Location Verification for AI Chips

This document is informed by interviews with the following professors: Ari Juels (Cornell), Yuval Shavitt (Tel-Aviv University), AbdelRahman Abdou (Carleton University), Katharina Kohls (Ruhr University Bochum), Nicolas Christin (CMU), Yong Liu (NYU Tandon), Zachary Weinberg (CMU). The document itself was not authored by them and should not be interpreted as a statement of their views, though their insights contributed to the content.

Executive summary

A **delay-based location verification system** for AI chips offers a practical, implementable solution to strengthen export control enforcement. This technology can help detect when high-performance AI chips are diverted to restricted countries by verifying their physical location through internet-based measurements.

Building on published research from Brass & Aarne (2024),[1] we outline how a location verification scheme for high-performance AI chips could practically be implemented. We focus on a delay-based verification system that measures round-trip communication time between AI chips and landmark servers to determine location. This method provides strong security against manipulation while leveraging existing chip capabilities.

This document outlines:

1. A high-level technical solution for implementing delay-based location verification
2. Methods to ensure geographic accuracy and security against potential attacks
3. Requirements from chip designers, users, and stakeholders for successful deployment

The verification system can serve as a scalable enforcement mechanism that complements existing export control frameworks while providing authorities with actionable intelligence about potential violations.

Our proposed system:

- **Builds on established literature.** Location verification leverages established research from academic and corporate sources previously developed for cloud data storage and region-locked content. These technical protocols can be adapted directly for AI chip geolocation without extensive new development.

[1] *See also* Tim Fist, Tao Burga & Vivek Chilukuri, *Technology to Secure the AI Chip Supply Chain: A Primer* at text accompanying nn. 67-70, Ctr. for a New Am. Sec. (Dec. 11, 2024); Onni Aarne, Tim Fist & Caleb Withers, *Secure, Governable Chips: Using On-Chip Mechanisms to Manage National Security Risks from AI & Advanced Computing* at 11-12, Ctr. for a New Am. Sec. (Jan. 8, 2024).

- **Relies on existing infrastructure.** The system can use existing cloud infrastructure and content delivery networks rather than building new hardware. Network security companies with monitoring and threat detection expertise can readily support landmark network operations.
- **Achieves sufficient accuracy for anti-smuggling purposes.** Research demonstrates location estimates with median error under 100km (Kohls et al., 2022), adequate for determining whether chips remain within permitted countries. Accuracy depends on landmark density and regional network characteristics, requiring careful design and testing of the landmark network to ensure reliable performance.
- **Defends against common evasion tactics** including false location reporting, VPN usage, artificial delay manipulation, and limited landmark compromise. Implementation requires collaboration with chip designers to integrate secure verification protocols into firmware with tamper-resistance features.
- **Would be even more effective if combined with a chip registry and/or other on-chip mechanisms.** A registry tracking ownership and expected locations makes suspicious movements easier to detect and investigate. On-chip mechanisms that throttle performance when verification fails create powerful compliance incentives.

Implementation Details

Delay-based location verification follows a seven-step process from network deployment to analysis and validation:

1. Landmark server network deployment
2. Landmark calibration
3. Configuration of AI chip's environment to enable connection to landmarks
4. Chip authentication and connection to landmarks
5. Delay measurement
6. Location estimation
7. Repetition of location verification and secondary analysis

1. Landmark server network deployment

Landmark servers require verifiable locations and secure timestamp capabilities. These servers can be deployed by renting existing infrastructure from cloud providers, significantly reducing implementation costs. Strategic deployment should prioritize:

- **Coverage in high-volume export destinations** where significant numbers of AI chips are legally shipped, ensuring comprehensive monitoring capability.
- **Increased density near borders of restricted countries** to distinguish between permitted and restricted locations with sufficient precision. Border regions between Taiwan/China and South Korea/North Korea require particularly dense coverage due to smuggling risks.

- **Consideration of regional network topologies** to prevent false readings from unusual routing patterns. In regions with less developed infrastructure, ISPs may route domestic traffic through external points, creating misleading delay measurements.
- **Selection of data centers with direct ISP connectivity** rather than those using private backbone networks that can introduce artificial latency. Direct connections to chip users' ISPs are especially important for measurement accuracy.

2. Landmark calibration

Landmarks must establish accurate delay-to-distance mappings to determine chip locations. These mappings convert communication delays into physical distance estimates and require regular updates to account for changing network conditions. Calibration accuracy depends on landmark location, local network topology, and overall delay magnitudes.

The proposed calibration method builds on Sheng et al. (2024), where landmarks measure delay-distance pairs by pinging other landmarks at known locations. These measurements create a statistical mapping that estimates maximum possible distance for each delay measurement, using a curve where approximately 95% of observed data points fall below the threshold. This approach provides reliable distance estimates even with normal network variability.

