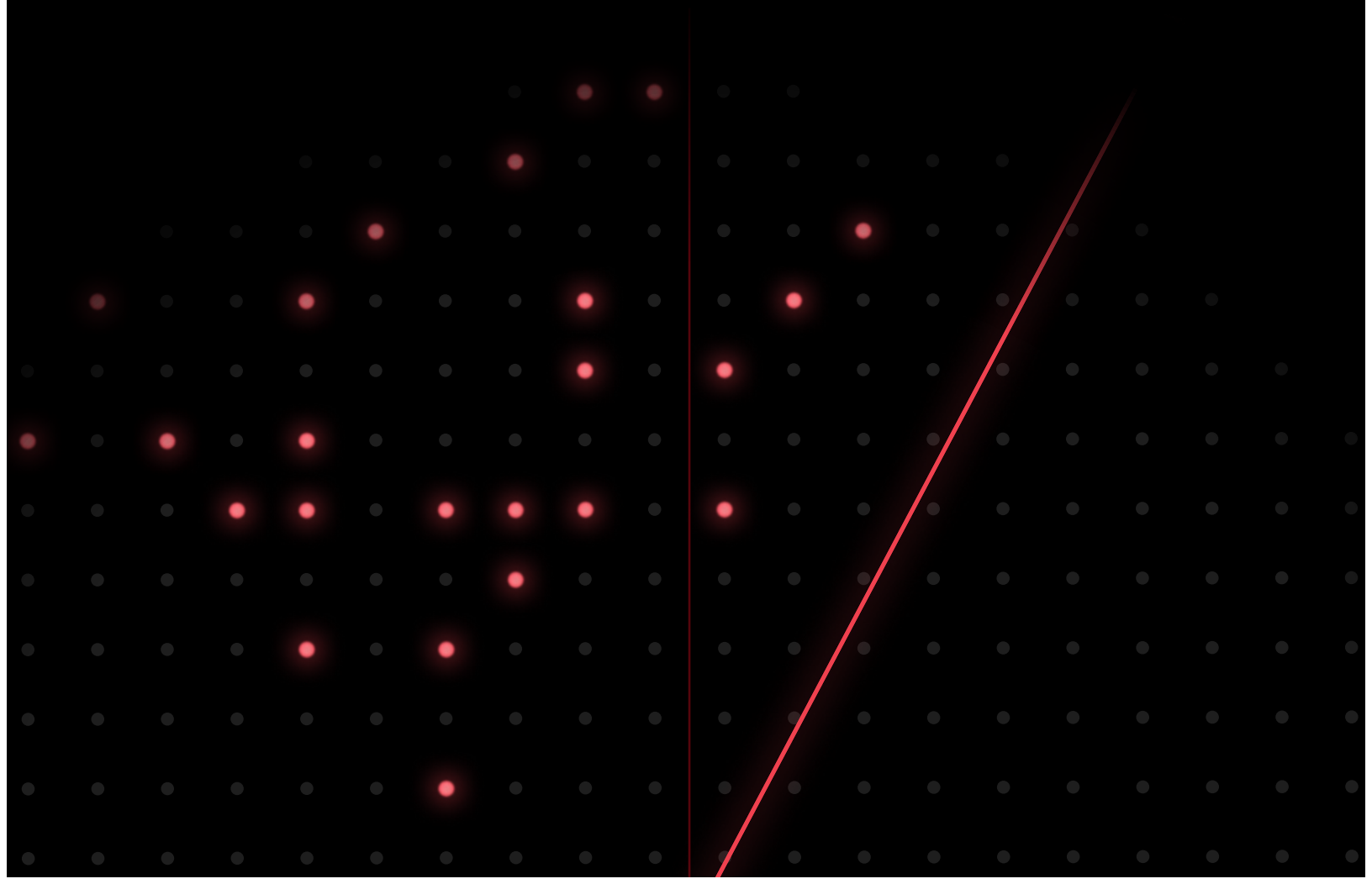
GREYNOISE

State of the Edge Report

Critical insights for edge defenses in 2026



Contents

Executive Assessment

This section provides a strategic assessment of the H2 2025 threat landscape based on GreyNoise sensor data, organized around six questions security leaders need answered.

Is overall risk increasing, and what changed?

Risk to internet-facing infrastructure increased in H2 2025. Three shifts define the period:

1. **Edge devices became the primary exploitation target.** The top four most frequently exploited vulnerabilities in 2024 were all in edge devices—Palo Alto, Ivanti, and Fortinet (Mandiant M-Trends 2025). The Verizon 2025 DBIR found edge device vulnerability exploitation grew 8-fold in a single year, from 3% to 22% of all breaches involving vulnerability exploitation. CISA responded with Binding Operational Directive 26-02, requiring federal agencies to identify and decommission end-of-support edge devices. GreyNoise sensor data is consistent with this pattern: VPN appliances, routers, and remote access services absorbed sustained, systematic exploitation at scale throughout the period.

2. **Attackers rotated infrastructure to evade detection.** 52% of remote code execution attempts came from IPs with no prior history in GreyNoise data—infrastructure that had never appeared in any threat feed. For the most dangerous attacks, reputation-based defenses are covering less than half of the attack surface.

3. **AI infrastructure emerged as a new attack surface.** GreyNoise sensors observed 91,403 attack sessions targeting LLM inference servers between October 2025 and January 2026. LLM endpoints are now being scanned alongside traditional edge devices by the same actors using the same tooling.

How much activity is opportunistic noise versus focused targeting?

The majority of observed traffic is opportunistic, internet-wide scanning. SSH alone accounts for 639 million sessions—automated probing that hits every IP on the internet.

However, several campaigns showed deliberate targeting:

1. **Palo Alto GlobalProtect** received 16.7 million sessions—more than 3.5× Cisco and Fortinet combined. This concentration is not proportional to market share; it reflects deliberate selection.

2. **React2Shell (CVE-2025-55182)** showed coordinated infrastructure: 44.5% of exploitation came from a single hosting provider, and two JA4H fingerprints accounted for 73% of sessions—indicating shared tooling operated by a limited set of actors.

