Unpacking the H200 Export Policy

January 16, 2026
By Janet Egan and James Sanders

AI Chips for China

With two new policies, President Donald Trump has implemented his pledge to allow sales of NVIDIA's H200 AI chips to China in exchange for a quarter of the revenue. On January 13, the Bureau of Industry and Security (BIS) issued a rule permitting U.S. companies to sell advanced AI chips to China. The next day, President Trump issued a proclamation establishing a 25 percent import tariff on H200s and AMD's MI325X. Through combining multiple legal mechanisms, this effectively imposes a 25 percent export fee on these chips to China. While these chips are not the most advanced in America's AI ecosystem, they nevertheless provide a significant step up over what China can access domestically.

Advanced AI chips are the essential hardware behind frontier AI models. For years, U.S. export controls have restricted China's access to these chips, forcing Chinese AI labs to train models on less capable hardware. Chinese AI companies themselves cite this chip shortage as the critical bottleneck to their progress. These exports erode the key advantage that currently distinguishes American AI capabilities from China's.

The rule establishes a threshold capping exports to China at 50 percent of total U.S. sales for each chip. This means China could acquire almost 900,000 H200-equivalent chips—over twice what Chinese fabs are expected to produce this year and representing an extraordinary level of compute potential.

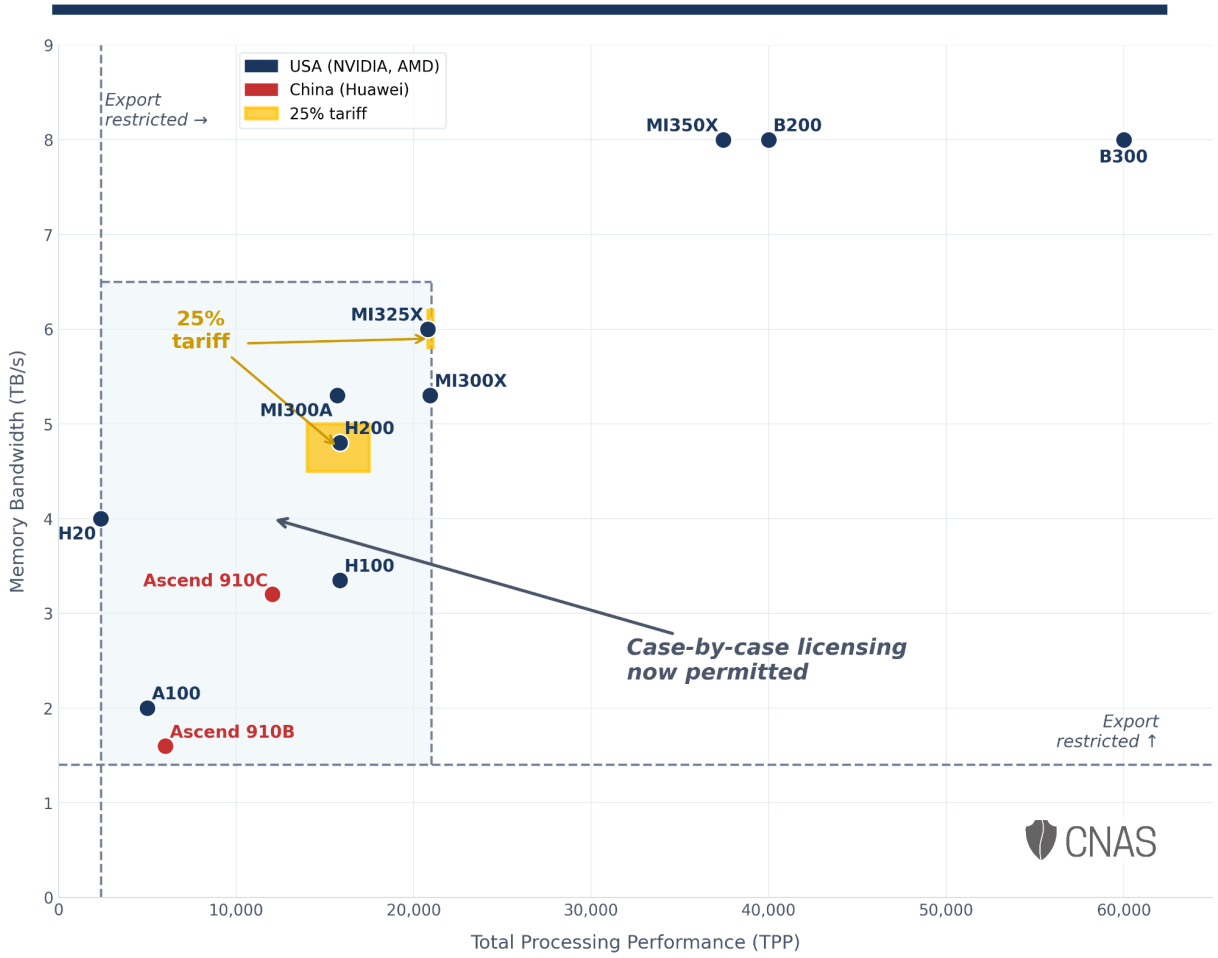
Additionally, it requires AI chip exporters to implement safeguards, including restrictions on end users of concern, mandatory testing in a third-party U.S. facility prior to export, and certification that any exports to China do not delay or reduce supply for U.S. customers. Yet, while these seem to promise strong safeguards, implementation carries significant risks, rendering these provisions almost entirely unenforceable, as outlined in the analysis below.

