CSAI Foundation

Cloud Security Alliance AI Safety Initiative

The Irremediable Attack Surface

SOHO Device Monoculture as Persistent Nation-State Infrastructure

Unofficial AI-assisted Research

2026-04-09

Table of Contents

Executive Summary 4

Introduction and Background 5

- The SOHO Device Category
- End-of-Life as an Irremediable State
- The Monoculture Condition

A Decade of Nation-State Exploitation 7

- VPNFilter: Establishing the Template (2018)
- Volt Typhoon and KV-Botnet: Operational Infrastructure for Long-Haul Intrusion (2023–2024)
- Salt Typhoon: Telecommunications Backbone Infiltration (2024–2025)
- Forest Blizzard: DNS Hijacking as Cloud Credential Theft (2025–2026)
- The Criminal Ecosystem: TheMoon and Faceless

The Technical Foundations of Persistence 11

- Why SOHO Devices Fail
- The Reconstitution Problem

Systemic Risk: Critical Infrastructure Dependencies 13

- The Relay Chain Problem
- Supply Chain and Third-Party Risk Amplification
- Interconnection with Critical Sectors

Regulatory and Policy Responses 14

- CISA Secure by Design Alert (January 2024)
- FCC National Security Determination (March 2026)
- Law Enforcement Disruption Operations

Recommendations 16

- For Critical Infrastructure Operators
- For Organizations Managing Remote Workers and Third Parties
- For SOHO Device Owners and Small Businesses
- For Manufacturers and Policymakers

CSA Resource Alignment 19

Conclusions 20

References 21

Executive Summary

Across the past decade, nation-state threat actors have systematically identified small office/home office (SOHO) routers and edge devices as a category of infrastructure ideally suited to long-term operational use: widely deployed, architecturally uniform, chronically unpatched, and almost entirely invisible to enterprise security monitoring. The resulting attack surface is not a temporary vulnerability waiting to be remediated. It is a structural condition—tens of millions of devices sharing identical firmware versions and unpatched vulnerabilities, with manufacturers who have formally declared the end of their support obligations and consumers who lack both the expertise and the incentive to replace equipment that appears to function normally.

The evidence that nation-states have recognized and operationalized this condition is now extensive and multi-sourced. The Chinese state-sponsored actor Volt Typhoon built and rebuilt a botnet from end-of-life Cisco and Netgear routers to obscure its intrusion traffic against U.S. critical infrastructure, with the FBI conducting a court-authorized disruption in January 2024 only to observe the botnet reconstituting itself within months. [1][2] Salt Typhoon exploited unpatched Cisco IOS XE vulnerabilities—some documented for seven years before exploitation—to compromise at least 200 organizations across 80 countries, including multiple U.S. telecommunications providers. [3][4] Russian military intelligence (Forest Blizzard, also known as APT28) has since at least May 2025 operated a campaign compromising SOHO routers to hijack DNS resolution and conduct adversary-in-the-middle attacks against cloud credentials, affecting more than 200 organizations before a DOJ-led disruption operation in April 2026. [6][22] Criminal actors simultaneously monetized the same device category through proxy services like Faceless, which drew on a botnet of more than 40,000 end-of-life SOHO devices harvested across 88 countries. [7]

What unites these campaigns is not the sophistication of the exploit—in most cases the vulnerabilities involved were old, well-documented, and unpatched precisely because the devices had exceeded their manufacturer support lifetime. What unites them is the structural advantage of the target: SOHO devices constitute a global monoculture of network infrastructure, where a single family of vulnerabilities can be reliably weaponized across hundreds of thousands of identically configured nodes. This paper examines that structural condition, its exploitation by multiple nation-state actors, the limits of both law enforcement disruption and regulatory intervention, and the risk management posture organizations must adopt given an attack surface that cannot be remediated at the source.

Introduction and Background

The SOHO Device Category

Small office and home office routers, along with associated edge devices such as network-attached storage units, wireless access points, and IP cameras, occupy a peculiar position in the security ecosystem. They are genuinely critical infrastructure—every packet transiting the home or small business network passes through the router—yet they are purchased, deployed, and managed with the same expectations applied to consumer appliances. Unlike enterprise network equipment, SOHO devices are sold largely on the basis of price and ease of setup, with security treated as an implicit feature rather than a specification. Users rarely log into management interfaces after initial configuration, firmware updates are rarely automatic, and default credentials remain in place far longer than any security standard would recommend.

The resulting installed base reflects these purchasing conditions. Market research places the global Wi-Fi router market at approximately USD 16.3 billion in 2024, with the SOHO and residential segments comprising roughly 55 percent of that revenue. [8] While precise global device counts are difficult to establish, the installed base of active SOHO routers numbers in the hundreds of millions, reflecting the near-ubiquity of broadband connectivity in developed economies and its rapid expansion in developing ones. The concentration of market share among a handful of manufacturers—Cisco (through its Linksys and Small Business lines), Netgear, ASUS, TP-Link, D-Link, and DrayTek account for the majority of global volume—means that firmware codebases are shared across enormous device populations. A vulnerability in the web management interface of a widely deployed chipset affects not one product but entire generations across multiple brands.

End-of-Life as an Irremediable State

End-of-life (EOL) is the industry term for a product that has passed the date after which the manufacturer will no longer release security patches or firmware updates. For enterprise equipment, EOL triggers procurement processes and replacement planning. For SOHO devices, it frequently means nothing changes at all. The device continues to pass traffic and appear functional; the household or small business has no visible indication that a material security boundary has been crossed; and the manufacturer has formally released itself from any obligation to address newly discovered vulnerabilities.

