

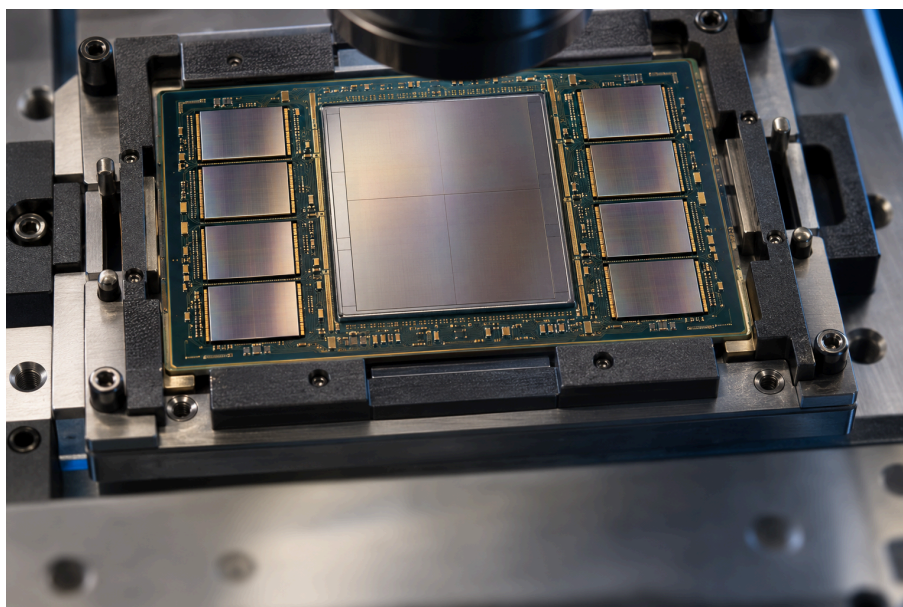


ECCN 3A090 and 4A090: TPP and Performance Density Explained

Plan and build private AI compute with GPU Smith. We connect workload requirements with GPU hardware, networking, deployment and operating decisions so your infrastructure fits the work it must perform.

Published by GPU Smith · 2026-10-29 · nvidia eccn · nvidia eccn list · nvidia eccn lookup · what is tpp · 3a090 total processing performance calculation · 4a090 server classification gpu · performance density export control gpu die area

Original article: <https://gpusmith.com/articles/en/eccn-3a090-4a090-tpp-performance-density-explained>



In this article

1. Why classification comes before price
 2. The ECCN map: 3A090, 4A090 and the .z paragraphs
 3. TPP: $2 \times \text{MacTOPS} \times \text{bit length}$
 4. Performance density: TPP per mm^2 of logic die
 5. The 3A090.a and 3A090.b parameter table
 6. Worked example 1: H100 SXM and H200 from datasheet numbers
 7. Worked example 2: HGX B200, HGX B300 and MI355X
 8. Where the rules stand (verified 2026-10-28)
 9. License Exceptions NAC and ACA
 10. The January 2026 policy: TPP under 21,000, DRAM bandwidth under 6,500 GB/s
 11. Buyer-side checklist for classification
 12. FAQ
 13. Sources
-

Why classification comes before price

For an accelerator or an AI server, the Export Control Classification Number (ECCN) decides whether an export, reexport or in-country transfer needs a license, which license exceptions might apply, and what the seller will ask you to certify. For data center GPUs and the servers built around them, the answer almost always comes down to two ECCNs: 3A090 for the integrated circuit and 4A090 for the computer, assembly or component that contains it. Both are decided by arithmetic on numbers from the datasheet.

Regulation text quoted here is the eCFR version of [15 CFR part 774, Supplement No. 1](#), up to date as of 2026-10-27 when we read it on 2026-10-28.

This is not legal advice. The TPP figures below are our own derivations from vendor marketing numbers, not classifications. The manufacturer's stated ECCN, or a BIS classification ruling, is what counts. For policy history and destinations, see our [NVIDIA GPU export restrictions overview](#).