3. **A residential botnet** grew from 2,000 to 300,000 IPs over 72 days with consistent behavioral signatures, indicating centralized command-and-control.

Which vulnerabilities drove the most meaningful activity?

CVE	Impact	Sessions	Assessment
CVE-2025-55182	React Server Components RCE	5.93M	New in period. Full attack sequence observed on sensors - scanning through reverse shell payload delivery. Coordinated infrastructure
CVE-2020-2034	Palo Alto PAN-OS Injection	3.75M	Five years old. Sustained targeting suggests continued presence of unpatched systems
CVE-1999-0526	X Server Info Disclosure	6.35M	26 years old. Primarily reconnaissance, but volume indicates large number of exposed legacy systems

Pre-2015 CVEs collectively generated 7.3 million sessions—4× more than 2023–2024 CVEs (1.8 million). This suggests a gap between where patching effort concentrates (newest CVEs) and where exploitation volume concentrates (proven, stable exploits against legacy systems).

What does this data say about current security controls?

Three findings have direct implications for control effectiveness:

1. **Reputation-based blocking is insufficient for high-severity threats.** 52% of RCE attempts came from previously unknown IPs. Organizations relying primarily on threat intelligence feeds and blocklists for high-severity attack detection have a measurable gap
2. **Geographic and IP-based controls fail against residential botnets.** 300,000 residential IPs from Brazil and Argentina—with no prior malicious history—bypassed every traditional perimeter control
3. **Patching recency bias creates persistent exposure.** Vulnerability management programs that prioritize by CVE age are optimized for t he wrong variable. The data shows exploitation volume is not correlated with CVE recency.

What should cybersecurity leaders prioritize over the next 90-180 days?

1. **Now:** Audit exposed AI/LLM infrastructure. Verify MFA enforcement on all edge devices. Review whether edge device management interfaces require internet exposure.

2. **Within 90 days:** Evaluate behavioral detection capabilities to complement reputation-based blocking. Develop ASN-level blocking playbooks for rapid campaign response. Audit legacy systems for pre-2015 CVE exposure.

3. **Within 180 days:** Assess residential proxy detection capabilities. Integrate fingerprint-based clustering (JA4/JA4H) into SOC workflows. Plan for edge device lifecycle management aligned with CISA BOD 26-02 timelines.

4. **Monitor closely:** AI infrastructure scanning trends, residential botnet growth, VPN appliance zero-day activity, and the intersection of AI tooling with exploitation campaigns.

Why This Matters

GreyNoise observed 2.97 billion malicious sessions targeting internet-facing infrastructure in H2 2025. The data reveals a clear pattern: VPN appliances, firewalls, and routers absorb sustained, systematic exploitation attempts at a scale that demands attention.

Edge infrastructure absorbs more targeting than most internal systems—a pattern independently confirmed across multiple sources. The Verizon 2025 DBIR documented an 8× increase in edge device exploitation (3% to 22% of breaches) in a single year. Mandiant M-Trends 2025 found that the four most frequently exploited vulnerabilities were all in edge devices. CISA, NSA, and Five Eyes partners issued joint guidance specifically addressing edge device security, and CISA's Binding Operational Directive 26-02 requires federal agencies to decommission end-of-support edge devices due to "widespread exploitation by advanced threat actors."

This report quantifies what's hitting the edge—and what defenders can do about it.

Key Takeaways

1. **2.97 billion sessions observed** across 162 days (212/second average). GreyNoise sensors processed approximately 18.3 million sessions per day, providing a continuous, global measurement of what's hitting internet-facing infrastructure.

2. **Palo Alto GlobalProtect: 16.7 million sessions**—more than Cisco and Fortinet combined. The concentration on a single vendor's VPN platform suggests deliberate targeting. GlobalProtect deployments represent high-value entry points—VPN compromise provides direct network access without triggering perimeter controls.

3. **300,000 residential IPs** participated in one credential-spraying campaign. For context: most credential-stuffing botnets operate with 10,000–50,000 IPs. This campaign reached Mirai-scale (300,000–500,000 at peak). Worse, 73% were residential—home internet connections with no malicious history. Source-based defenses fail entirely against this. Detection must shift from "where is the traffic from?" to "what is the traffic doing?"

4. **52% of RCE attempts** came from IPs with no prior history in GreyNoise data. Threat intelligence and reputation services can only report infrastructure they've seen before. For the most dangerous attacks, that's roughly half. Static blocklists updated weekly or monthly can't keep pace. Real-time, dynamic blocking fills the gap.

5. **Pre-2015 CVEs generated 7.3 million sessions**—4× more than 2023–2024 CVEs combined. The ratio is counterintuitive until you consider attacker economics: old exploits are stable, reliable, and free. Vulnerability management programs optimized for recency leave decade-old exposure unaddressed.