The scale of this problem becomes concrete in specific cases. Cisco's RV320 and RV325 routers, widely deployed in small businesses, passed their End of Software Maintenance dates and Cisco explicitly stated it would not release patches for subsequently discovered vulnerabilities. [9] When SecurityScorecard researchers analyzed the KV-Botnet in early 2024, they found that Volt Typhoon had compromised

approximately 30 percent—325 of 1,116 observable devices—of all internet-facing Cisco RV320/325 routers in a 37-day window. [10] The devices were not targeted because they were strategically significant individually; they were targeted because they were available, uniform, and permanently unpatched.

This permanence is the defining characteristic that transforms SOHO EOL devices from a routine security hygiene problem into a structural threat. Enterprises can patch or replace managed assets. SOHO routers are not managed assets in any meaningful sense; they are utilities, and the notion of a "patch cycle" is entirely foreign to the context in which they are owned and operated. When Lumen Technologies' Black Lotus Labs disclosed the resurgence of TheMoon malware in early 2024, they found it actively targeting EOL devices from Asus, D-Link, Linksys, Netgear, and other manufacturers—not because patch-compatible versions did not exist, but because the EOL devices were the most reliably vulnerable nodes in the global network topology. [7]

The Monoculture Condition

The risk amplifier that transforms individual device vulnerabilities into systemic threats is monoculture: the concentration of an enormous number of network nodes on identical or near-identical hardware and firmware. In computer science, monoculture risk is understood to mean that when a large population of systems shares the same vulnerabilities, a single successful exploit class can propagate or be operationalized across that entire population without adaptation. Agricultural analogies are frequently invoked—and apt—because they capture how uniformity that appears efficient in normal conditions becomes catastrophic fragility under adversarial pressure. [11][38]

The CrowdStrike software outage of July 2024 provided a modern, non-adversarial illustration of monoculture's systemic risk. A single flawed sensor configuration update disabled Windows systems protected by CrowdStrike's Falcon platform across organizations representing as many as 45 percent of Fortune 100 companies, causing cascading disruptions to healthcare, aviation, and supply chain operations. [12] The analogy to SOHO device vulnerability is direct: where CrowdStrike presented a single point of systemic failure through software update, SOHO device monoculture presents a standing, structurally irremediable point of failure through shared EOL firmware. The difference is that the CrowdStrike outage was accidental and recoverable; SOHO monoculture exploitation is deliberate, persistent, and by design invisible.

For nation-state actors operating at scale, the monoculture condition is operationally invaluable. A campaign developed to exploit a specific vulnerability in Cisco RV320 firmware can be applied to every member of that device population globally without modification. The marginal cost of each additional compromised node approaches zero once the initial exploit is developed, and the geographic distribution of the installed base provides relay infrastructure across jurisdictions that complicate law enforcement response.

A Decade of Nation-State Exploitation

VPNFilter: Establishing the Template (2018)

The operational template for nation-state SOHO exploitation was established at scale by VPNFilter, a multi-stage modular malware platform attributed to APT28, the Russian military intelligence directorate (GRU). [13] When Cisco Talos and the FBI disclosed the campaign in May 2018, VPNFilter had infected more than 500,000 devices across 54 countries, with affected manufacturers including Asus, D-Link, Huawei, Linksys, Mikrotik, Netgear, QNAP, TP-Link, Ubiquiti, Upvel, and ZTE. [13] The breadth of that list reflects not a campaign requiring device-specific development but one exploiting common vulnerability classes present across the consumer router market.

VPNFilter was notable for its sophistication relative to prior router malware. Its architecture included a persistent first stage that survived reboots, a second stage that provided command execution and data exfiltration capabilities, and a suite of third-stage plugins enabling packet-capture, network traffic interception, and search for SCADA industrial control system protocols. [13] The campaign included a destructive capability—the ability to overwrite firmware and brick devices—that the operators could trigger across the entire botnet simultaneously. At the time of FBI disruption via sinkholing of the command-and-control domain, a large portion of the infected devices were concentrated on Ukrainian networks, suggesting pre-positioning for offensive infrastructure operations.

VPNFilter demonstrated three things that adversaries and defenders alike absorbed: first, that SOHO routers could sustain sophisticated, multi-stage malware without any visible degradation in device function; second, that the global distribution of SOHO devices provided relay infrastructure that was practically impossible to attribute or disrupt without law enforcement action in dozens of jurisdictions; and third, that the same device population could be weaponized by different threat actors for qualitatively different purposes—espionage, traffic manipulation, destructive attack.

Volt Typhoon and KV-Botnet: Operational Infrastructure for Long-Haul Intrusion (2023–2024)

The most extensively documented case of SOHO infrastructure weaponization by a nation-state actor is the KV-Botnet operated in connection with Volt Typhoon (also tracked as Insidious Taurus, Bronze Silhouette), a People's Republic of China state-sponsored threat actor assessed by U.S. intelligence agencies to be pre-positioning access across U.S. critical infrastructure for potential use in a conflict scenario involving Taiwan. [1]

Volt Typhoon's use of SOHO infrastructure is qualitatively different from a conventional botnet. The actor's primary objective was not to leverage the botnet for denial-of-service attacks or spam; it was to route traffic through legitimate U.S. IP address space to blend intrusion activity against critical infrastructure targets with normal internet traffic. By routing command-and-control communications through compromised SOHO routers belonging to U.S. residential and small-business subscribers, Volt Typhoon was able to present its traffic as originating from the same geographic and network context as its targets—a technique often called "living off the land" in network routing terms, analogous to the use of built-in operating system tools to avoid triggering endpoint detection. [1][14]

The device composition of the KV-Botnet reflects the EOL condition at the core of the attack surface. The botnet's routers included Cisco RV320s, DrayTek Vigor routers, NETGEAR ProSAFE devices, and Axis IP cameras. [3] With the exception of the Axis cameras, all of these devices had passed their End of Software Maintenance dates, and their manufacturers had either stated they would not issue patches or had released advisories recommending replacement rather than remediation. [9] Cisco's posture was unambiguous: it would not release patches for the RV320/325 vulnerabilities because the products were past their end-of-life dates, regardless of the severity of the vulnerabilities discovered.