The ECCN map: 3A090, 4A090 and the .z paragraphs

3A090 covers integrated circuits. Since the December 2024 rule it has three paragraphs ([FR 2024-28270](#)):

- **3A090.a:** the most capable logic ICs (thresholds below).
- **3A090.b:** a lower band of logic ICs.
- **3A090.c:** high-bandwidth memory with a 'memory bandwidth density' above 2 GB/s per mm^2 . The note excludes co-packaged parts whose dominant function is processing, so a GPU with HBM on the package is classified as a logic IC, not as 3A090.c.

According to Note 3 to 3A090, the entry covers GPUs, TPUs, neural processors, co-processors and accelerators, FPGAs and ASICs. **4A090** covers "computers, 'electronic assemblies,' and 'components' containing integrated circuits, any of which meets or exceeds the limits in 3A090.a" (4A090.a) or 3A090.b

(4A090.b). One qualifying chip is enough. An eight-GPU baseboard, a PCIe accelerator card and a complete server can all land in 4A090.a if a single IC inside meets 3A090.a.

One limit on the "computer" side: answering comments in October 2023, BIS said "for control under ECCN 4A090, an item must be a general purpose computer", and that it does not classify network security or DNA sequencing appliances in Category 4 ([FR 2023-23055](#)). Such appliances classify elsewhere, and the **.z paragraphs** (3A001.z, 4A003.z, 5A002.z, 5A992.z and others) then apply the same advanced computing controls to items that "meet or exceed the performance parameters in 3A090".

TPP: $2 \times \text{MacTOPS} \times \text{bit length}$

Technical Note 2 to 3A090.a and 3A090.b defines the metric:

'Total processing performance' ('TPP') is $2 \times \text{'MacTOPS'} \times \text{'bit length of the operation'}$, aggregated over all processing units on the integrated circuit.

Four rules in the note do most of the work ([eCFR, Supp. 1 to part 774](#)):

1. **The 2 is the datasheet convention.** One multiply-accumulate ($D = A \times B + C$) counts as two operations, so " $2 \times \text{MacTOPS}$ may correspond to the reported TOPS or FLOPS on a datasheet." In practice, TPP is the datasheet TFLOPS or TOPS figure multiplied by the bit length.
2. **Bit length is the widest multiply input.** For FP8 $\times$ FP8 it is 8; for FP16 it is 16.
3. **Use the highest TPP across precisions.** MacTOPS is taken "at its maximum value theoretically possible", meaning the highest figure the manufacturer claims. If the chip supports several bit lengths, "the highest 'TPP' value should be evaluated."
4. **Dense, not sparse.** If a chip processes both sparse and dense matrices, TPP uses the dense figure. Most NVIDIA datasheets headline sparse numbers, so you halve them.

The note's own check: 600 TOPS at 8 bits or 300 TFLOPS at 16 bits both give 4,800. Because tensor throughput usually doubles as precision halves, many chips have the same TPP at FP16, FP8 and INT8. Where that pattern breaks, the highest product wins, as the Blackwell and MI355X examples show.

TPP is compute-only. Interconnect and memory bandwidth are not part of it; memory bandwidth enters only through the separate 'total DRAM bandwidth' test in the January 2026 licensing policy.

Performance density: TPP per mm^2 of logic die

The same note defines 'performance density' (PD) as TPP divided by 'applicable die area'. That area is "measured in millimeters squared and includes all die area of logic dies manufactured with a process node that uses a non-planar transistor architecture". In practice that means FinFET or gate-all-around logic dies. Memory dies made on planar processes are not counted.

BIS added PD in October 2023 because it "prevents the workaround of simply purchasing a larger number of smaller datacenter AI chips which, if combined, would be equally powerful as restricted chips" ([FR 2023-23055](#)). PD only matters between 1,600 and 4,800 TPP; at 4,800 or above a chip is 3A090.a regardless of die area.

