026. https://arxiv.org/abs/2410.21680

8. Story of Two GPUs: Characterizing the Resilience of Hopper H100 and Ampere A100 GPUs. arXiv 2503.11901 v4. arXiv, 10 December 2025, viewed 8 November 2026. https://arxiv.org/abs/2503.11901

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

10. Dell Limited Hardware Warranty. Dell Technologies, viewed 8 November 2026. https://www.dell.com/en-us/legal/lp/art-limited-hardware-warranties

11. NVIDIA Enterprise Support Details. NVIDIA Enterprise Support and Services User Guide. NVIDIA, viewed 8 November 2026. https://docs.nvidia.com/enterprise-support-and-services-user-guide/support-details

12. Chandramouli and Hibbard. Guidelines for Media Sanitization, NIST SP 800-88 Rev. 2. National Institute of Standards and Technology, September 2025, viewed 8 November 2026. https://csrc.nist.gov/pubs/sp/800/88/r2/final https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-88r2.pdf

13. Industry Guidance to Prevent Diversion of Advanced Computing Integrated Circuits. Bureau of Industry and Security, U.S. Department of Commerce, 13 May 2025 (file regenerated 15 October 2026), viewed 8 November 2026. https://www.bis.gov/media/documents/ai-counter-diversion-industry-guidance-may-13-2025.pdf

14. Amazon.com, Inc. Form 10-K for the fiscal year ended 31 December 2025. SEC EDGAR, filed 6 February 2026, viewed 8 November 2026. https://www.sec.gov/Archives/edgar/data/1018724/000101872426000004/amzn-20251231.htm

15. The Llama 3 Herd of Models. arXiv 2407.21783 v3. arXiv, 23 November 2024, viewed 8 November 2026. https://arxiv.org/abs/2407.21783

16. Supported GPUs. NVIDIA GPU Memory Error Management. NVIDIA, last updated 4 November 2026, viewed 8 November 2026. https://docs.nvidia.com/deploy/a100-gpu-mem-error-mgmt/supported-gpus.html

17. Row Remapping. NVIDIA GPU Memory Error Management. NVIDIA, last updated 4 November 2026, viewed 8 November 2026. https://docs.nvidia.com/deploy/a100-gpu-mem-error-mgmt/row-remapping.html

18. RAS Repair. NVIDIA GPU Memory Error Management. NVIDIA, last updated 4 November 2026, viewed 8 November 2026. https://docs.nvidia.com/deploy/a100-gpu-mem-error-mgmt/hbm-channel-repair.html

19. Error Recovery and Response Flags. NVIDIA GPU Memory Error Management. NVIDIA, last updated 4 November 2026, viewed 8 November 2026. https://docs.nvidia.com/deploy/a100-gpu-mem-error-mgmt/error-recovery-and-response-flags.html

20. Working with Xid Errors. NVIDIA Xid Errors. NVIDIA, last updated 4 November 2026, viewed 8 November 2026. https://docs.nvidia.com/deploy/xid-errors/working-with-xid-errors.html

21. Analyzing Xid Errors with the Xid Catalog. NVIDIA Xid Errors. NVIDIA, last updated 4 November 2026, viewed 8 November 2026. https://docs.nvidia.com/deploy/xid-errors/analyzing-xid-catalog.html

22. GPU Node Triage. NVIDIA GPU Debug Guidelines. NVIDIA, last updated 4 November 2026, viewed 8 November 2026. https://docs.nvidia.com/deploy/gpu-debug-guidelines/gpu-node-triage.html

23. Best Practices. NVIDIA GPU Debug Guidelines. NVIDIA, last updated 4 November 2026, viewed 8 November 2026. https://docs.nvidia.com/deploy/gpu-debug-guidelines/best-practices.html

24. DCGM Diagnostics. NVIDIA DCGM User Guide. NVIDIA, last updated 17 June 2026, viewed 8 November 2026. https://docs.nvidia.com/datacenter/dcgm/latest/user-guide/dcgm-diagnostics.html

25. Enterprise Customer Support. NVIDIA, viewed 8 November 2026. https://www.nvidia.com/en-us/support/enterprise/

26. Support Throughout Product Lifecycle. NVIDIA Enterprise Support and Services User Guide. NVIDIA, viewed 8 November 2026. https://docs.nvidia.com/enterprise-support-and-services-user-guide/product-lifecycle-support

27. The PCI Express Advanced Error Reporting Driver Guide HOWTO. The Linux Kernel documentation, kernel.org, viewed 8 November 2026. https://docs.kernel.org/PCI/pcieaer-howto.html

28. Reliability, Availability and Serviceability (RAS). The Linux Kernel documentation, kernel.org, viewed 8 November 2026. https://docs.kernel.org/admin-guide/RAS/main.html

29. IEEE 2883-2022, IEEE Standard for Sanitizing Storage. IEEE Standards Association, published 17 August 2022, viewed 8 November 2026. https://standards.ieee.org/ieee/2883/10277/

30. R2v3 Appendix Applicability Guidance. Sustainable Electronics Recycling International (SERI), last updated 11 September 2024, viewed 8 November 2026. https://sustainableelectronics.org/knowledge-base/r2v3-appendix-applicability-guidance/

31. R2v3 Appendix Determination Tool. Sustainable Electronics Recycling International (SERI), last updated 26 September 2024, viewed 8 November 2026. https://sustainableelectronics.org/knowledge-base/r2v3-appendix-determination-tool/

32. CISA and NSA Release Joint Guidance on Hardening Baseboard Management Controllers (BMCs). Cybersecurity and Infrastructure Security Agency, 14 June 2023, viewed 8 November 2026. https://www.cisa.gov/news-events/alerts/2023/06/14/cisa-and-nsa-release-joint-guidance-hardening-baseboard-management-controllers-bmcs

33. CVE-2023-25528 Detail. National Vulnerability Database, NIST, published 20 September 2023, viewed 8 November 2026. https://nvd.nist.gov/vuln/detail/CVE-2023-25528

34. Supplement No. 3 to Part 732, BIS's "Know Your Customer" Guidance and Red Flags. Title 15, Code of Federal Regulations, eCFR, Office of the Federal Register, viewed 8 November 2026. https://www.ecfr.gov/current/title-15/subtitle-B/chapter-VII/subchapter-C/part-732/appendix-Supplement%20No.%203%20to%20Part%20732

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

About this report

Published by Rillor on 10 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. What to check before buying a used GPU server. 10 November 2026. https://rillor.com/insights/buying-used-gpu-servers

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 workload, system, quantity, delivery window and destination. Contact us at rillor.com/contact?type=gpu-systems or contact@rillor.com. More on GPU systems: rillor.com/compute/gpu-systems