When SecurityScorecard analyzed the botnet's composition in early 2024, the results illustrated the speed with which EOL devices can be operationalized at scale: approximately 30 percent of all observable internet-facing Cisco RV320/325 routers had been compromised within 37 days, representing 325 of 1,116 tracked devices. [10] This figure is not a total device population count; it is the fraction of devices observable through public scanning infrastructure. The actual number of compromised devices was almost certainly substantially larger.

The U.S. government's response demonstrated both the feasibility and the limits of law enforcement disruption. In January 2024, the FBI obtained court authorization to remotely access compromised routers and delete the KV-Botnet malware, severing the devices from their command-and-control infrastructure. [15][16] The operation was technically successful: the malware was removed from hundreds of routers. Within months, however, Volt Typhoon had reconstituted its botnet infrastructure using the same device category and the same underlying vulnerability conditions. By September 2024, researchers confirmed the botnet was operational again, exploiting the same end-of-life Cisco and Netgear devices. [5] The takedown had disrupted the infrastructure but had done nothing to address its root cause: the devices remained EOL, remained unpatched, and remained connected to the internet.

Salt Typhoon: Telecommunications Backbone Infiltration (2024–2025)

While Volt Typhoon operated through SOHO devices as relay infrastructure, Salt Typhoon (also tracked as RedMike, GhostEmperor, FamousSparrow) took a different approach—exploiting vulnerabilities in enterprise-class Cisco network equipment used by telecommunications providers. [3][17] The distinction

matters less than it might appear: the vulnerability exploitation method was the same (unpatched network devices at the perimeter), and the affected equipment, while carrier-grade rather than consumer-grade, had in many cases gone unpatched for years.

Between December 2024 and January 2025, Salt Typhoon attempted to exploit more than 1,000 internet-facing Cisco network devices globally, primarily those associated with telecommunications providers. [17] The primary vulnerability chain exploited two privilege escalation flaws in Cisco's IOS XE software: CVE-2023-20198, which allows an unauthenticated remote attacker to create a high-privilege account, and CVE-2023-20273, which allows that account to be used to inject commands with root-level privileges. [18] In at least one case, attackers exploited a Cisco router vulnerability that had been publicly documented for seven years before the intrusion, underscoring that the problem is not the discovery of novel zero-days but the persistence of known, unpatched vulnerabilities across operational network infrastructure. [19]

The scope of the Salt Typhoon campaign established it as among the most significant telecommunications security breaches in U.S. history. By early 2025, the campaign had compromised major U.S. telecommunications providers including Verizon, AT&T, and Lumen Technologies, with reporting indicating the campaign had affected at least 200 organizations across 80 countries. [20] Five additional telecom networks were subsequently identified as compromised: a U.S. internet service provider, a U.K.-affiliated telecom provider, an Italian ISP, a large Thai telecom company, and a South African provider. [21] The intrusions provided access to metadata and content of communications across significant subscriber populations, and to systems used for court-authorized interception—the telecommunications infrastructure operated in compliance with CALEA, the U.S. Communications Assistance for Law Enforcement Act.

The Salt Typhoon campaign highlights a dimension of the SOHO-and-edge-device risk that extends beyond the residential router: the same pattern of deferred patching, extended device lifecycles, and inadequate network perimeter monitoring that makes consumer SOHO routers exploitable applies across a continuum of network edge devices, including carrier-grade equipment that has aged past the active patch window. The attackers exploited CVEs published in 2023 against equipment that remained connected and unpatched more than a year later.

Forest Blizzard: DNS Hijacking as Cloud Credential Theft (2025–2026)

The most recent major disclosure, published by Microsoft Threat Intelligence on April 7, 2026, documents a campaign by Forest Blizzard (APT28, STRONTIUM)—a unit of Russian military intelligence—that has operated since at least May 2025 against a fundamentally different target: cloud service credentials accessible through DNS manipulation. [22][23]

Forest Blizzard's technique exploits the position of SOHO routers as the network gateway through which all DNS queries flow. After gaining access to SOHO devices, the threat actor altered the default DNS resolver configuration on compromised routers to redirect DNS queries to actor-controlled resolvers. [22] This

manipulation is largely invisible to users: the browser still displays the expected hostname in the address bar, certificates may still appear valid, and the connection appears to succeed. What changes is that the attacker's DNS infrastructure returns fabricated records pointing to adversary-controlled servers, enabling adversary-in-the-middle (AiTM) interception of TLS-encrypted sessions by serving fraudulent certificates or redirecting authentication flows.

Microsoft identified more than 200 organizations and approximately 5,000 consumer devices affected by Forest Blizzard's malicious DNS infrastructure. [22] The operational objective was credential harvesting for cloud services, with particular emphasis on Microsoft 365 accounts. Microsoft characterized this as the first observed instance of Forest Blizzard using DNS hijacking at scale to enable AiTM interception of TLS connections following edge device exploitation—a notable evolution in the actor's tradecraft. [22] The campaign's sub-group, tracked as Storm-2754, operated the DNS infrastructure component. The DOJ conducted a court-authorized disruption of the DNS hijacking network in coordination with the Microsoft disclosure. [24][25]

The Forest Blizzard campaign illustrates an important evolution in SOHO exploitation doctrine. Earlier campaigns used compromised routers as passive relays, routing traffic through them to obscure attribution. Forest Blizzard's DNS manipulation made the compromised routers active participants in credential theft—an exploitation of the trusted, authoritative position routers hold in the DNS resolution chain. The same device that obscures traffic origin for Volt Typhoon can, with a different configuration change, intercept every authentication session conducted by every user on the network.