The 3A090.a and 3A090.b parameter table

Paragraph	TPP	Performance density
3A090.a.1	$\geq 4,800$	any
3A090.a.2	$\geq 1,600$	≥ 5.92
3A090.b.1	$\geq 2,400$ and $< 4,800$	≥ 1.6 and < 5.92
3A090.b.2	$\geq 1,600$	≥ 3.2 and < 5.92

Source: current 3A090 text in the [eCFR](#), unchanged in substance since the [October 2023 rule](#).

Note 2 adds a market test: 3A090.a and .b "do not apply to items that are not designed or marketed for use in datacenters and do not have a 'total processing performance' of 4800 or more." A workstation or gaming part below 4,800 TPP is outside 3A090 whatever its density; at or above 4,800 it is caught but can use License Exceptions NAC/ACA. "Designed or marketed for use in datacenters" is a factual question, so keep the datasheet and marketing material on file.

Hypothetical chips (illustrative inputs, not real products) show how density moves a part between bands:

TPP	Applicable die area	PD (derived)	Result if datacenter-marketed
3,000	400 mm ²	7.50	3A090.a.2
3,000	1,000 mm ²	3.00	3A090.b.1 (also b.2)
2,000	500 mm ²	4.00	3A090.b.2
2,000	1,000 mm ²	2.00	Not 3A090 (fails b.1 TPP floor and b.2 PD floor)

Worked example 1: H100 SXM and H200 from datasheet numbers

NVIDIA's [H100 page](#) lists the H100 SXM at 3,958 teraFLOPS FP8, 1,979 teraFLOPS FP16 and 3,958 TOPS INT8, each footnoted "With sparsity". It lists FP64 Tensor Core at 67 teraFLOPS without a sparsity footnote.

Precision	Datasheet (sparse)	Dense	Bit length	TPP (derived)
FP8	3,958 TFLOPS	1,979	8	15,832
FP16	1,979 TFLOPS	989.5	16	15,832
INT8	3,958 TOPS	1,979	8	15,832
FP64 Tensor	67 TFLOPS (dense)	67	64	4,288

The highest value, 15,832, is more than three times the 4,800 line. For density, NVIDIA's architecture write-up gives the GH100 die as "80 billion transistors, a die size of 814 mm²" on TSMC 4N ([NVIDIA Technical Blog, 2022-03-22](#)). That gives PD = 15,832 / 814 $\approx$ 19.4 (derived). TPP alone already meets 3A090.a.1, so an HGX

H100 baseboard or H100 server is a 4A090.a candidate.

The [H200 page](#) lists the same 3,958 TFLOPS FP8 with sparsity, so its derived TPP is also 15,832; what differs is memory, 141 GB at 4.8 TB/s (4,800 GB/s), which matters for the January 2026 policy. The H100 NVL, at 3,341 TFLOPS FP8 sparse, works out to $1,670.5 \times 8 = 13,364$ (derived): lower, but still 3A090.a.

Worked example 2: HGX B200, HGX B300 and MI355X

NVIDIA's [HGX page](#) gives 8-GPU system totals only. HGX B200 is listed at FP4 "144 PFLOPS | 72 PFLOPS" (sparse | dense), FP8/FP6 at 72 PFLOPS sparse ("Dense is ½ sparse spec shown"), FP16/BF16 at 36 PFLOPS sparse and INT8 at 72 POPS sparse. HGX B300 is listed at FP4 "144 PFLOPS | 108 PFLOPS", FP8/FP6 72 PFLOPS, FP16/BF16 36 PFLOPS and INT8 3 POPS. The [DGX B200 page](#) gives 64 TB/s of HBM3e bandwidth across eight GPUs. **Every per-GPU figure below is derived by dividing these system totals by eight.**

Per GPU (derived)	Dense TFLOPS	Bit length	TPP (derived)
B200, FP4	9,000	4	36,000
B200, FP8	4,500	8	36,000
B200, FP6	4,500	6	27,000
B200, FP16	2,250	16	36,000
B300, FP4	13,500	4	54,000
B300, FP8	4,500	8	36,000
B300, INT8	187.5	8	1,500

B300 shows why the "highest TPP across precisions" rule matters. Its FP4 rate rises relative to B200 while FP8 and FP16 stay flat, so FP4 sets its TPP. For B200, the derived per-GPU DRAM bandwidth is $64 / 8 = 8$ TB/s.