Ultimately, the new BIS rule functions as a statement of current policy, rather than a hard constraint: The administration retains discretion to change this rule at any stage and approve exports above the threshold. U.S. policymakers face difficult negotiations ahead as they attempt to balance the commercial ambitions of chipmakers against America's national security objectives. As our analysis shows, that balance is exceedingly difficult to maintain.

Scope and Volume of Chip Exports

The BIS rule permits exports of previously restricted chips that fall below specified memory and processing thresholds. The administration will implement the rule by licensing exports on a case-by-case basis. The subsequent proclamation sets the technical criteria for which chips will face a 25 percent tariff. Only two chips fall under both the rule and the 25 percent tariff: the NVIDIA H200 and the AMD MI325X (see Figure 1), indicating that they are positioned to receive the administration's green light for export.

Figure 1: Advanced Chips for China: Where the BIS Rule and Tariffs Intersect

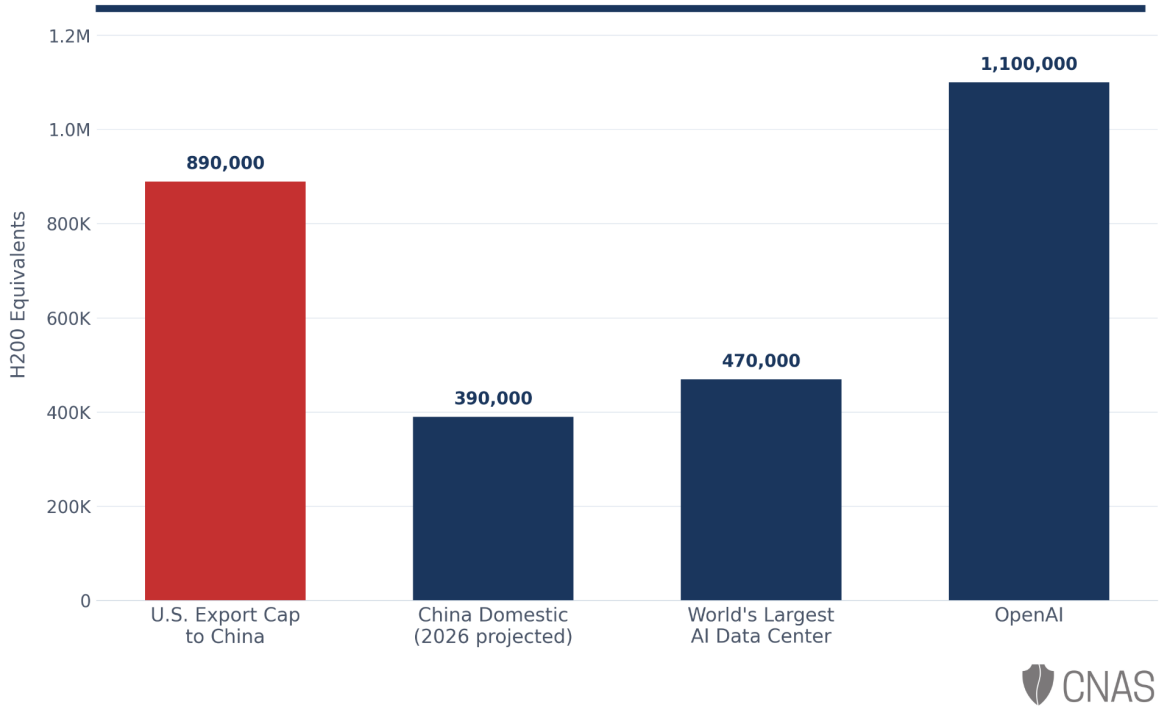


Source: Janet Egan and James Sanders. Analysis used data from Epoch AI, the new BIS rule, the presidential proclamation and associated annex, and previous export control analysis from Lennart Heim.

The rule caps total shipments to China at 50 percent of how many of each chip has already been bought for use in the United States. Because U.S. buyers shifted to more advanced chips in early 2025, cumulative U.S. purchases of H200s have largely plateaued, with NVIDIA estimated to have sold a total of two million H200s. Assuming 85 percent were sold to U.S.

customers, China could purchase up to 850,000 H200 chips under this threshold. Combined with AMD's MI325X chips, China could acquire the equivalent aggregate compute of nearly 900,000 H200s (see Appendix 1). While we assume that the administration will only approve licenses for chips where they collect the 25 percent fee, total exports to China could rise to the equivalent compute of 2.3 million H200s if the administration allows the export of all chips newly allowable under this rule (see Figure 1 and Appendix 1).