The Criminal Ecosystem: TheMoon and Faceless

Nation-state actors do not operate in isolation. Parallel to government-sponsored campaigns, a criminal market has developed around SOHO device compromise—both to monetize the compromised devices directly and to rent the resulting infrastructure to other actors, including those conducting sophisticated targeted attacks. [26]

The most thoroughly documented example is TheMoon malware and its relationship to the Faceless proxy service. TheMoon first emerged in 2014; Lumen's Black Lotus Labs documented its resurgence in early 2024 targeting EOL SOHO devices from a range of manufacturers. [7] By January and February 2024, TheMoon had built a botnet of more than 40,000 devices across 88 countries. [7][37] In March 2024 alone, a focused campaign compromised more than 6,000 ASUS routers in under 72 hours. [27] Lumen's researchers assessed that TheMoon functions as the primary botnet feeding the Faceless proxy service—a criminal proxy-for-hire that allows customers to route traffic through compromised residential and small-business internet connections, appearing as legitimate users in the target country without any know-your-customer verification and accepting only cryptocurrency payment. [7]

The Faceless service illustrates the dual-use nature of SOHO compromise infrastructure. The same device population that Volt Typhoon uses for living-off-the-land intrusion routing can be and is rented to ransomware operators, fraud networks, and credential-stuffing campaigns through commercial criminal services. Trend Micro researchers documented this convergence explicitly in 2024, observing that compromised SOHO routers serve as a shared substrate for both nation-state advanced persistent threat operations and commodity cybercrime. [26] The implication is that SOHO device compromise is not merely a nation-state problem amenable to diplomatic or intelligence responses; it is a criminal economy that nation-state actors can access, benefit from, or contract with as operational circumstances require.

The Technical Foundations of Persistence

Why SOHO Devices Fail

The vulnerability characteristics of SOHO devices are well-documented and consistent. Independent Security Evaluators' "SOHOpelessly Broken 2.0" research, assessing 13 SOHO routers and network-attached storage devices, found 125 CVEs across the tested population. [28] In almost every tested device, researchers achieved remote root-level access. The vulnerabilities found were not novel classes—they included cross-site scripting, command injection, authentication bypass, and improper input validation, attack categories catalogued in security research for decades. [28] Subsequent research by Cisco Talos, analyzing routers in the post-VPNFilter disclosure period, documented hundreds of additional vulnerabilities across SOHO and small-office wireless router models. [29][39]

The persistent failure to address these vulnerability classes in SOHO firmware reflects structural incentives. Consumer routers compete primarily on price, wireless range, and ease of setup. Security does not register as a purchase criterion for most buyers. Manufacturers therefore have weak market incentives to invest in secure development practices or to maintain long patch cycles. The CISA and FBI Secure by Design Alert issued in January 2024 described the design conditions enabling Volt Typhoon's exploitation: manufacturers building routers with web management interfaces exposed to the public internet by default, lacking automatic update capabilities, using default credentials shared across entire product lines, and including exploitable defects in web application code. [30][31] Each of these conditions is known and preventable; each persists because market competition does not price it in.

The absence of automatic updates is particularly consequential for the EOL condition. Even for devices still within their support lifetimes, patches require user action to apply. Studies of patch deployment across consumer networking equipment consistently find that large fractions of devices run firmware versions years

behind current releases. When a device reaches EOL, this chronic underpatching becomes permanent: there are no future patches to be applied, even by a diligent user. The device freezes at whatever firmware version it ran at the time the manufacturer discontinued support, accumulating vulnerability debt indefinitely.

The Reconstitution Problem

Law enforcement disruption operations against SOHO botnets have demonstrated consistent effectiveness and consistent impermanence. The FBI's court-authorized deletion of KV-Botnet malware from hundreds of Volt Typhoon-controlled routers in January 2024 was technically successful. The underlying vulnerability in each affected device remained. The device remained connected to the internet. The actor retained knowledge of the exploitation method. Within months the botnet was rebuilt. [5] The pattern is not unique to Volt Typhoon: VPNFilter's disruption via sinkholing in 2018 disrupted the campaign's command-and-control but could not eliminate the infection from the 500,000 devices spread across 54 countries. The Forest Blizzard disruption of April 2026 similarly removed the immediate threat while leaving the structural conditions enabling it fully intact.

This reconstitution problem is intrinsic to the target category. An enterprise network can be remediated: endpoints are managed, patches can be deployed, access controls can be enforced, and configuration drift can be detected. The global population of SOHO routers is not a managed environment. There is no mechanism by which a government or industry body can enforce firmware updates across hundreds of millions of consumer devices in private hands. Law enforcement operations can excise specific malware from specific devices at a specific moment in time. They cannot prevent the re-exploitation of vulnerabilities that remain present in the firmware.

The MITRE ATT&CK framework's documentation of the KV-Botnet campaign (C0035) captures the operational sophistication of this persistence: the actor used living-off-the-land techniques on compromised routers to avoid detection, routed traffic through multiple compromised devices to complicate attribution, and chose target devices specifically because their EOL status guaranteed the absence of monitoring by security vendors or automated patching systems. [14] The choice of target was not incidental; it was a deliberate selection of infrastructure engineered to be irremediable.

