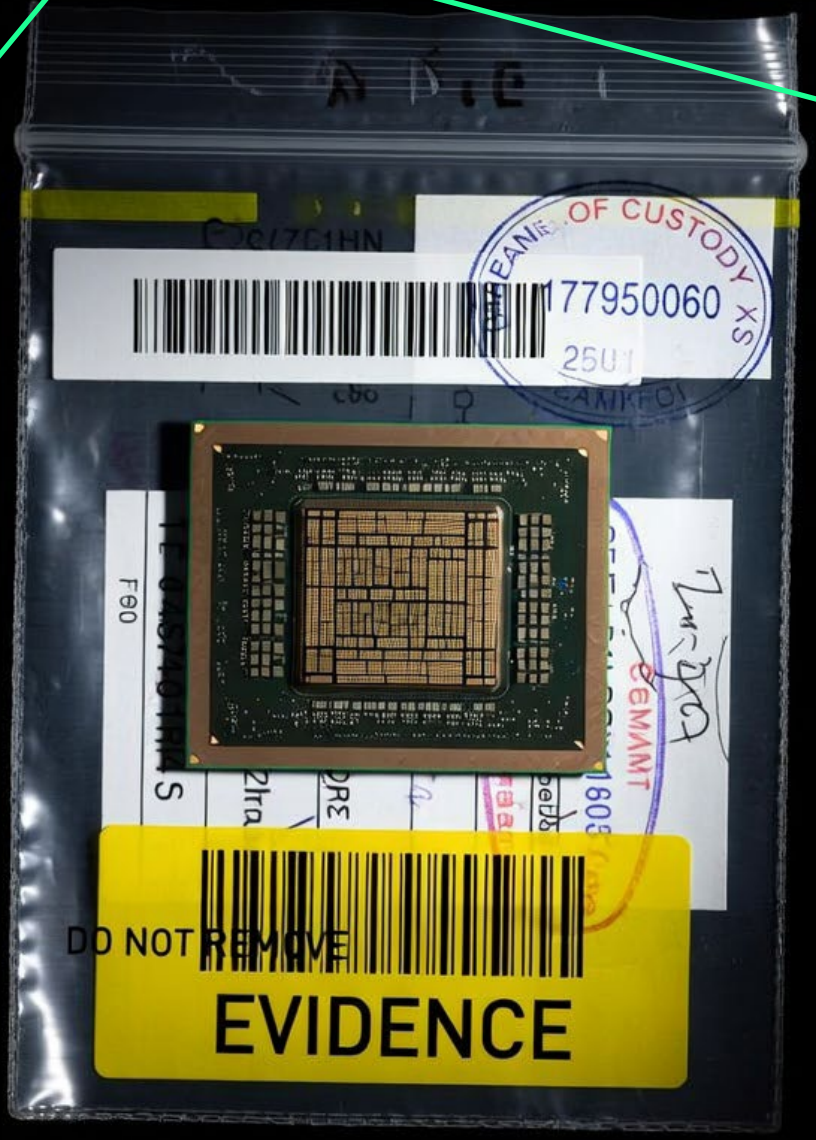
OVERHAUL

CARGO CRIME SERIES

#4 | July 2026



When the chips are down

The chain of custody in AI infrastructure

Artificial intelligence has become one of the defining technologies of the twenty-first century,

and for years, the conversation has centred on algorithms, models, data governance, and the race to build ever more powerful systems. Beneath that discussion lies a less visible touchstone; the physical infrastructure that powers it. **Advanced graphics processing units (GPUs), AI accelerators, high-bandwidth memory, networking equipment, and hyperscale servers are now some of the most valuable and strategically important assets on earth.**

The landscape has changed irrevocably, and hardware is no longer just tech. It is geopolitical leverage. It is national security infrastructure. A strategic asset caught between governments, regulators, hyperscalers, organized crime groups, and a rapidly expanding black market. As regulators seek to control the movement of this technology, **the concept of 'chain of custody' has emerged as one of the most critical components of AI governance.**

In our first paper, 'Inside the Billion Dollar Black Market of Stolen Supply Chains', we focussed on how illicit enterprises operate in the shadows. Our second paper, 'The Attraction of AI Infrastructure, The New Gold Rush and the Modern Outlaw,' looked at the infrastructure of AI at the height of the new gold rush. It examined how high value, high demand, high consequence physical components, coupled with weak physical security and supply chain oversight, were drawing the attention of organised crime groups. Our third paper, 'The Hybrid Heist, The Physical and Digital Convergence in Cargo Crime', examined the marriage of the physical and cyber criminals and the consequences for the global supply chain. In this paper, 'When the Chips Are Down, The Chain in AI Infrastructure,' we look at how the chain of custody in AI infrastructure is one of the most crucial but often misunderstood sequences in AI governance.