3. AI chip environment configuration

Chip owners must configure their networks to enable verification communications. Standard security measures like firewalls typically block external pings needed for verification. Owners in participating countries need clear notification that their high-performance AI chips require regular landmark communication to verify compliance with export controls.

Effective implementation requires appropriate compliance incentives. While regulatory sanctions provide baseline enforcement, on-chip mechanisms could **throttle or deactivate chips after failed verification attempts**. Such technical enforcement should only be deployed after thorough validation of the verification system's accuracy and security, potentially requiring hardware-level tamper resistance in future chip designs.

4. Chip authentication and connection

Each AI chip requires a secure digital identity for verification. Chips initiate the process by establishing encrypted TLS connections with landmark servers, requiring owners to have access to at least one landmark address. Multiple landmark addresses increase resilience against network outages.

The authentication system uses **signed certificates issued during manufacturing** and stored in on-chip Trusted Platform Modules (TPMs). These certificates enable mutual authentication between chips and landmarks while ensuring encrypted communications. When integrated with

a chip registry, these digital identities link location data to specific chips and their ownership history.

Most AI accelerators (GPUs) cannot directly perform verification due to their specialized architecture. Instead, a CPU handles the verification protocol on behalf of the accelerator. **Security measures must prevent the CPU from duplicating or remotely accessing the GPU's authentication credentials.** The CPU should only relay signed messages from the accelerator without generating them independently. Newer AI chips with co-located CPU/GPU components provide additional security by limiting the physical distance between the components.

5. Delay measurement

Accurate distance estimation relies on precise network delay measurements. After initial connection, the first landmark shares the chip's address with other landmarks within range of the chip's claimed location. Each landmark establishes a secure connection and performs multiple ping measurements to calculate the Round Trip Time (RTT).

The system measures the time between sending a ping and receiving a response, with half the RTT providing an estimated one-way delay. This approach assumes symmetric network speeds, though more sophisticated methods (Abdou et al., 2015) can account for directional differences. To minimize the impact of transient network conditions, **landmarks perform multiple measurements in sequence** (typically 20 as in Sheng et al., 2024) and use the minimum RTT for final distance calculations.

6. Location estimation

Multiple landmark measurements enable precise triangulation of chip location. Each landmark converts its minimum delay measurements into distance estimates using the calibration mappings established earlier. These distance estimates are cryptographically signed and shared with other landmarks to ensure integrity.

The system combines multiple distance measurements to determine the chip's likely location through triangulation. Current research supports several approaches, with the method outlined by Sheng et al. (2024) particularly promising for AI chip verification. This approach **identifies a probable region rather than a single point**, offering flexibility in landmark placement unlike more rigid geometric configurations. It also incorporates **Byzantine fault-tolerance**, maintaining accuracy even when some landmarks are compromised or unreliable.

7. Verification analysis and validation

Repeated measurements over time reduce false positives and strengthen enforcement credibility. Network variations occasionally produce inaccurate location estimates that could wrongly suggest export control violations. To prevent unjustified enforcement actions, the

system should compare multiple measurements taken over hours or days, flagging chips only when a consistent pattern of suspicious readings emerges.

Verification can be further strengthened through secondary analysis methods:

- **Network path analysis** using traceroute and DNS records to verify consistency with measured distances
- **IP topology database cross-referencing** to validate location estimates against known network configurations
- **ISP consultation** to gather network topology information from service providers and chip users

These additional validation techniques help authenticate suspicious readings before triggering formal investigations, balancing effective enforcement with fairness to compliant chip owners.

Effectiveness

The location verification system must deliver **accurate, reliable location estimates** to support enforcement actions and maintain **security against evasion attempts**. This approach specifically addresses anti-smuggling use cases where we verify that chips remain within permitted countries rather than being diverted to restricted locations.

Accuracy Capabilities

Research demonstrates country-level accuracy sufficient for export control enforcement. The system needs precision in the range of tens to low hundreds of kilometers—enough to determine whether a chip remains within its permitted country. Current research shows impressive results:

- Accuracy within a 100km radius achieved in Europe using 80 landmarks sampled from a 3,500-node network (Kohls et al., 2022)
- Similar precision demonstrated in the US using 450 landmarks (Sheng et al., 2024)
- Ongoing research shows potential for reducing landmark requirements while maintaining accuracy (Cho et al., 2024)

Regional network characteristics affect achievable precision. Network infrastructure in different countries creates varying challenges for location verification. For example, China's Great Firewall and hierarchical network topology may produce longer delays or unexpected routing patterns. However, these challenges primarily affect verification precision within restricted countries, not the system's ability to detect when chips leave permitted locations.