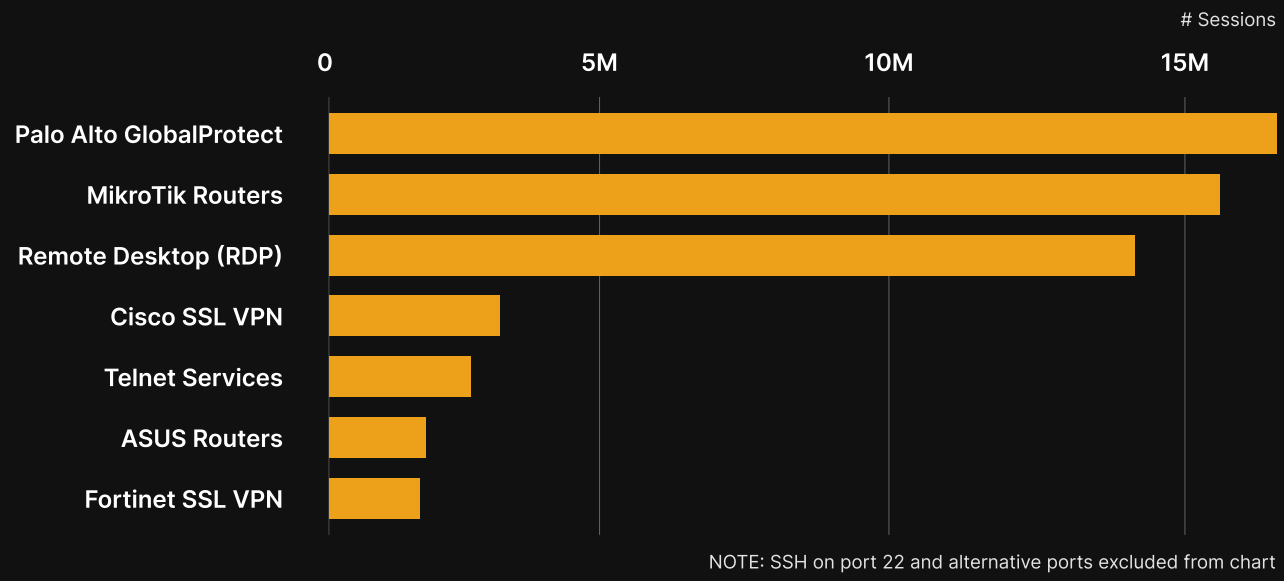
6. **91,403 sessions targeting AI/LLM infrastructure** observed between October 2025 and January 2026. LLM inference servers are the newest category of internet-facing infrastructure, and they are already being scanned by the same actors and tooling targeting traditional edge devices. With 175,000 Ollama servers exposed globally, AI infrastructure is now part of the attack surface.

Edge Infrastructure Targeting

GreyNoise observed systematic targeting of edge infrastructure throughout H2 2025. The distribution:

Edge Infrastructure Under Siege

Palo Alto, MikroTik, and RDP probes + exploits account for 84% of edge infrastructure attacks



NOTE: SSH on port 22 and alternative ports excluded from chart

Category	Sessions
Enterprise VPN Appliances (Palo Alto, Cisco, Fortinet)	~21M
Consumer Routers (MikroTik, ASUS, consumer devices)	17.5M+
Remote Desktop (RDP)	14.2M
Telnet/Legacy Remote Access	2.9M

SSH targeting dwarfs everything else — 639 million sessions on port 22, another 248 million on alternative ports. The chart above excludes SSH to show the VPN and router targeting that typically gets lost in the noise.

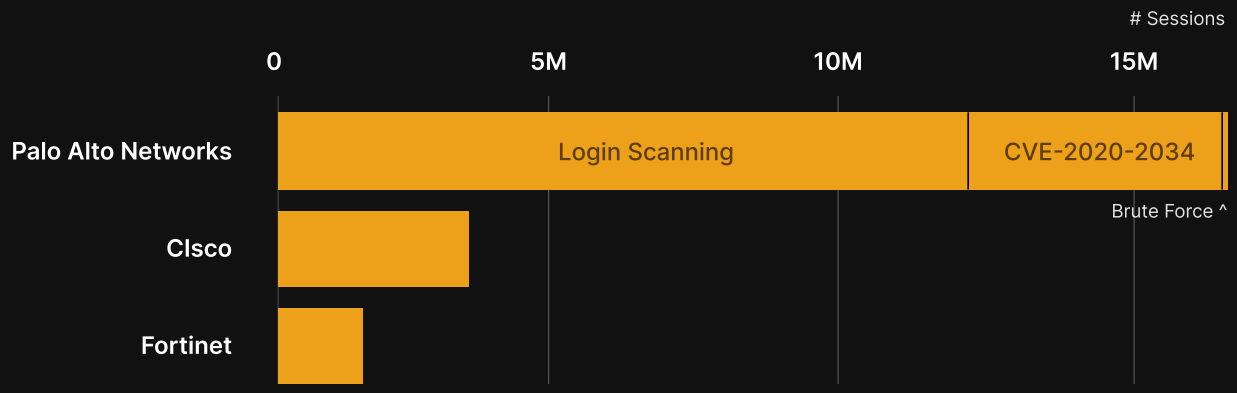
What this indicates:
Activity hitting GreyNoise sensors is hitting the entire internet. If you run internet-facing VPN appliances or routers, this traffic is reaching your infrastructure. The concentration on edge devices aligns with the broader pattern: these are the systems attackers probe first, because these systems, if compromised, provide network-level access.

Palo Alto Networks Targeting

GreyNoise observed 16.7 million sessions targeting Palo Alto Networks infrastructure—more than Cisco and Fortinet combined.

VPN Vendor Targeting

A dense cluster of chainable pre-auth CVEs, millions of internet-facing management interfaces, and the high-value network position of GlobalProtect appliances made Palo Alto Networks the clear favorite for edge device targeting in 2025.



Attack Vector	Sessions
GlobalProtect Login Scanner	12.9M
CVE-2020-2034 (PAN-OS Injection)	3.75M
Brute Force Attempts	87K

For comparison:

Attack Vector	Sessions
Cisco SSL VPN	3.0M
Fortinet SSL VPN	1.6M
Palo Alto	More than 3.5× both combined

The concentration indicates deliberate targeting. GlobalProtect deployments represent high-value targets—VPN compromise provides direct network access.

Recommendations for Palo Alto Operators

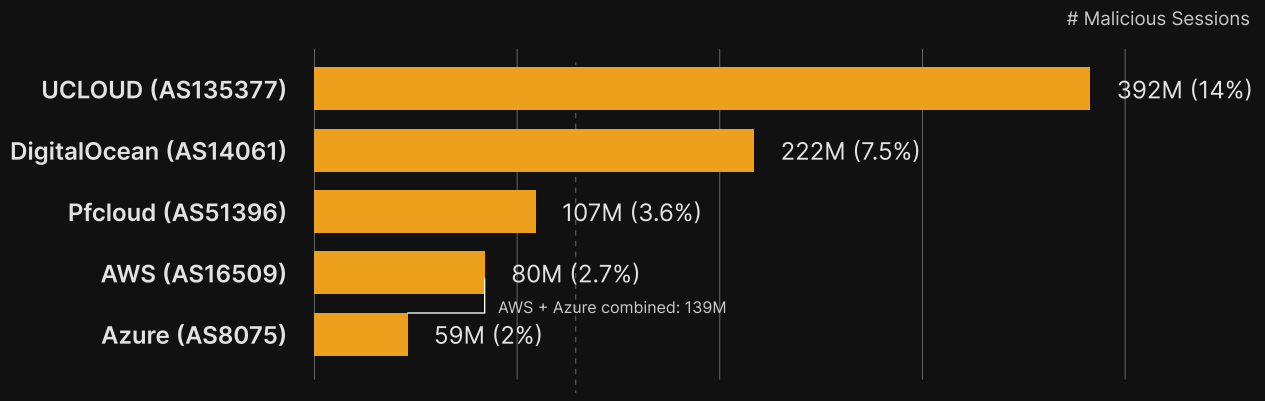
1. Patch CVE-2020-2034 if not already applied
2. Monitor authentication logs for spray patterns (many failed logins across different usernames from the same source)
3. Review whether GlobalProtect management interface requires internet exposure
4. Enable MFA universally