What is chain of custody?

Essentially, chain of custody is the **detailed, documented record showing who handled an asset, when they handled it, where it was located, and under what authority it moved.** While long established in other sectors such as pharmaceuticals, nuclear materials, and forensic evidence management, that concept is now becoming absolutely essential for the AI sector. Without it, trust collapses, and the consequences are extraordinary. In an environment where a single shipment of advanced AI servers can be worth hundreds of millions of dollars, the ability to demonstrate where hardware has been, who handled it, and whether it has been diverted **has become a matter of regulatory compliance and national security.** This was seen recently in the US, when on March 26 of this year, the **U.S. House Foreign Affairs Committee advanced the Chip Security Act**, legislation that requires advanced chips to embed technical security and location-verification mechanisms. The proposal reflects a growing recognition among policymakers that traditional, paperwork-based chain-of-custody controls may be insufficient for monitoring the movement of export-controlled AI hardware at scale.[1]

Chairman Mast, HFAC, advances Chip Security Act

> “Small distinctions become critically important when it comes to determining responsibility for theft, diversion, or export-control violations”.

Considering an AI hardware supply chain may involve semiconductor fabrication facilities, original equipment manufacturers (OEMs), freight forwarders, third-party logistics providers (3PLs), customs brokers, distributors, cloud providers, colocation facilities, and end users, it’s common that at each stage ownership and custody may be held by different entities. For example, a manufacturer may own the hardware, while a logistics provider physically controls it, or a cloud provider may operate hardware that is leased rather than owned. These small distinctions become critically important when it comes to determining responsibility for theft, diversion, or export-control violations.

Bailees and supply chain security

Third-party logistics providers occupy an extremely delicate position within this chain. Most infrastructure relies on transportation networks involving warehouses, freight forwarders, customs intermediaries, and carriers. Legally, these firms often act as bailees; possessing goods temporarily without taking ownership. **Advanced AI hardware carries enormous value and strategic significance, and as a result, any failure in custody can have serious consequences.** Every custody transfer presents an opportunity for theft, diversion, espionage, or sanctions and evasion, and for some, the financial incentives are worth the risk. Advanced AI accelerators such as NVIDIA’s H100 and H200 GPUs can command prices of tens of thousands of dollars per unit, while complete AI server racks may be worth millions. Their high value density, and strong global demand, has contributed to the emergence of black markets and organized criminal networks.

As outlined in our previous papers, the evidence suggests the growth of illicit trading networks incorporating offshore shell companies, third-country distributors, secondary brokers, and complex procurement arrangements designed to obscure the ultimate destination of hardware. **Complexity creates opportunity.** Cargo theft, shipment diversion, fraudulent documentation, and carrier impersonation schemes have all become growing concerns as demand for restricted AI chips increases. The modern AI gold rush is not taking place inside chatbots or foundation models. It is happening inside data centres, semiconductor fabrication plants, logistics hubs, cargo aircraft, and warehouses across the globe.

Massive fines & extraordinary consequences

The scale of the problem became undeniable in 2025 and 2026. Between late April and mid-May 2025 alone, at least $510 million worth of servers carrying controlled US AI technology were diverted in violation of export regulations. Further to that, in December 2025, federal prosecutors secured the first-ever conviction in an AI chip smuggling case: a $160 million operation that moved export-controlled NVIDIA H100 and H200 GPUs to China using falsified shipping documents. The operator faced up to ten years in federal prison[2]. That case was quickly overshadowed when, in March 2026, the co-founder of a major IT company was indicted for allegedly directing the diversion of $2.5 billion in Nvidia-powered servers to China via Southeast Asian intermediaries.[3]

The Bureau of Industry and Security (BIS) responded with their full force. Congress approved a 23% increase in BIS's Fiscal Year 2026 budget, with specific funding earmarked for semiconductor enforcement. In February 2026, BIS announced a $252 million settlement for export control violations by a materials engineering company. In January 2026, a separate $1.5 million settlement targeted an in-country transfer to a blacklisted Chinese foundry.[4]