Figure 2: These Chips Provide a Significant Boost for China's AI Progress



Computing capacity measured in H200 chip equivalents. U.S. Export Cap to China reflects newly permitted advanced chip exports to China following the January 2026 BIS rule (Source: See Appendix 1). China Domestic represents China's estimated 2026 production capacity (Source: See Appendix 2). World's Largest AI Data Center is based on announced capacity as of December 2025 (Source: Epoch AI). OpenAI represents OpenAI's estimated total deployed compute worldwide as of October 2025 (Source: Epoch AI).

This is a significant amount of computing power: It's roughly twice what Chinese fabs are expected to produce domestically in 2026 (see Appendix 2), twice the capacity of the world's largest data center, and nearly OpenAI's entire deployed compute worldwide at the end of 2025. This is sufficient to train models matching or exceeding current American frontier models.

Restrictions on End Users of Concern

The rule requires exporters to certify that chips will not be accessed by prohibited end users, including military-linked and blacklisted entities. This may appear like a reasonable safeguard, but in practice, these certifications face significant verification challenges.

Chips exported to one entity can be easily redirected to another once inside China. Even for entities that appear to be legitimate, China's military civil fusion strategy removes meaningful distinctions between private sector and military entities. Shell companies present an additional challenge. In late 2024, researchers discovered that Huawei had likely used the company Sophgo as a front to procure nearly three million TSMC-manufactured chips for its Ascend 910B processor, illustrating how sanctioned entities can exploit intermediaries to access controlled technologies.

One partial remedy: BIS could require location verification for export-controlled chips, which could help flag diversion to military-affiliated facilities or data centers in restricted markets. Thankfully, NVIDIA has reportedly developed technology that starts to enable this.