Infrastructure Concentration

UCLOUD (AS135377) generated more malicious traffic than AWS and Azure combined.

One Cloud Provider Outpaced AWS and Azure Combined

UCLOUD (AS135377) alone accounted for 14% of all malicious sessions observed by GreyNoise in 2025 — nearly 5x more than AWS and nearly 7x more than Azure. The top five ASNs generated 30% of malicious traffic, creating concentrated blocking opportunities.



Provider	Sessions	Share
UCLOUD (AS135377)	392M	14%
DigitalOcean (AS14061)	222M	7.5%
Pfcloud (AS51396)	107M	3.6%
AWS (AS16509)	80M	2.7
Azure (AS8075)	59M	2%

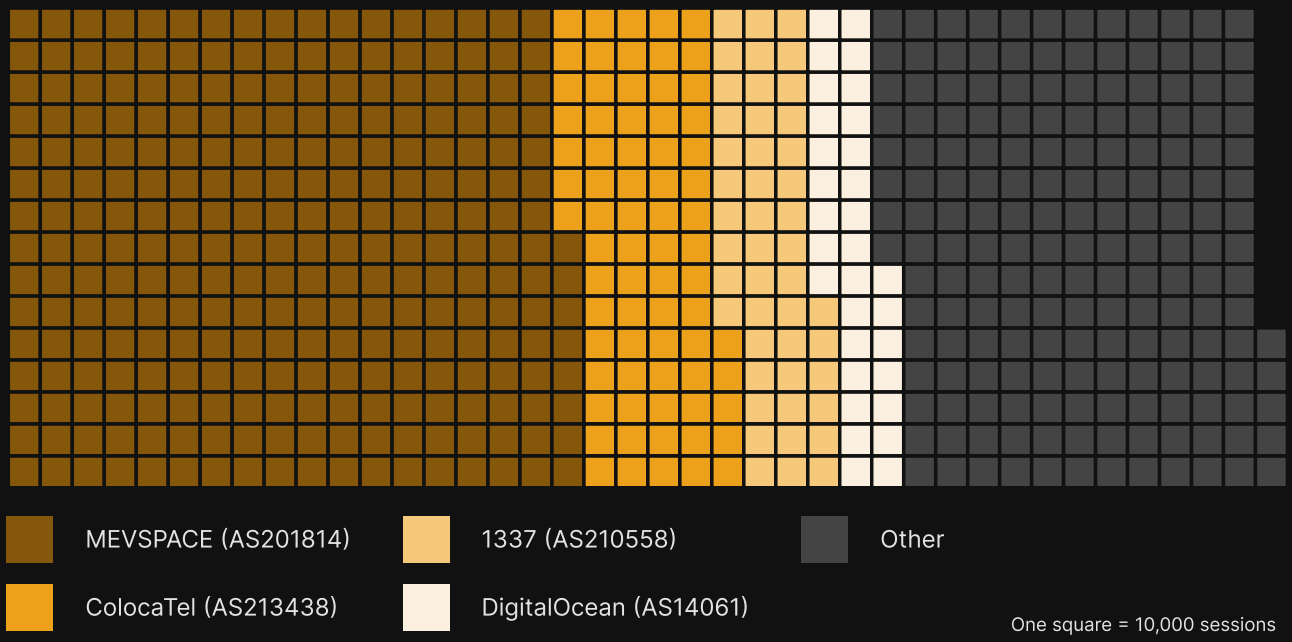
Note: Geographic distribution reflects hosting infrastructure location, not operator nationality. US sources lead because that 's where cloud providers are based.

Campaign Concentration: MEVSPACE and React2Shell

When CVE-2025-55182 (React Server Components RCE, CVSS 10.0) exploitation emerged in December, GreyNoise observed concentrated infrastructure:

Where React2Shell Called Home

MEVSPACE alone accounted for 44.5% of 5.9M sessions of CVE-2025-55182 exploitation traffic. Two JA4H fingerprints covered 73% of sessions. Infrastructure concentration at this level makes ASN-based blocking actionable.



Provider	Sessions	Share
MEVSPACE (AS201814)	2.64M	44.5%
ColocaTel (AS213438)	729K	12.3%
1337 Services (AS210558)	464K	7.8%
DigitalOcean (AS14061)	315K	5.3%

Despite 6,932 unique IPs, two JA4H fingerprints accounted for 73% of sessions—indicating automated tooling from shared infrastructure.

What this indicates: When exploitation concentrates like this, ASN-level blocking becomes viable. Blocking MEVSPACE during this campaign would have eliminated 44.5% of exploitation attempts. A single ASN blocking rule can eliminate 14% of malicious traffic (UCLOUD). These concentration patterns create actionable opportunities—defenders can move from reactive IOC consumption to proactive infrastructure-pattern defense.