AMD's [MI355X specification page](#) gives per-GPU figures. It lists MXFP4 at 10.1 PFLOPs, MXFP6 at 10.1 PFLOPs, OCP-FP8 at 5 PFLOPs dense (10.1 with structured sparsity), FP16 at 2.5 PFLOPs dense, INT8 at 5 POPS dense, and 8 TB/s of HBM3E bandwidth.

MI355X precision	Dense TFLOPS	Bit length	TPP (derived)
OCP-FP8	5,000	8	40,000
FP16	2,500	16	40,000
MXFP4	10,100	4	40,400
MXFP6	10,100	6	60,600

The MI355X runs MXFP6 at the MXFP4 rate, so the 6-bit product is largest. Caveat: microscaling formats carry a shared block scale alongside the 4- or 6-bit elements (see our [NVFP4 vs MXFP4 vs FP8 comparison](#)), and the regulation text does not say how block-scaled formats count toward "bit length". Every reading lands far above 4,800, so 3A090.a.1 holds either way. AMD lists the process as "TSMC 3nm | 6nm FinFET", so every logic die would count toward applicable die area; die areas are not published there, so we did not compute PD, which would not change the outcome at this TPP.

For context on how these parts compare as hardware, see [H200 vs B200 vs B300 vs MI325X](#).

Where the rules stand (verified 2026-10-28)

We diffed the eCFR text of 3A090 and 4A090 at dates between 2024-01-01 and 2026-10-26 and matched each change to the Federal Register:

Date	Document	Effect on 3A090 / 4A090
2023-10-25 (eff. 2023-11-17)	FR 2023-23055	TPP/PD structure, .a/.b split, Note 2, NAC/ACA
2024-12-05 (eff. 2024-12-02)	FR 2024-28270	Added 3A090.c (HBM); notes renumbered
2025-01-15	FR 2025-00636 (AI Diffusion)	3A090.a/4A090.a license requirement made worldwide; License Exceptions AIA, ACM, LPP added; ACA scope widened
2025-01-16	FR 2025-00711	Note 1 to 3A090.a: foundry/OSAT presumption, transistor-count tests
2025-05-13	BIS press release	AI Diffusion rule rescission announced; enforcement officials instructed "not to enforce" it
2026-01-15	FR 2026-00789	Case-by-case review for China/Macau under TPP and DRAM-bandwidth caps (ECCN text unchanged)
2026-04-09	FR 2026-06851	Authorized IC designer date moved from April 13 to December 31, 2026
2026-05-31	BIS guidance	License still required for .a items to D:5/Macau-headquartered entities anywhere
2026-07-14 (eff. 2026-07-10)	FR 2026-14132	UAE moved to A:5; UAE written into the 3A090.b/4A090.b license requirement, with approved-entity carve-outs

The TPP and PD thresholds have not changed since November 2023; the destination and enforcement layer around them has.

One point needs care. The AI Diffusion text was never formally removed. On 2026-10-28 the eCFR (Title 15, up to date as of 2026-10-27) still codifies the worldwide license requirement for 3A090.a and 4A090.a in [§ 742.6\(a\)\(6\)\(iii\)\(A\)](#), the 4E091 model-weights requirement in [§ 742.6\(a\)\(13\)](#), and License Exceptions AIA, ACM and LPP in §§ 740.27 to 740.29. A Federal Register API search of BIS documents since May 2025 found no rescission rule; the only rule that mentions the rescission is the UAE rule. Its preamble sets out the BIS position is set out in the July 2026 UAE rule's preamble: BIS "is only enforcing this license requirement to destinations in Country Groups D:1, D:4, or D:5 (excluding those also specified in Country Groups A:5 or A:6),

and to entities headquartered in or with an ultimate parent headquartered in Country Group D:5 or Macau, wherever located" (FR 2026-14132). Have counsel confirm how much weight that gap between codified text and enforcement position can bear.