Systemic Risk: Critical Infrastructure Dependencies

The Relay Chain Problem

Nation-state actors using SOHO devices as relay infrastructure are not merely obfuscating their own traffic—they are systematically undermining the ability of critical infrastructure operators to identify the true origin of attacks against them. When intrusion traffic arrives at an energy grid management system or a water treatment facility's operational technology network appearing to originate from a residential IP address in the same metropolitan area, the investigative and defensive calculus becomes substantially more difficult. Source IP reputation blocklists are ineffective. Geographic filtering is ineffective. Behavioral anomaly detection must work harder to distinguish traffic from a compromised local router from traffic from a legitimate local user.

This relay chain dynamic is central to why Volt Typhoon chose SOHO infrastructure as its primary operational mechanism. The targeted critical infrastructure sectors—communications, energy, transportation, water, and government facilities—are not compromised through the SOHO routers directly. The routers provide the camouflage layer that makes the intrusion traffic appear normal. By the time investigators identify that specific traffic flows are malicious, the SOHO relay nodes are typically the most recent identifiable hop, and attribution beyond that point requires the kind of international law enforcement cooperation that moves on timescales measured in years, not the hours or days that matter for incident response.

Supply Chain and Third-Party Risk Amplification

The SOHO device risk is further amplified by the third-party network access patterns common in enterprise environments. Remote workers, contractors, managed service providers, and technical staff routinely access enterprise resources from home networks traversing SOHO routers. In the vast majority of these cases, the enterprise has no visibility into the security posture of the SOHO router: whether it is EOL, what firmware version it runs, whether it has been compromised. The Forest Blizzard DNS hijacking campaign, which targeted credentials for cloud services including Microsoft 365, is a direct exploitation of this third-party access pattern. [22] An employee accessing corporate email from a home network behind a compromised router may have their credentials captured before the TLS connection to the corporate authentication infrastructure is even established.

The organizational security perimeter, already substantially eroded by cloud adoption and remote work normalization following the COVID-19 pandemic, cannot be extended to encompass SOHO router security. Zero trust architecture principles address the logical consequence—no implicit trust based on network

location—but do not eliminate the credential-theft risk of a DNS hijacking attack upstream of the authentication flow. If DNS resolution for the corporate identity provider is compromised at the user's router, the user may be directed to a phishing page that precedes any zero trust verification.

Interconnection with Critical Sectors

The targeting patterns documented across Volt Typhoon, Salt Typhoon, and related Chinese and Russian campaigns reveal specific focus on the sectors most critical to government and military operations in a conflict scenario: telecommunications, which carries government and emergency communications; energy, which powers defense installations and civilian infrastructure alike; water and wastewater systems; and transportation networks. [1][20] CISA's advisory on Volt Typhoon assessed that the actor was seeking to develop capabilities that could be activated to disrupt critical services during a conflict, rather than conducting immediate collection. [1] The SOHO infrastructure network serves as both the relay layer obscuring this pre-positioning and, in some cases, a staging platform from which to pivot toward higher-value enterprise targets.

Salt Typhoon's access to telecommunications infrastructure is particularly significant because it represents compromise of the sector responsible for communications security during any national emergency response. Telecommunications companies operate the systems used for court-authorized interception, emergency alert distribution, and the routing of government communications. Salt Typhoon's intrusion into the lawful intercept systems of multiple U.S. carriers—itself an exploitation of an edge device vulnerability at the network perimeter—means that the infrastructure designed to provide visibility into adversary communications was itself compromised by an adversary.

Regulatory and Policy Responses

CISA Secure by Design Alert (January 2024)

The immediate U.S. government policy response to the Volt Typhoon disclosures included a January 2024 joint alert from CISA and the FBI urging SOHO router manufacturers to adopt secure-by-design principles. [30] The alert identified specific design failures enabling the campaign—management interfaces exposed to the public internet by default, absence of automatic updates, use of default shared credentials—and called on manufacturers to eliminate exploitable defects in web management interfaces during product design, require unique or time-limited setup credentials, enable automatic security updates by default, and restrict management interface access to local network connections. [30][31]

The alert represents a meaningful evolution in the framing of SOHO security risk: rather than addressing the problem as user education or consumer awareness, it placed responsibility with manufacturers as the parties capable of addressing the structural conditions at the design stage. However, the alert is voluntary guidance, not regulation. Manufacturers are invited to adopt these principles; they are not required to do so. The market conditions that produced the existing insecure design baseline—price competition that does not price security—remain unchanged by advisory-based intervention.

FCC National Security Determination (March 2026)

A more consequential regulatory action came in March 2026, when a federal interagency body convened by the White House completed a National Security Determination finding that foreign-produced consumer routers posed unacceptable risks to the United States. [32] The determination cited two categories of risk: the introduction of supply chain vulnerabilities capable of disrupting the U.S. economy, critical infrastructure, and national defense; and the establishment of cybersecurity risks that could be leveraged to immediately and severely disrupt U.S. critical infrastructure. [32] The FCC subsequently added foreign-produced consumer-grade routers to its Covered List under the Secure Networks Act, effectively banning new imports of foreign-manufactured consumer routers absent a Conditional Approval from the Department of War (DoW) or the Department of Homeland Security. [33][40]

The FCC action directly cited the Volt Typhoon, Flax Typhoon, and Salt Typhoon campaigns as evidence base for the national security determination. [33] The scope is notable: the determination applies to consumer routers manufactured in any foreign country, not only those associated with designated foreign adversaries. The conditional approval process, which can grant 18-month exemptions, acknowledges the practical impossibility of immediately substituting domestic production for the global SOHO router supply chain. But the broader regulatory intent is clear: SOHO device security is now understood as a national security issue requiring regulatory intervention, not merely industry voluntary action.