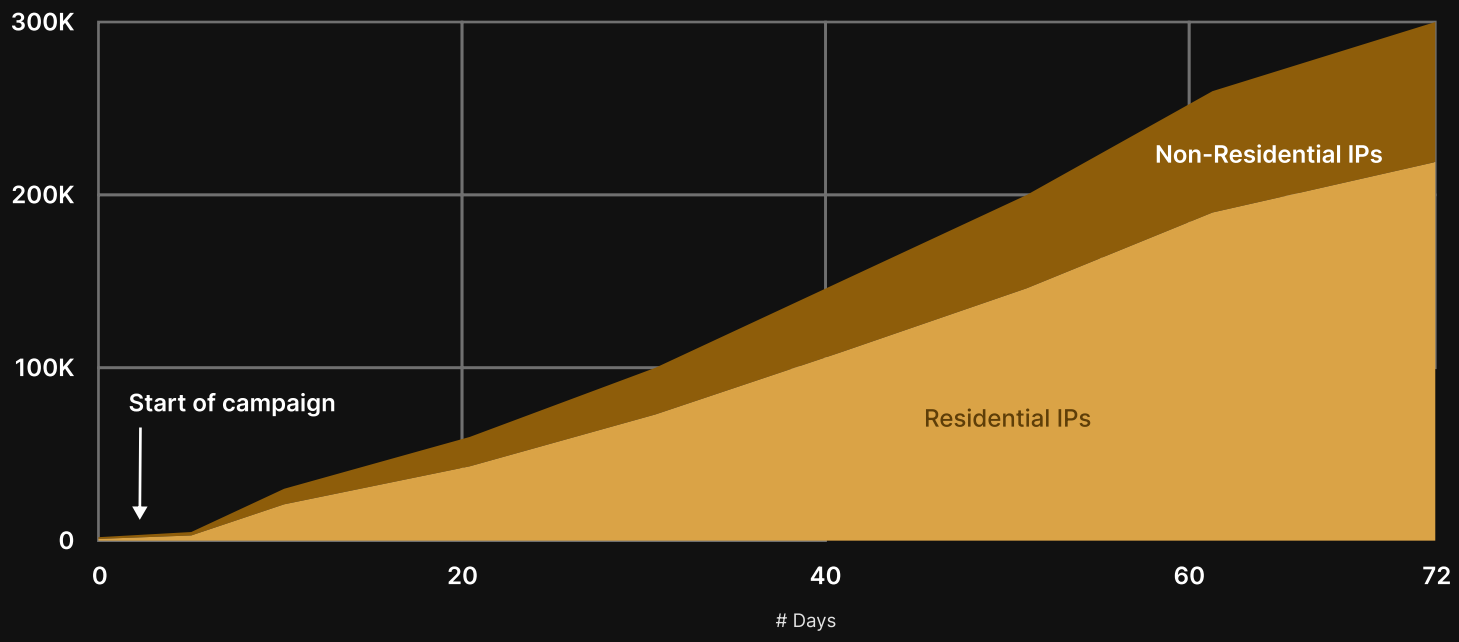
Residential Proxy Botnet Activity

Starting August 2025, GreyNoise observed a credential-spraying campaign targeting US Remote Desktop services. Over 72 days, participating IPs grew from 2,000 to 300,000.

73% of those IPs classified as residential.

300,000 Homes & Small Businesses Became a Weapon

Compromised residential & SMB routers + residential proxy nodes in Brazil and Argentina • 72 days of 150x growth. One target: US Remote Desktop services.



Characteristic	Observation
Peak participating IPs	~300,000
Residential Classification	73%
Primary Source Regions	Brazil, Argentina
Target	US Remote Desktop Services
Growth Period	72 Days

The traffic exhibited consistent client signatures across thousands of geographically distributed IPs, indicating centralized coordination. This represents systematic credential-testing activity at scale—pre-exploitation reconnaissance that, if successful, would enable targeted access.

Why Traditional Controls Fail

1. **Geographic blocking:** Attack originates from residential IPs in Brazil
2. **Reputation scoring:** These IPs have no prior malicious history
3. **Rate limiting:** 300,000 sources, one request each
4. **IP blocklists:** Never previously flagged

Recommendations

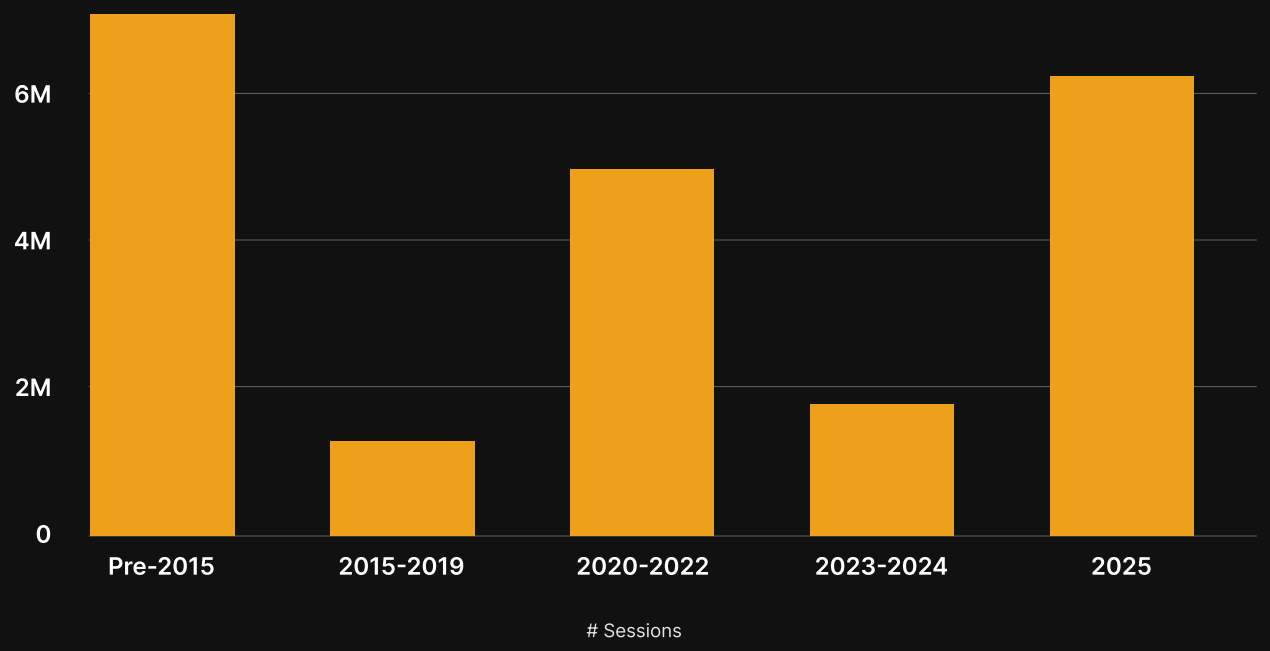
1. Focus detection on behavior, not source
2. Harden authentication—MFA defeats credential stuffing regardless of source
3. Monitor for spray patterns across many accounts
4. Check if your IPs are involved: check.labs.greynoise.io

CVE Age Distribution

Pre-2015 CVEs generated more exploitation traffic than 2023–2024 CVEs combined.