License Exceptions NAC and ACA

§ 740.8 defines two exceptions that cover 3A090 (except .c), 4A090 and the related .z items. Both exclude "items designed or marketed for use in a datacenter and meeting the parameters of 3A090.a", so in practice they serve datacenter parts in the 3A090.b band and non-datacenter parts at or above 4,800 TPP.

- **NAC (Notified Advanced Computing)** covers exports and reexports to Macau, Country Group D:5, or entities headquartered there. It is a notification, not a license. You file in SNAP-R at least 25 calendar days before shipping and include TPP, performance density, a datasheet showing how the item is designed and marketed, memory bandwidth, NAC and license history for the end user over the past 12 months, and cluster details (aggregate TPP, internal use or cloud). BIS either confirms eligibility or tells you a license is required.
- **ACA (Advanced Computing Authorized)** covers the rest of the world, plus in-country transfers within D:5 and Macau. It needs no prior notification.

Both require a written purchase order and are unavailable where a part 744 or 746 license requirement applies (with a narrow carve-out in § 740.8(b)). A supplier that cannot state TPP and PD cannot file NAC.

The January 2026 policy: TPP under 21,000, DRAM bandwidth under 6,500 GB/s

FR 2026-00789, effective 2026-01-15, did not reclassify anything. It changed the license review policy for exports from the United States to end users located in China or Macau. Commodities "with a TPP less than 21,000 ... and a 'total DRAM bandwidth' less than 6,500 GB/s (e.g., NVIDIA H200 or AMD MI325X)" moved from a presumption of denial to case-by-case review, if conditions are met. The applicant must certify that U.S. supply is sufficient and that foundry capacity is not diverted. Aggregate TPP shipped to China and Macau must be no more than 50% of that shipped to U.S. end users. The consignee must have KYC and remote-access controls, and a qualified U.S. third-party lab must review each shipment before export. Reexports to China, Macau or D:5, and exports to D:5- or Macau-headquartered entities located elsewhere, stay under a presumption of denial.

'Total DRAM bandwidth' is the aggregate bandwidth between the IC and DRAM, HBM and GDDR included. Checking the two named examples with our derived numbers:

Part	TPP (derived)	DRAM bandwidth	Under both caps?
H200 SXM	15,832	4,800 GB/s	Yes
MI325X	$2,614.9 \times 8 = 20,919$	6 TB/s = 6,000 GB/s (AMD)	Yes, by 81 TPP
B200	36,000	8,000 GB/s (derived)	No

Part	TPP (derived)	DRAM bandwidth	Under both caps?
MI355X	40,000 to 60,600	8 TB/s	No

The MI325X margin shows how closely the caps track specific products. Applications must also report 'interconnect bandwidth', 'copackaged DRAM capacity' and peak power at max TPP, and the third-party lab verifies them.

Buyer-side checklist for classification

1. **Get the ECCN from the manufacturer in writing**, for the chip and the finished system SKU. NVIDIA's [parts classification lookup](#) is searchable by part number but is "provided for informational purposes only" and "AS IS".
2. **Ask for TPP for every supported precision, dense**, so you can see which one governs. Don't accept a sparse headline number.
3. **Ask for applicable die area and process** whenever TPP is under 4,800. Without them, PD cannot be computed.
4. **Ask for total DRAM bandwidth** in GB/s if a China or Macau transaction is anywhere in the chain.
5. **Get the "datacenter design or marketing" position** in writing for workstation-class parts, because Note 2 depends on it.
6. **Write ECCN and license-exception representations into the contract**, with a duty to notify you if a rule change alters them.
7. **Screen ownership**. For .a items, an end user headquartered in D:5 or Macau triggers a license requirement wherever it is located.
8. **Re-check the eCFR at each order**. The timeline above lists eight changes to the rules around this entry since the October 2023 rule.