The FCC action addresses future procurement without touching the installed base problem. Foreign-produced routers already deployed in U.S. homes and businesses—including all the EOL devices currently being exploited by Volt Typhoon—are not recalled or patched by the Covered List determination. The irremediable attack surface remains irremediable; the policy intervention slows its expansion rather than contracting it.

Law Enforcement Disruption Operations

Between January 2024 and April 2026, U.S. law enforcement conducted at least three court-authorized operations targeting SOHO-based nation-state infrastructure: the FBI disruption of the KV-Botnet in January 2024; an operation disrupting Forest Blizzard's DNS hijacking network announced in April 2026

contemporaneously with Microsoft's technical disclosure; and earlier actions targeting Russia's Forest Blizzard infrastructure under prior operational names. [15][16][24][25][36] Each operation was operationally successful in temporarily disabling the targeted infrastructure.

These operations are significant for several reasons beyond their immediate impact. They establish legal precedent for court-authorized access to and modification of privately owned network devices that have been incorporated into foreign state espionage infrastructure. They demonstrate the value of public-private partnerships between law enforcement and technology companies like Microsoft, whose threat intelligence identified and documented the Forest Blizzard campaign. And they impose meaningful operational costs on threat actors, forcing reconstitution of infrastructure that has been disrupted.

The limits of disruption operations are equally significant. Each operation addresses infrastructure that has already been built and already been used for operational purposes. The underlying device vulnerabilities are not corrected. The device population remains exploitable. A threat actor with sufficient operational patience—and nation-state actors have demonstrated exactly this patience in their SOHO campaigns—can rebuild compromised infrastructure faster than law enforcement can repeatedly obtain the court authorization, technical access, and international cooperation required to disrupt it.

Recommendations

For Critical Infrastructure Operators

The starting premise for organizations responsible for critical infrastructure must be that the SOHO and edge-device attack surface cannot be eliminated and that traffic appearing to originate from domestic IP addresses in normal geographic and network contexts cannot be implicitly trusted. Zero trust architecture, as articulated in NSA's Zero Trust Implementation Guidelines and CISA's Zero Trust Maturity Model, provides the appropriate framework for network access policy in this environment: explicit verification of identity and device posture at every access boundary, with no trust granted on the basis of network location alone. [34][35]

Practically, this means critical infrastructure operators should implement multi-factor authentication for all remote access, not as a supplementary control but as the baseline condition for any user accessing operational systems from any network. The Forest Blizzard DNS hijacking campaign specifically targeted cloud service credentials; phishing-resistant MFA methods—hardware security keys compliant with FIDO2/WebAuthn—resist the credential-theft objective of AiTM attacks that mere push-notification MFA does not. Organizations should also deploy network monitoring capable of identifying DNS resolution anomalies for corporate resources, since the first indicator of a DNS hijacking attack affecting remote workers is DNS queries for corporate hostnames resolving to unexpected IP addresses.

Microsegmentation within operational technology environments is essential to limit the blast radius of any intrusion that exploits the SOHO relay chain. CISA's Zero Trust Microsegmentation Guidance, released in July 2025, provides a phased implementation framework specifically addressing the complexity of transitioning legacy operational environments to microsegmented network architectures. [35] The goal is to ensure that a threat actor who successfully routes traffic through a compromised SOHO router and gains a foothold in an enterprise environment encounters internal network controls that limit lateral movement.

For Organizations Managing Remote Workers and Third Parties

Organizations with significant remote workforces or contractor access populations face the most direct exposure to the SOHO relay and DNS hijacking risk. The most effective near-term mitigation is deploying enterprise DNS resolution for all remote access sessions: through VPN configurations that route all DNS queries through enterprise-controlled resolvers, or through cloud-based secure DNS services that provide corporate DNS resolution regardless of the user's local router configuration. When DNS is resolved by an enterprise-controlled service, the effectiveness of router-level DNS hijacking is substantially reduced.

Supply chain risk management programs should extend beyond software vendors and cloud service providers to include the network access conditions of high-privilege remote users, contractors with access to sensitive systems, and managed service providers with privileged administrative access. While requiring SOHO router security audits for individual employees is impractical, organizations can require that remote workers with privileged access use provided or approved hardware security tokens, route all work traffic through enterprise VPN with DNS control, and use dedicated work devices that do not share the same network segment as other household traffic.

Third-party access risk assessments should evaluate whether contractors and managed service providers have their own network security standards that address SOHO device management. For the highest-risk access relationships—those with administrative access to operational technology or sensitive data systems—organizations may consider providing hardware routers with specific security configurations rather than relying on third-party network security posture.

For SOHO Device Owners and Small Businesses

SOHO device owners face a genuine information asymmetry: manufacturers rarely communicate clearly that a device has reached EOL, and the device continues to function normally even when it is no longer receiving security updates. The most direct mitigation is to determine whether devices in use have reached EOL status, using the manufacturer's product lifecycle documentation, and to replace EOL devices on a planned schedule rather than waiting for device failure.

For devices still within their support lifetime, enabling automatic firmware updates is the single highest-impact available action. CISA's guidance consistently identifies unpatched firmware as the primary vector enabling SOHO exploitation. [30] Web management interfaces should be restricted to local network access and the public internet management interface should be disabled. Default credentials should be replaced with strong, unique passwords, and unused services (remote management, UPnP, and similar) should be disabled.

For small businesses unable to maintain internal IT capability, managed network service providers can deliver enterprise-grade security controls—including monitoring, automated patching, and threat detection—to SOHO environments. The market for managed SOHO security services has matured significantly in response to the documented threat environment. While not eliminating the underlying device risk, managed services can substantially reduce the exposure window by monitoring for indicators of compromise and ensuring timely patch application.