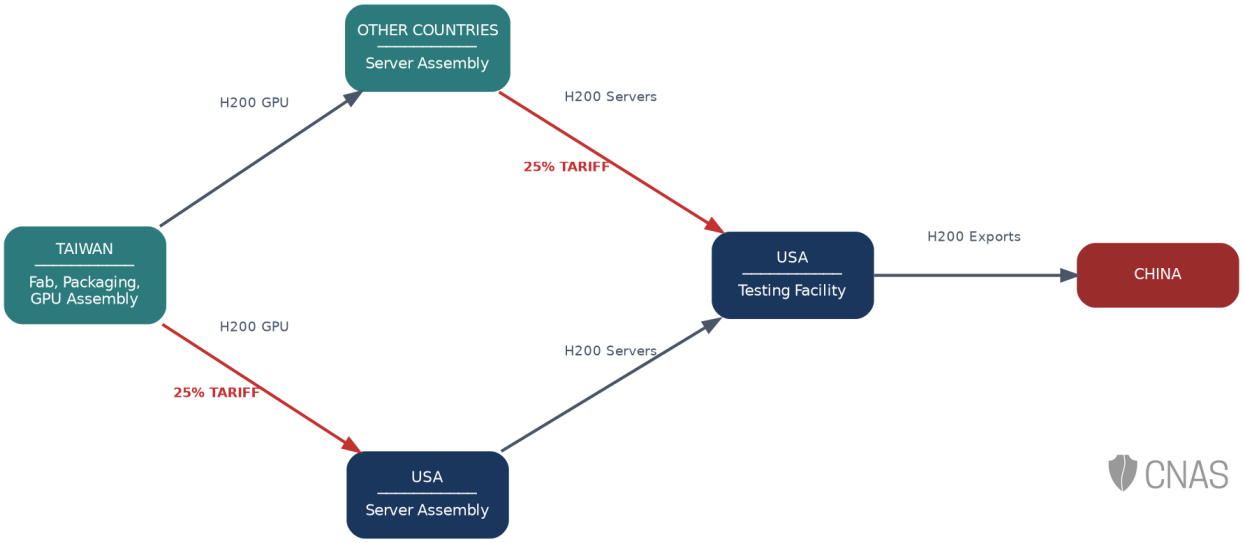
Even with workable location verification, however, it remains virtually impossible to prevent prohibited actors from accessing these chips remotely via the cloud. The rule seeks to address this through know your customer requirements: Exporters must obtain descriptions of customers' screening procedures and any remote end users from countries of concern. But the rule assumes Chinese companies like Alibaba will screen and report customers even when doing so conflicts with Chinese Communist Party directives. That is unrealistic, especially considering how Chinese law *requires* commercial firms to assist military and national security organizations.

Moreover, AI technology in China increasingly flows from private sector companies to military end users. A new dataset and report from the Center for Security and Emerging Technology found that nearly three-quarters of entities winning multiple AI-related contracts for the People's Liberation Army were civilian firms with no self-reported state ownership ties. We should expect this trend to continue, with AI blurring between civil and military models, even in democratic countries. The Pentagon has contracts totaling $800 million with xAI, Anthropic, Google, and OpenAI.

Third-Party Review and the 25 Percent Tariff

The new policy requires AI chips to undergo a third-party review by a U.S.-based testing facility before export. This serves a dual purpose: It creates an independent verification layer that confirms technical capabilities match the license application, and it might plausibly function as a workaround to constitutional restrictions on export tariffs. By requiring chips to pass through U.S. facilities before re-export, the administration can collect its 25 percent fee as an import tariff.

Figure 3: Stylized H200 Supply Chain: U.S. Tariff and Testing Touchpoints



Source: Janet Egan and James Sanders

America First Chip Access

The rule requires exporters to certify that China-bound shipments will not delay U.S. orders or divert foundry capacity from the domestic market. The intent is to prevent exports from constraining American supply or inflating U.S. prices. However, this condition seems impossible to credibly meet when critical parts of the AI chip supply chain are already oversubscribed.