We can help with steps 2 to 4 as part of hardware selection through [GPU Smith](#). The classification itself stays with the vendor and your counsel.

FAQ

The datasheet only shows sparse numbers. What do I use? Halve them where the footnote says dense is half of sparse, as on NVIDIA's HGX page; Technical Note 2 requires dense.

Does 4A090 apply to a single server? Yes, if one IC inside meets 3A090.a. Thresholds apply per chip; the eight GPUs in a server are not added together. Aggregate TPP appears only in NAC filings and LPP's annual quota.

Is TPP the same as the APP number for supercomputers? No. BIS says 3A090/4A090 "control items based on the performance of a single chip", while the APP formula in 4A003 "describes aggregate performance across multiple chips" ([FR 2023-23055](#)).

Where do I verify the current text? The eCFR ([Supplement No. 1 to part 774, § 740.8](#)), then BIS guidance for enforcement positions the codified text does not reflect.

Sources

- [Implementation of Additional Export Controls: Certain Advanced Computing Items; Supercomputer and Semiconductor End Use \(FR 2023-23055\)](#) — BIS, Federal Register, 2023-10-25, accessed 2026-10-28
- [15 CFR part 774, Supplement No. 1 \(Commerce Control List\)](#) — eCFR, up to date as of 2026-10-27, accessed 2026-10-28
- [15 CFR § 740.8 License Exceptions NAC and ACA](#) — eCFR, up to date as of 2026-10-27, accessed 2026-10-28
- [Foreign-Produced Direct Product Rule Additions, and Refinements to Controls for Advanced Computing and Semiconductor Manufacturing Items \(FR 2024-28270\)](#) — BIS, Federal Register, 2024-12-05, accessed 2026-10-28
- [Framework for Artificial Intelligence Diffusion \(FR 2025-00636\)](#) — BIS, Federal Register, 2025-01-15, accessed 2026-10-28
- [Implementation of Additional Due Diligence Measures for Advanced Computing Integrated Circuits \(FR 2025-00711\)](#) — BIS, Federal Register, 2025-01-16, accessed 2026-10-28
- [Department of Commerce Announces Rescission of Biden-Era Artificial Intelligence Diffusion Rule, Strengthens Chip-Related Export Controls](#) — BIS, 2025-05-13, accessed 2026-10-28
- [Revision to License Review Policy for Advanced Computing Commodities \(FR 2026-00789\)](#) — BIS, Federal Register, 2026-01-15, accessed 2026-10-28
- [Extension of Authorized Integrated Circuit \(IC\) Designer Status and Application Deadline \(FR 2026-06851\)](#) — BIS, Federal Register, 2026-04-09, accessed 2026-10-28
- [Guidance Regarding Enforcement of License Requirements for Advanced Computing Items for Entities Headquartered in Country Group D:5 and Macau](#) — BIS, 2026-05-31, accessed 2026-10-28
- [Enhanced Favorable Treatment for the United Arab Emirates Under the Export Administration Regulations \(FR 2026-14132\)](#) — BIS, Federal Register, 2026-07-14, accessed 2026-10-28
- [NVIDIA H100 Tensor Core GPU](#) — NVIDIA, accessed 2026-10-28
- [NVIDIA H200 Tensor Core GPU](#) — NVIDIA, accessed 2026-10-28
- [NVIDIA HGX Platform](#) — NVIDIA, accessed 2026-10-28
- [NVIDIA DGX B200](#) — NVIDIA, accessed 2026-10-28
- [NVIDIA Hopper Architecture In-Depth](#) — NVIDIA Technical Blog, 2022-03-22, accessed 2026-10-28
- [AMD Instinct MI355X GPUs](#) — AMD, accessed 2026-10-28
- [AMD Instinct MI325X](#) — AMD, accessed 2026-10-28
- [NVIDIA Export Regulations: Parts Classification](#) — NVIDIA, accessed 2026-10-28