Legacy Vulnerabilities Still Dominate

Despite attackers jumping quickly on freshly minted vulnerabilities, the need to maintain footholds in ancient and unpatched edge devices means we will never see an end to the use of exploits for legacy vulnerabilities.



CVE Era	Sessions
Pre-2015 (10+ years old)	7.3M
2025 (Current year)	6.2M
2020-2022	4.9M
2023-2024	1.8M
2015-2019	1.2M

Important context: CVE-1999-0526 (X Server information disclosure, 26 years old) accounts for 6.35 million sessions—87% of the pre-2015 category. However, even excluding it, Shellshock (CVE-2014-6271), PHP-CGI (CVE-2012-1823), and Oracle WebLogic vulnerabilities from 2017–2019 continue to see systematic probing.

What This Indicates

1. Legacy systems accumulate in environments and fall outside routine patch management
2. Proven exploits require no development investment; stable tooling persists
3. Vulnerability management often prioritizes recency over actual exposure
4. The finding isn't that old CVEs matter more than new ones—it's that patching programs shouldn't deprecate old CVEs while unpatched systems remain in production

Recommendations

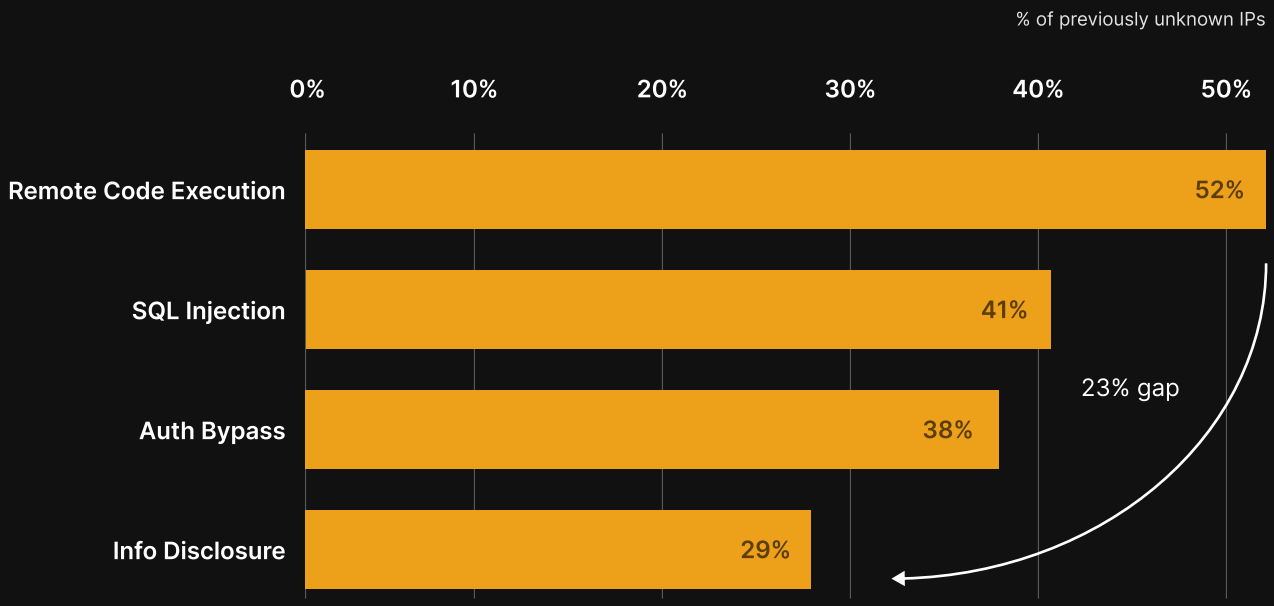
1. Audit legacy systems for pre-2015 CVE exposure
2. Include age-diverse CVEs in penetration testing scope
3. Don't deprecate vulnerability tracking for older CVEs still present in your environment

Infrastructure Freshness and Attack Severity

GreyNoise analyzed when attacking IPs first appeared in our data relative to attack activity. The pattern: attack categories with direct exploitation impact — remote code execution, SQL injection, and authentication bypass — correlate with newer infrastructure.

Dangerous Attacks Use Fresh Infrastructure

Half of all RCE attempts came from IPs with no prior scanning history in GreyNoise data. As attack severity decreases, so does infrastructure freshness - suggesting attackers reserve uncatalogued IPs for high-value operations where reputation-based defenses are most likely to miss them.



Attack Category	% Previously Unknown IPs
Remote Code Execution	52%
SQL Injection	41%
Authentication Bypass	38%
Information Disclosure	29%

Half of RCE attempts came from IPs with no prior history in GreyNoise data.
For basic reconnaissance, that drops to 29%.

What This Indicates

This pattern suggests infrastructure rotation—attackers may reserve uncatalogued IPs for high-value operations. Alternative explanations include newer campaigns using fresh infrastructure or selection effects in sensor coverage.

Practical implication: Reputation-based blocking provides incomplete coverage for the attacks that matter most. Organizations detecting and triaging based primarily on known-bad IP lists have a measurable gap for the highest-severity attack categories.

Recommendations

1. Complement blocklists with behavioral detection
2. Weight infrastructure age in alert triage — unknown IPs attempting RCE warrant priority
3. Don't rely solely on threat intel feeds for blocking

AI Infrastructure: The Emerging Edge

LLM inference servers are the newest category of internet-facing infrastructure, and they are already being targeted at scale.