The H200 relies on TSMC's four-nanometer fabrication process and advanced packaging—the same infrastructure used for its superior Blackwell line of chips. Even before this rule, TSMC's CEO stated that advanced-node capacity was "about three times short" of demand. High bandwidth memory (HBM), another critical component, also faces supply constraints. SK Hynix, which controls roughly 60 percent of the global HBM market, is sold out of HBM through 2026. Micron's HBM production is fully committed through 2026, and even so, it can only fulfill 50 to 66 percent of demand from its core customers. With critical inputs fully committed, it is difficult to see how exporters could credibly certify that shipments to China will not impact U.S. supply. Exporters should bear a significant burden of proof to demonstrate that their sales to customers in China do not divert capacity from those in the United States.

Notably, the rule only addresses U.S. supply. It does not account for whether China-bound shipments divert capacity from allies and partners. The United States will need to weigh exporting chips to China directly against America's AI exports program, which promotes full-stack AI technology sales to allies, and its Pax Silica initiative, which aims to secure AI supply chains among trusted partners.

China's Response: Uncertainty and Ambiguity

Beijing's response has been inconsistent. Rather than issuing a formal ban, Chinese customs authorities have reportedly instructed agents that H200s are not permitted to enter China, while also telling domestic tech companies not to purchase H200s "unless necessary" and reportedly discussing exemptions for research and development and universities. This approach of restricting through customs enforcement and informal guidance rather than published regulation has precedent: China's 2021 fertilizer export controls,which used customs inspections rather than explicit bans, allowed Beijing to calibrate restrictions in real time. Whether deliberate or not, this ambiguity serves multiple purposes: It maintains leverage ahead of the April Trump-Xi summit, sustains demand uncertainty that keeps domestic producers like Huawei growing, and preserves the option to cut off U.S. imports once China's ecosystem no longer needs them.

The Way Forward

Continued strong AI chip export controls remains the surest path to ensuring the United States leads the global AI transition. The implementation challenges outlined here make clear that balancing advanced chip exports to China with national security is exceptionally difficult. And these new policies set a concerning precedent. U.S. chipmakers will inevitably use this opening to push for exports of even more advanced chips. U.S. interests require discipline, not further concessions.

If U.S. compute continues growing near its historical 3x annual rate while chip exports to China remain fixed, then America's AI compute lead will widen. This would mitigate but not eliminate the national security harm from H200 exports. Should the administration follow this model for Blackwell-generation chips down the line, it would enable a quantum leap in China's AI capacity. Further relaxation risks providing China not just chips for domestic use, but sufficient capacity to export AI infrastructure and expand its Digital Silk Road, undermining America's AI exports program and leading to lost revenue for American companies and third countries becoming locked-in on the Chinese tech stack. Every advanced chip sent to China strengthens its AI ecosystem. Policy should make this harder, not easier.

Janet Egan is a senior fellow and deputy director of the Technology and National Security Program at the Center for a New American Security.

James Sanders is a research associate with the Technology and National Security Program at the Center for a New American Security.