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. <https://www.ecfr.gov/current/title-15/subtitle-B/chapter-VII/subchapter-C/part-774/appendix-Supplement%20No.%201%20to%20Part%20774>
2. <https://www.federalregister.gov/documents/2024/12/05/2024-28270/foreign-produced-direct-product-rule-additions-and-refinements-to-controls-for-advanced-computing>

3. <https://www.federalregister.gov/documents/2023/10/25/2023-23055/implementation-of-additional-export-controls-certain-advanced-computing-items-supercomputer-and>
4. <https://www.nvidia.com/en-us/data-center/h100>
5. <https://developer.nvidia.com/blog/nvidia-hopper-architecture-in-depth>
6. <https://www.nvidia.com/en-us/data-center/h200>
7. <https://www.nvidia.com/en-us/data-center/hgx>
8. <https://www.nvidia.com/en-us/data-center/dgx-b200>
9. <https://www.amd.com/en/products/accelerators/instinct/mi350/mi355x.html>
10. <https://www.federalregister.gov/documents/2025/01/15/2025-00636/framework-for-artificial-intelligence-diffusion>
11. <https://www.federalregister.gov/documents/2025/01/16/2025-00711/implementation-of-additional-due-diligence-measures-for-advanced-computing-integrated-circuits>
12. <https://www.bis.gov/press-release/departments-commerce-announces-rescission-biden-era-artificial-intelligence-diffusion-rule-strengthens>
13. <https://www.federalregister.gov/documents/2026/01/15/2026-00789/revision-to-license-review-policy-for-advanced-computing-commodities>
14. <https://www.federalregister.gov/documents/2026/04/09/2026-06851/extension-of-authorized-integrated-circuit-ic-designer-status-and-application-deadline-to-become-an>
15. <https://www.bis.gov/media/documents/bis-guidance-may-31-2026.pdf>
16. <https://www.federalregister.gov/documents/2026/07/14/2026-14132/enhanced-favorable-treatment-for-the-united-arab-emirates-under-the-export-administration>
17. <https://www.ecfr.gov/current/title-15/subtitle-B/chapter-VII/subchapter-C/part-742/section-742.6>
18. <https://www.ecfr.gov/current/title-15/subtitle-B/chapter-VII/subchapter-C/part-740/section-740.8>
19. <https://www.amd.com/en/products/accelerators/instinct/mi300/mi325x.html>
20. <https://www.nvidia.com/en-us/about-nvidia/company-policies/export-regulations>

THE PUBLISHER

About GPU Smith

GPU Smith provides independent engineering and research for private AI infrastructure. We help technical buyers and operators reason about GPU workload sizing, hardware selection, procurement, deployment and inference operations, from the compute system to the networking, power and cooling requirements around it.

Start with the workload

Useful infrastructure decisions begin with the models, data, throughput, latency and operating constraints a team actually has. GPU Smith helps connect those requirements with system design and deployment choices. The goal is a privately controlled AI environment whose capacity and operational demands are understood before a purchasing decision.

Hardware and supplier research

Our public reference library covers GPUs, complete systems and networking components. The server-vendor directory and manufacturing research help teams investigate suppliers, compare options and follow sources behind technical claims. We distinguish manufacturer specifications from measured benchmarks and advertised capabilities from tested configurations. Reference pages are research resources, not a live inventory listing or a binding equipment quote.

Deployment and operations

GPU Smith's engineering scope includes infrastructure integration, networking, power, cooling and the practical operation of inference workloads. Our published articles explain the tradeoffs behind deployment, procurement and ongoing operations so teams can ask better questions and document decisions.

Work with GPU Smith

Explore the [hardware library](#), [GPU references](#), [networking references](#) and [vendor research](#). [Contact GPU Smith](#) with your workloads and deployment constraints to discuss a project and confirm current scope, pricing and availability.

Inclusion of a manufacturer or product in our research does not imply a partnership, certification or endorsement.

Website: <https://gpusmith.com>

This article is educational. Verify consequential medical, legal, financial, regulatory and technical decisions with appropriate professionals and the cited primary sources. Company services are described separately from the article's research findings.