For Manufacturers and Policymakers

The CISA Secure by Design Alert identified specific design decisions that enable SOHO exploitation at scale. Manufacturers should adopt those principles as baseline product requirements: unique default credentials, automatic security update capability enabled by default, management interface restricted to local network access, and defined patch support commitments communicated clearly to purchasers at the time of sale. Extended support commitments—analogous to the long-term support arrangements common in enterprise software—would reduce the rate at which deployed devices transition into the EOL irremediable state.

Policymakers should complement the FCC's forward-looking procurement restrictions with consideration of measures targeting the installed base. Options include mandatory disclosure programs requiring manufacturers to notify purchasers when products reach EOL status, incentive structures (tax credits, trade-in programs) to accelerate the retirement of EOL devices from sensitive network positions, and minimum security labeling requirements that allow purchasers to assess the patch lifecycle commitment of new devices at purchase. The European Union's Cyber Resilience Act, which entered into force in late 2024 and mandates security requirements and vulnerability disclosure obligations for connected devices sold in the EU market, provides a regulatory model that addresses manufacturer obligations more directly than current U.S. voluntary guidance.

CSA Resource Alignment

The threat landscape described in this paper intersects with several foundational CSA frameworks and guidance documents. Organizations operationalizing the recommendations above should reference these resources alongside the primary government advisories.

CSA's MAESTRO (Multi-layer Analysis of Emerging Security Threats, Risks, and Objectives) threat modeling framework provides structured methodologies for analyzing threats to network infrastructure across multiple abstraction layers—precisely the kind of layered analysis required when SOHO devices function simultaneously as relay infrastructure, DNS manipulation points, and credential-theft enablers. The SOHO exploitation pattern, in which edge-device compromise enables both traffic relay and active AiTM attacks at higher layers of the network stack, is a canonical MAESTRO analysis case.

The AI Controls Matrix (AICM), as a superset of the Cloud Controls Matrix (CCM), provides control specifications applicable to the network access context of remote AI workloads and AI-integrated enterprise systems. Organizations deploying AI systems accessible from remote networks traversing SOHO infrastructure should apply AICM controls specifically addressing network perimeter integrity, authentication assurance, and supply chain risk for network devices in the access path. CCM control domains covering Identity and Access Management, Infrastructure and Virtualization Security, and Supply Chain Management and Transparency are directly applicable to the risk conditions this paper analyzes.

CSA's Zero Trust Guidance for Critical Infrastructure provides sector-specific implementation guidance aligned with the recommendations in this paper. The core principles—verify explicitly, use least-privilege access, assume breach—directly address the network-location trust assumption that SOHO router compromise exploits. Critical infrastructure organizations should treat this guidance as complementary to the NSA Zero Trust Implementation Guidelines and CISA Zero Trust Microsegmentation Guidance cited above.

CSA's STAR (Security Trust Assurance and Risk) program offers a framework for assessing and communicating the security posture of cloud and network service providers. Organizations evaluating managed SOHO security service providers or cloud-based DNS security vendors as part of their response to the threats documented here should use STAR-aligned assessment criteria to evaluate vendor security commitments and verify claimed controls.

Conclusions

The exploitation of SOHO devices by nation-state actors is not a temporary threat that will resolve as awareness improves or as vendors release patches. The structural conditions enabling it—concentrated market share among a small number of vendors, massive installed base of EOL devices frozen in permanently unpatched firmware states, consumer deployment context that provides no mechanism for monitoring or management, and global distribution that defeats geographic and network attribution—are durable conditions built into the architecture of the SOHO device market. They will remain material for as long as the currently deployed installed base of EOL devices remains connected to the internet, a period that could extend a decade or more given the typical replacement lifecycles of consumer networking equipment.

The campaigns documented in this paper—VPNFilter, KV-Botnet and Volt Typhoon, Salt Typhoon, Forest Blizzard's DNS hijacking operation, and the TheMoon/Faceless criminal ecosystem—represent not a series of discrete incidents but the progressive operational development of a global, multi-actor exploitation infrastructure. Each campaign has added sophistication: from passive relay (VPNFilter, Volt Typhoon) to active credential interception (Forest Blizzard), from single-actor operations to shared criminal proxy markets (Faceless) available to state and non-state actors alike. The trajectory points toward continued escalation in the operational uses to which compromised SOHO infrastructure will be put.

The appropriate security posture for organizations operating in this environment begins with accepting the irremediability of the attack surface rather than treating it as a problem awaiting a technical solution. SOHO devices will remain exploitable. Remote workers and third parties will access enterprise systems from compromised networks. The defensive priority is therefore to ensure that the enterprise systems and data those remote connections reach are protected by controls that do not depend on the integrity of the network layer through which the connection arrives—authentication that resists credential relay attacks, DNS resolution not dependent on the local router, network access policies that verify device posture rather than network location, and monitoring that can detect the anomalies of AiTM and relay-chain attacks even when the traffic appears to originate from a trusted geographic context.

Regulatory interventions—the FCC's Covered List determination, the CISA Secure by Design Alert, the potential adoption of EU Cyber Resilience Act-style mandatory requirements in U.S. law—address the conditions under which future devices will be manufactured and sold. They do not and cannot remediate the hundreds of millions of EOL devices already deployed. The irremediable attack surface is, for the foreseeable future, a permanent feature of the network environment in which critical infrastructure must operate.

References

[1] CISA, FBI, NSA, and International Partners. "PRC State-Sponsored Actors Compromise and Maintain Persistent Access to U.S. Critical Infrastructure." CISA Advisory AA24-038A, February 2024.