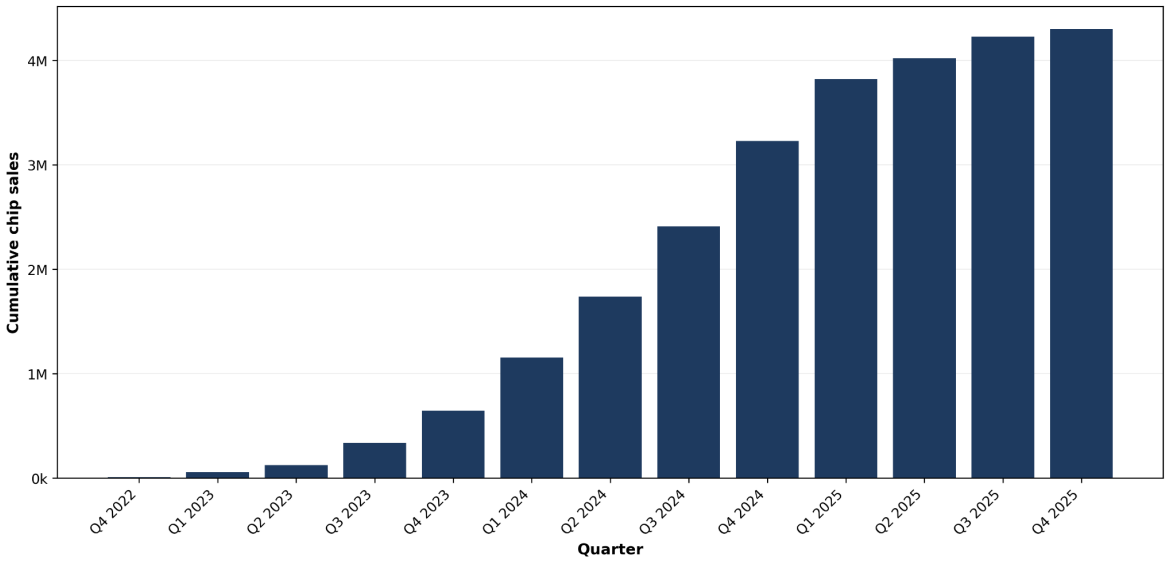
Appendix 1: American Compute

We use Epoch AI's data on chip sales to estimate aggregate compute from sales of different types of chips. They estimate that, as of the end of 2025, NVIDIA sold 4.3 million Hopper chips (including H100s and H200s),with sales peaking around the end of 2024 before dropping sharply, leading to a plateau of cumulative installed Hoppers over the past year (see Figure 4 below). Given that the H200 wasn't commercially deployed until the second quarter of 2024, we can estimate that out of these 4.3 million Hoppers, two million were H200s. Similarly, they estimate that AMD has sold enough MI325Xs to equal the compute of 97,000H200s (when normalized by dense 8-bit FLOP/s).

To estimate the proportion of compute that was sold for end use in America, we use Epoch AI's report on AI supercomputers. As of early 2025, they estimate that America had 74.4 percent of the world's compute, with the rest of the world (excluding China) having 11.5 percent, implying that America has ~85 percent of the compute that isn't in China. Combining these estimates implies that 1.8 million H200-equivalents of compute were sold for end use in America as of January 2026, implying that China could currently buy 890,000H200-equivalents of compute. We assume this number will not increase much going forward, as cumulative sales of the H200 and MI325X have mostly plateaued in America due to the availability of more advanced chips (see Figure 4).

We assume that America will only export the H200 and MI325X, since these are what the tariff policy most incentivizes. However, if America were to export all the types of chips newly allowable under this rule (including H100s and other AMD chips), we estimate that China could import 2.3 million H200-equivalents worth of compute.

Figure 4: Cumulative Sales of Hopper Chips (H100s and H200s) Have Plateaued



Source: Epoch AI data on chip sales

Appendix 2: Chinese Compute

For estimating China's domestic chip production from 2026, we took estimates from Semianalysis and *Bloomberg*. Semianalysis estimates that China will be able to produce 300,000 Huawei 910C in 2026. Since Semianalysis' estimate is based on China being bottlenecked by HBM, and Huawei using all accessible HBM, we use this as an estimate for all of China's compute production in 2026. *Bloomberg* estimates that Huawei will produce 600,000 910Cs and 100,000 910Ds. We convert these figures into H200-equivalents by normalizing by total processing performance (TPP). We assume that the 910D will have comparable TPP to the H100 based on early reporting. We then average the estimates from Semianalysis and *Bloomberg* to get a final estimate of China producing the equivalent compute in 2026 of 390,000 H200s. However, there is still significant uncertainty from the estimates of both Semianalysis and *Bloomberg*, and production could be significantly higher if China solves HBM bottlenecks.