Export, control, and strategic commodities

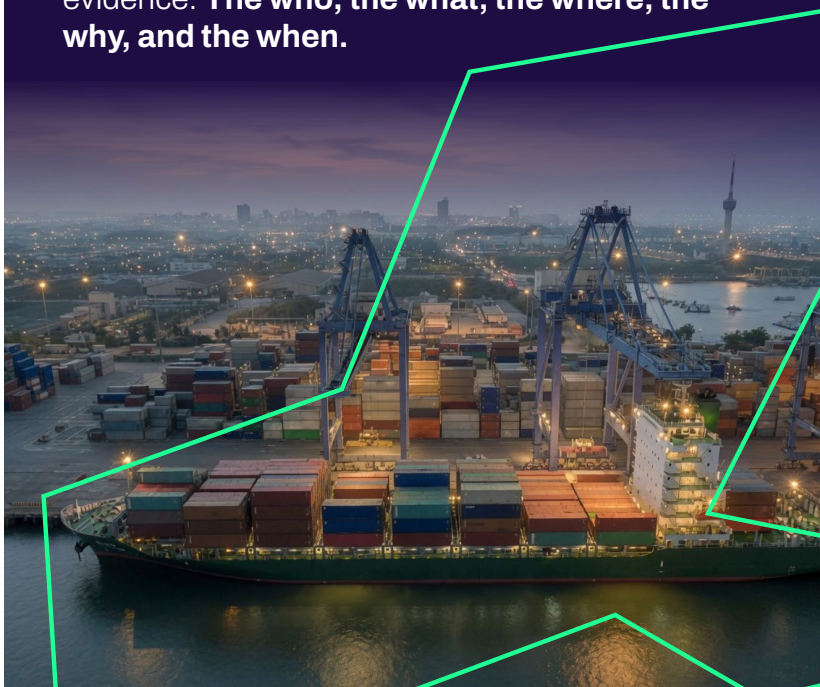
Historically, export controls focused on weapons systems, cryptography, and sensitive military technologies. AI infrastructure has now joined that list.

Advanced AI chips are highly valuable, globally desired, portable, and increasingly difficult to acquire through legitimate channels.

Advanced semiconductors have become strategic commodities, sitting alongside energy resources, telecommunications infrastructure, and critical minerals, as valuable assets governments are determined to control. The problem is in the movement. A shipment leaving a manufacturing facility may pass through freight forwarders, warehouse operators, customs brokers, and local carriers before arriving at its destination. Each handoff represents a potential point of failure in the chain, or a point of opportunity for an interception and another opportunity for manipulation, diversion, theft, or misrepresentation.

The challenge becomes even more complex once hardware reaches a data center. Modern AI deployments often involve multiple parties across multiple sites, including hardware manufacturers, systems integrators, co-location operators, hyper scale cloud providers, and customers leasing compute capacity. **If AI hardware is stolen or misplaced after delivery, determining liability becomes a minefield.** Investigators, insurers and lawyers may assess ownership, contractual custody arrangements, and potential negligence through a fact-specific investigation of risk allocation, transfer-of-title provisions, meticulous custody and control arrangements, insurance obligations, tight security requirements, and the circumstances surrounding the loss. They may examine shipping records, warehouse logs, access controls, surveillance footage, audit trails, and communications between parties to determine who had possession or responsibility for the hardware when it went missing. At that point, any breaches of contractual or regulatory obligations can be identified. Where export controls apply, what begins as a commercial dispute may evolve into a broader regulatory investigation.[5]

Since 2022, the Bureau of Industry and Security (BIS) has expanded restrictions on advanced AI infrastructure through successive rounds of export controls.[6] The January 2025 Framework for Artificial Intelligence Diffusion accelerated that effort, introducing worldwide licensing requirements for advanced AI chips and certain AI model weights while tightening controls on overseas transfers and third-country intermediaries.[7] The objective is straightforward: **stop advanced infrastructure from reaching restricted jurisdictions.** The chain itself now becomes evidence. **The who, the what, the where, the why, and the when.**



Every link in the chain

In many respects, **AI infrastructure resembles another heavily regulated resource: nuclear material.** Not because of its destructive capability, but because of its strategic value. Such is the urgency around keeping the chain intact, policymakers now view advanced infrastructure as dual-use technologies capable of delivering both commercial and military advantages. As a result, any link in the chain, be that a distributor, a cloud provider, or goods inward on site at the data center are now expected to execute enhanced due diligence. Documentation, monitoring, due diligence, customer verification, and end-use validation are non-negotiable within the AI infrastructure ecosystem and the ability to account for every last chip, every server and every single movement through the life cycle is now a military style operation.[8] If restricted AI hardware is discovered in a prohibited jurisdiction, regulators are likely to investigate every participant in the supply chain. Companies must be able to show that enhanced due diligence by implementing safeguards to prevent diversion, customer vetting, shipment monitoring, and end-use verification procedures. Proposed location-verification technologies, embedded tracking mechanisms, enhanced cloud-computing controls, and post-sale monitoring requirements all point toward the same destination: **continuous and unequivocal accountability.**

> *"When something goes wrong and there is a loss, it is often a mess to understand which party is liable and will suffer the loss if the goods aren't fully insured. With goods changing hands so many times in most cases,* ***there are many contracts and chain of custody documents*** *to sort through to determine where the risk of loss is. Those documents often raise more questions than they answer. It is so much easier and cost effective to prevent the loss on the front end."*
>
> **Shannon Fohn**
> General Counsel, Overhaul