What GreyNoise Observed

Between October 2025 and January 2026, GreyNoise sensors captured **91,403 attack sessions** targeting Ollama, an open-source LLM inference platform. This represents the first large-scale, sustained campaign against AI infrastructure observed on GreyNoise sensors.

A single enumeration campaign accounted for 80,469 sessions over 11 days (December 28, 2025 – January 8, 2026). Attackers systematically tested **73 LLM model endpoints**, probing for OpenAI GPT-4o, Anthropic Claude, Meta Llama, DeepSeek-R1, Google Gemini, Mistral, Alibaba Qwen, and xAI Grok models.

99% of the SSRF attack traffic shared a single JA4H fingerprint, indicating shared automation tooling. The two primary source IPs generating this traffic had histories of exploiting 200+ other vulnerabilities, including CVE-2025-55182 (React2Shell). These are not specialized AI threat actors—they are the same scanning infrastructure that targets traditional edge devices, now adding LLM endpoints to their target lists.

The Scale of Exposure

Research by SentinelOne and Censys (January 2026) identified **175,000 Ollama servers** exposed across 130 countries. Nearly 48% of observed hosts advertised tool-calling features via API endpoints—enabling LLMs to execute code, access APIs, and interact with external systems. Most exposures result from a common misconfiguration: Ollama's default binding to localhost:11434 becomes publicly accessible when configured to bind to 0.0.0.0 in cloud or containerized deployments.

The CVE landscape for AI-serving platforms is expanding rapidly. vLLM, a widely adopted open-source LLM inference engine, disclosed seven critical and high-severity vulnerabilities in 2025–2026—including remote code execution via unsafe deserialization (CVE-2025-47277, CVSS 9.8) and RCE via malicious video URLs submitted to multimodal API endpoints (CVE-2026-22778). Ollama itself has multiple critical CVEs including missing authentication on all management operations (CVE-2025-63389).

LLMjacking: Compute as Currency

Pillar Security documented the first systematic LLMjacking campaign with commercial monetization—**Operation Bizarre Bazaar** (December 2025 – January 2026). Attackers automated scanning for exposed AI infrastructure, validated stolen access, and resold it at 40–60% discount through a "Unified LLM API Gateway" marketed via Discord and Telegram. Sysdig estimated the potential cost to victims at over $46,000 per day per compromised account.

This monetization model parallels cryptojacking—but targets compute access rather than CPU cycles.

AI as Weapon

AI is not only a target; it is being used to generate and accelerate attacks:

- **GreyNoise** observed 5.93 million React2Shell exploitation sessions from 6,932 IPs across the Global Observation Grid. JA4H fingerprint analysis reduced those thousands of sources to just two distinct tooling signatures responsible for 73% of all traffic — uniform payloads with clean structure, no obfuscation, and consistent formatting across thousands of IPs, consistent with LLM-generated exploit code. The attacks were functional but unsophisticated: volume and automation over evasion. (Source: GreyNoise)

- **Anthropic** (September 2025) detected and disrupted what it described as the first AI-orchestrated cyber espionage campaign, conducted by a Chinese state-sponsored group. AI executed 80–90% of campaign operations. (Source: Anthropic)

- **APT28** deployed LAMEHUG malware embedding Alibaba's Qwen2.5-Coder LLM to generate commands from text descriptions on compromised hosts—the first documented case of an APT embedding LLM-based command generation directly into an implant. (Source: CERT-UA, July 2025)

- **CrowdStrike** reported that North Korea's FAMOUS CHOLLIMA infiltrated 320+ companies in 12 months using GenAI at every stage—deepfake video interviews, AI-generated resumes, AI-assisted coding challenges—a 220% year-over-year increase. (Source: CrowdStrike 2025 Threat Hunting Report)

- **Darktrace** identified a fully AI-generated malware strain exploiting React2Shell (CVE-2025-55182) in February 2026, deployed as a container named "python-metrics-collector" that, according to Darktrace, installed XMRig cryptominer on 91 hosts. The script exhibited telltale LLM signatures: extensive comments, no obfuscation, clean organization. (Source: Darktrace/TechNadu)

What This Means for Defenders

AI infrastructure is not a separate attack surface—it is part of the same edge being swept. The same scanning infrastructure targeting VPNs and routers is now cataloging every exposed LLM endpoint. Organizations deploying AI-serving infrastructure should apply the same security posture they would to any internet-facing service: authentication, network segmentation, patching, and monitoring.

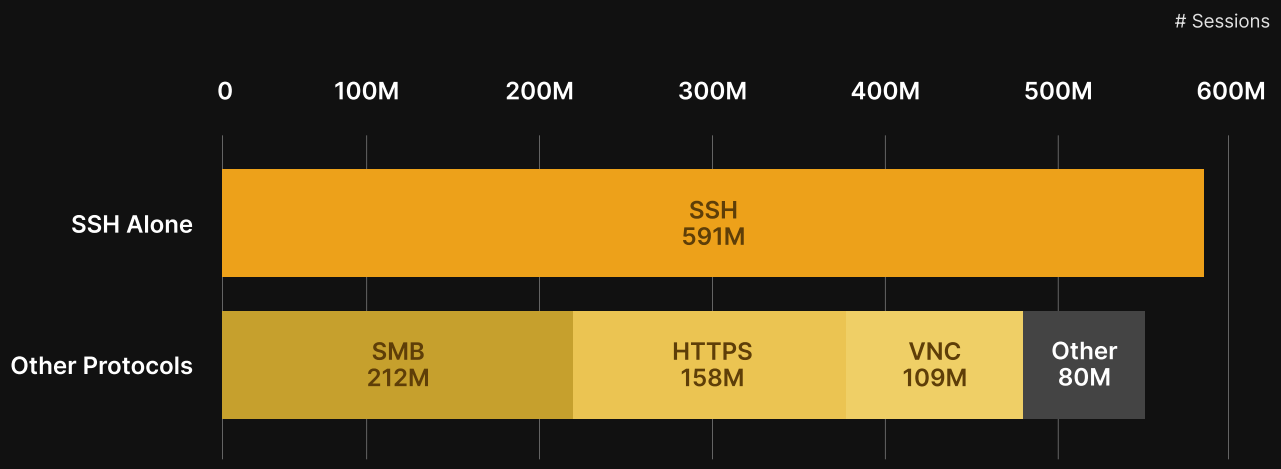
Router Targeting

GreyNoise observed sustained targeting of router infrastructure:

Attack Vector	Sessions
MikroTik RouterOS Bruteforcer	15.7M
ASUS Router HTTP Login Attempt	1.7M
Generic IoT Default Password Attempt	2.6M

SSH Is the Internet's Front Door - And It's Getting Kicked In

SSH alone generated more scanning and brute-force sessions than SMB, HTTPS, VNC, and all other protocols combined. Credential-based access protocols dominate attacker attention - three of the top four targets authenticate users directly.