[2] U.S. Department of Justice. "U.S. Government Disrupts Botnet People's Republic of China Used to Conceal Hacking of Critical Infrastructure." DOJ Press Release, January 2024.

[3] SecurityWeek. "Chinese APT Volt Typhoon Linked to Unkillable SOHO Router Botnet." SecurityWeek, January 2024.

[4] Recorded Future / Insikt Group. "RedMike (Salt Typhoon) Exploits Vulnerable Cisco Devices." Recorded Future, February 2025.

[5] BleepingComputer. "Volt Typhoon Rebuilds Malware Botnet Following FBI Disruption." BleepingComputer, September 2024.

[6] Microsoft Threat Intelligence. "SOHO Router Compromise Leads to DNS Hijacking and Adversary-in-the-Middle Attacks." Microsoft Security Blog, April 7, 2026.

[7] Lumen Black Lotus Labs. "The Darkside of TheMoon." Lumen Blog, March 2024.

[8] Cognitive Market Research. "The Global Wi-Fi Routers Market Size Will Be USD 16,251.2 Million in 2024." Cognitive Market Research, 2024.

[9] Cisco. "Cisco Small Business RV042, RV042G, RV320, and RV325 Routers Vulnerabilities." Cisco Security Advisory, 2024.

[10] GlobalSecurityMag. "SecurityScorecard Threat Research: Volt Typhoon Compromises 30% of Cisco RV320/325 Devices in 37 Days." GlobalSecurityMag, January 2024.

[11] Wikipedia. "Monoculture (computer science)." Wikipedia.

[12] ResearchGate. "When Trust Fails: Examining Systemic Risk in the Digital Economy from the 2024 Crowd Strike Outage." ResearchGate, 2024.

[13] Infosec Institute. "The VPNFilter: A Powerful Botnet of More Than 500k Devices Ready to Attack." Infosec Institute, 2018.

[14] MITRE ATT&CK. "KV Botnet Activity, Campaign C0035." MITRE ATT&CK, 2024.

[15] The Hacker News. "U.S. Feds Shut Down China-Linked 'KV-Botnet' Targeting SOHO Routers." The Hacker News, February 2024.

[16] Dark Reading. "Feds Confirm Remote Killing of Volt Typhoon's SOHO Botnet." Dark Reading, 2024.

[17] CyberScoop. "Salt Typhoon Remains Active, Hits More Telecom Networks via Cisco Routers." CyberScoop, January 2025.

[18] Security Affairs. "China-Linked APT Salt Typhoon Breached Telecoms by Exploiting Cisco Router Flaws." Security Affairs, January 2025.

[19] Nextgov/FCW. "Salt Typhoon Hackers Exploited Stolen Credentials and a 7-Year-Old Software Flaw in Cisco Systems." Nextgov/FCW, February 2025.

[20] Wikipedia. "Salt Typhoon." Wikipedia.

[21] SecureWorld. "Salt Typhoon Expands Espionage Campaign, Targets Cisco Routers." SecureWorld, January 2025.

[22] The Hacker News. "Russian State-Linked APT28 Exploits SOHO Routers in Global DNS Hijacking Campaign." The Hacker News, April 2026.

[23] Windows Forum. "Forest Blizzard Hijacks SOHO Routers via DNS to Enable AiTM Cloud Espionage." Windows Forum, April 2026.

[24] U.S. Department of Justice. "Justice Department Conducts Court-Authorized Disruption of DNS Hijacking Network Controlled by a Russian Military Intelligence Unit." DOJ Press Release, April 2026.

[25] SecurityWeek. "US Disrupts Russian Espionage Operation Involving Hacked Routers and DNS Hijacking." SecurityWeek, April 2026.

[26] Trend Micro. "Router Roulette: Cybercriminals and Nation-States Sharing Compromised Networks." Trend Micro, 2024.

[27] BleepingComputer. "TheMoon Malware Infects 6,000 ASUS Routers in 72 Hours for Proxy Service." BleepingComputer, March 2024.

[28] Independent Security Evaluators. "SOHOpelessly Broken 2.0." ISE, 2019.

[29] Cisco Talos. "The Many Vulnerabilities Talos Discovered in SOHO and Industrial Wireless Routers Post-VPNFilter." Cisco Talos Intelligence Blog, 2023.

[30] CISA. "Secure by Design Alert: Security Design Improvements for SOHO Device Manufacturers." CISA, January 2024.

[31] CISA. "Security Design Improvements for SOHO Device Manufacturers (PDF)." CISA, January 31, 2024.

[32] FCC. "National Security Determination on the Threat Posed by Foreign-Produced Consumer Routers." FCC, March 2026.

[33] Morgan Lewis. "FCC Adds New Foreign-Made Consumer Routers, Signaling the Administration's Shifting ICTS Approach." Morgan Lewis, April 2026.

[34] National Security Agency. "Zero Trust Implementation Guideline Phase One." NSA, January 2026.

[35] CISA. "Releases Part One of Zero Trust Microsegmentation Guidance." CISA, July 2025.

[36] BleepingComputer. "Authorities Disrupt Router DNS Hijacks Used to Steal Microsoft 365 Logins." BleepingComputer, April 2026.

[37] The Hacker News. "TheMoon Botnet Resurfaces, Exploiting EoL Devices to Power Criminal Proxy." The Hacker News, March 2024.

[38] Lawfare Institute. "The Cyber Monoculture Risk." Lawfare, October 2021.

[39] TechTarget. "Hundreds of New Vulnerabilities Found in SOHO Routers." TechTarget Security, 2024.

[40] TechRepublic. "US Bans New Foreign-Made Routers, Citing 'Unacceptable' Security Risks." TechRepublic, 2026.