For anti-smuggling purposes, the primary goal is confirming chips remain in permitted countries, not precisely locating them within restricted ones. So long as accuracy in permitted countries meets standards—which requires empirical validation—the system can achieve its enforcement objectives.

Future implementations could incorporate secure GPS for enhanced precision. Authenticated GPS signals like those from Galileo's OS-NMA could complement network measurements with tamper-resistant positioning data. While this would require additional cooperation from chip owners (installing antennas), it could significantly improve accuracy in border-adjacent facilities where network-based verification faces greater challenges.

Security

The verification system must defend against both evasion and disruption attempts from sophisticated adversaries. Given the strategic importance of AI hardware, threat actors may include nation-states with significant resources. The system's security architecture addresses multiple attack vectors falling into two broad categories: evasion resistance and system resilience.

Evasion Resistance

The system inherently defends against false location claims by measuring network delays rather than trusting reported positions. When chips are powered on and connected, adversaries cannot simply declare false locations. For periods when chips are legitimately offline, the verification protocol includes reporting requirements for events like resale or decommissioning to prevent undocumented diversion.

VPNs, proxies, and artificial delay manipulation cannot defeat the verification system. These techniques typically increase network delays, creating wider uncertainty ranges in location estimates rather than helping adversaries appear to be in permitted countries. The system can flag suspicious measurements or implement maximum delay thresholds that trigger re-verification.

Sophisticated attacks using specialized infrastructure face practical limitations. While theoretically possible, attacks using dark fiber or satellite links to artificially decrease distances require significant resources, specialized knowledge, and risk detection through secondary validation techniques.

System Resilience

The landmark network design provides robust defense against compromise attempts. Using the approach from Sheng et al. (2024), the system maintains accuracy even when a fraction of landmarks are compromised. Landmarks can operate within secure enclaves with remote attestation capabilities to verify their measurement protocols remain uncorrupted.

Distributed landmark deployment mitigates denial-of-service attacks. A large, decentralized network across multiple providers and regions makes comprehensive DDoS attacks prohibitively expensive. Regular measurement repetition further reduces vulnerability since most DDoS attacks last only hours.

Border-adjacent installations may require enhanced verification measures. For data centers near restricted country borders, the system can deploy higher landmark density or supplement network measurements with secure GPS verification for greater precision.

Protection against physical tampering relies on secure hardware features. While determined attackers might modify verification software, the protocol reduces tampering value by using server-side timestamps and secure communication channels. Future implementations could incorporate additional hardware-based security measures as chip designs evolve.

This system has certain limitations, such as adversaries using cloud-based AI chips in permitted countries or indigenously produced hardware. These scenarios fall outside the scope of the proposed verification approach and would require complementary enforcement mechanisms.

References

AbdelRahman Abdou, Ashraf Matrawy, and Paul C. Van Oorschot. *CPV: Delay-based location verification for the Internet*. TDSC 2015.

Mohammed Jubaer Arif, Shanika Karunasekera, and Santosh Kulkarni. *GeoWeight: internet host geolocation based on a probability model for latency measurements*. ACSC 2010.

Shinyoung Cho, Zachary Weinberg, Arani Bhattacharya, Sophia Dai and Ramsha Rauf. *Selection of Landmarks for Efficient Active Geolocation. TMA 2024*

Bamba Gueye, Artur Ziviani, Mark Crovella, and Serge Fdida. *Constraint-based geolocation of internet hosts*. IMC 2004.

Katharina Kohls and Claudia Diaz. *VerLoc: Verifiable Localization in Decentralized Systems*. USENIX 2022.

Sándor Laki, Péter Mátray, Péter Hága, Tamás Sebők, István Csabai, and Gábor Vattay. *Spotter: A model based active geolocation service*. INFOCOM 2011

Deepak Maram, Iddo Bentov, Mahimna Kelkar, and Ari Juels. *GoAT: File geolocation via anchor timestamping*. Cryptology ePrint Archive, 2021.

Venkata N. Padamanabban and Lealkshminarayanan Subramanian. *Determining the geographic location of Internet hosts*. SIGMETRICS 2001.

Peiyao Sheng, Vishal Sevani, Himanshu Tyagi, and Pramod Viswanath. *BFT-PoLoc: A Byzantine Fortified Trigonometric Proof of Location Protocol using Internet Delays*. arXiv preprint, 2024.

Bernard Wong, Ivan Stoyanov, and Emin Gün Sirer. *Octant: a comprehensive framework for the geolocalization of internet hosts*. NSDI 2007.