Enter 'compute smuggling'

The emergence of cloud computing has introduced another layer of complexity. Policymakers have become increasingly concerned about what some analysts describe as 'compute smuggling'[9] which is the ability to access export-controlled AI hardware remotely. It has been reported that some Chinese firms and other restricted entities have sought access to advanced computing resources by renting server capacity from offshore cloud providers and third-country data centers. The response by lawmakers were proposals to extend export controls beyond physical hardware to encompass remote access to advanced infrastructure, as seen by the passing of the Remote Access Security Act, by the U.S. House of Representatives in 2026. **Export controls are evolving from border-focused restrictions into comprehensive governance frameworks covering procurement, transportation, deployment, cloud access, and end-of-life disposition.**

End of life care

The most underexamined chain of custody gap is at hardware end-of-life.

AI infrastructure turns over rapidly. GPU generations cycle in 18-24 months as operators chase performance but decommissioning practices have not kept pace. NIST SP 800-88 Revision 2 (updated September 2025) requires chain-of-custody documentation to record personnel identity, asset condition, and timestamp at every handling point, down to per-GPU serial number tracking and NVMe drive manifesting. **Yet the IT Asset Disposition (ITAD) market, valued at $17.5 billion globally in 2025, remains fragmented**, with wide variance in how rigorously operators apply these standards.[10]

Ultimately, **chain of custody has become the most important component in AI governance.** It is no longer simply a logistics or compliance exercise. It is the mechanism by which organizations demonstrate that advanced infrastructure is where it is supposed to be, operated by authorized parties, and protected from diversion into unauthorized markets.

As AI infrastructure becomes more valuable and strategically significant, the ability to maintain an unbroken chain of custody is as important as cybersecurity itself.

In the coming decade, the organizations that can effectively document, monitor, and secure the movement of AI hardware will not only reduce regulatory risk but also gain a competitive advantage in a market increasingly defined by trust, transparency, and national security considerations. **The infrastructure race is no longer just a race for performance. It is a race for control.** And in a world where AI has become power, chain of custody will prove to be one of the most important governance mechanisms of all.

For leaders who ask the question, "How do I protect the chain?"

Overhaul is the answer.

Committed to continuity, resilience and security as cargo moves across the globe, we support the world's most critical supply chains with data driven intelligence — keeping our customers in control.

Overhaul's platform combines SKU-level visibility, predictive risk intelligence, and compliance monitoring to offer true dynamic control over goods in motion and the certainty that every SKU is tracked, protected, and delivered as planned.

Trusted by companies across industries, Overhaul supports global freight operations for major brands like Microsoft, Google, Apple and Dell. With Overhaul's SaaS platform, organisations can monitor sensitive, high value shipments anytime, anywhere. Advanced risk management tools help detect and prevent potential threats before they escalate. Overhaul collaborates with law enforcement to support cargo theft recovery and leverage insights from global intelligence partners to keep customers informed of emerging risks.

For a free demo or to speak with one of representatives' contact us https://over-haul.com/request-a-demo/

[1] https://www.congress.gov/bill/119th-congress/house-bill/3447/text
[2] https://introl.com/blog/nvidia-160m-smuggling-operation-gatekeeper-december-2025
[3] https://tech-insider.org/super-micro-nvidia-chip-smuggling-china-2026/
[4] https://www.gtlaw.com/en/insights/2025/12/navigating-gpu-export-controls-and-ai-use-restrictions-in-data-center-operations
[5] See https://www.uniformlaws.org/acts/ucc. See also: AI in Supply Chains: Perspectives from Global Thought Leaders - Google Books
[6] https://public-inspection.federalregister.gov/2025-00636.pdf
[7] https://www.gtlaw.com/en/insights/2025/12/navigating-gpu-export-controls-and-ai-use-restrictions-in-data-ce https://labs.cloudsecurityalliance.org/research/csa-research-note-nsa-allied-ai-supply-chain-security-guidan/nter-operations
[8] Sidley Austin – New U.S. Export Controls on Advanced Computing Items and Artificial Intelligence Model Weights Covington & Burling – Export Control Framework Limiting the Diffusion of Advanced AI WilmerHale – BIS Issues Long-Awaited Export Controls on AI Simpson Thacher – BIS Announces Worldwide Export Controls on Advanced Chips and AI Models Council on Foreign Relations – What to Know About the New U.S. AI Diffusion Policy and Export Controls Tom's Hardware – Remote Access Security Act and Offshore AI Compute Access Reuters – Trump Administration Recommends Location Verification for AI Chips
[9] https://www.bruegel.org/analysis/europe-needs-strategy-close-artificial-intelligence-compute-gap
[10] https://www.stselectronicrecyclinginc.com/ai-gpu-data-center-itad-2026