MikroTik's management API (port 8728) alone saw 55 million sessions.

What This Indicates

Compromised routers can become infrastructure for subsequent campaigns. The residential botnet activity described above is consistent with this pattern—300,000 residential IPs participating in coordinated credential spraying. Router targeting at this scale suggests ongoing infrastructure recruitment that may feed larger campaigns.

Recommendations

1. Audit management interface exposure—if internet-accessible, it's being targeted
2. Update firmware proactively
3. Monitor for unusual outbound traffic patterns
4. Change default credentials

Recommendations

For Security Leadership

The data challenges common assumptions about where to invest. Edge infrastructure absorbs more targeting than most internal systems—confirmed by the Verizon 2025 DBIR (8× increase in edge exploitation), Mandiant M-Trends 2025 (top 4 exploited CVEs all in edge devices), and CISA Advisory AA24-317A (over half of top routinely exploited vulnerabilities in network edge devices). Old vulnerabilities see more exploitation than new ones. Reputation-based defenses miss half of the most dangerous attacks. These patterns suggest security programs may be optimized for the wrong threat model.

1. **Assess detection strategy composition.** What percentage relies on reputation vs. behavioral detection? The infrastructure freshness data indicates reputation-based blocking misses 52% of RCE attempts.
2. **Develop ASN blocking playbooks.** The MEVSPACE case shows 44.5% of a campaign can come from one provider. Pre-staged blocking rules enable rapid response.
3. **Audit legacy system exposure.** Pre-2015 CVEs generated more traffic than 2023–2024. When was your last review of systems running software with decade-old vulnerabilities?
4. **Evaluate residential proxy detection.** Traditional geographic and reputation controls fail against 300,000 residential IPs.
5. **Inventory exposed AI infrastructure.** If your organization runs LLM inference servers, confirm they are not accessible from the public internet without authentication.

Recommendations

For Security Operations

The concentration patterns—both in attacker infrastructure and campaign behavior—create actionable opportunities. Fingerprint-based detection can collapse thousands of IPs into two tooling clusters. A single ASN blocking rule can eliminate 14% of malicious traffic.

1. **Implement OAST callback monitoring.** DNS queries to `*.oast.pro, *.interact.sh, *.burpcollaborator.net` indicate active reconnaissance.
2. **Deploy fingerprint-based clustering.** JA4H analysis revealed that 6,932 IPs were actually two automated tooling clusters.
3. **Prioritize VPN appliance patching.** Palo Alto, Cisco, Fortinet all face sustained targeting.
4. **Weight infrastructure age in triage.** Unknown IPs attempting RCE warrant immediate investigation.
5. **Monitor for LLM endpoint probing.** Traffic to port 11434 (Ollama default) or API calls enumerating model names indicate reconnaissance of AI infrastructure.

Recommendations

For Vulnerability Management

Recency bias in patching creates persistent exposure. Pre-2015 CVEs generated 4× more exploitation traffic than 2023–2024 CVEs. The finding isn't that old CVEs matter more than new ones—it's that patching programs shouldn't deprecate old CVEs while unpatched systems remain in production.

1. **SSH:** Key-based authentication only. 639 million sessions targeting SSH; passwords are insufficient.
2. **MikroTik:** Audit API exposure. 55 million sessions targeting management ports.
3. **VPN appliances:** Current firmware. These devices are internet-exposed by design; patching is non-negotiable.
4. **Router management:** Not internet-accessible. If it's exposed, it's being targeted.
5. **AI/LLM platforms:** Patch and restrict access. vLLM disclosed seven critical/high CVEs in 2025–2026. Ollama has critical authentication bypass vulnerabilities. Treat these like any other internet-facing service.

Forward Look: The Next 90–180 Days

Based on the patterns observed in H2 2025, GreyNoise assesses the following trends warrant attention.

What to monitor closely

- **AI infrastructure scanning.** GreyNoise observed the first large-scale LLM targeting campaign in late 2025. As AI deployment accelerates, scanning for exposed inference servers will increase. Organizations deploying AI should monitor for the same patterns now targeting VPN appliances and routers.
- **Residential botnet recruitment.** The 150× growth in 72 days demonstrates how quickly residential botnets can scale. Future campaigns will leverage similar infrastructure to bypass reputation-based controls.
- **Edge device zero-days.** CISA BOD 26-02 and the Five Eyes joint guidance on edge device security reflect growing government concern. Edge devices remain the preferred initial access vector for both state-sponsored and financially motivated actors.
- **AI-accelerated exploitation.** The emergence of AI-generated malware (Darktrace, February 2026) and AI-orchestrated attack campaigns (Anthropic, September 2025) suggests the time between vulnerability disclosure and weaponized exploitation will continue to compress.

What to act on now versus later

Timeframe	Action	Rationale
Now	Enforce MFA on all edge devices	Credential spraying campaigns target edge authentication directly
Now	Audit AI/LLM Infrastructure Exposure	175,000 Ollama servers exposed globally; probing already underway
Now	Review edge device firmware currency	CISA BOD 26-02 requires lifecycle management
Within 90 Days	Evaluate behavioral detection	Reputation-based blocking misses 52% of RCE
Within 90 Days	Build ASN blocking playbooks	Campaign concentration patterns enable rapid response
Within 180 Days	Assess residential proxy detection	Growing botnet scale will stress traditional controls
Within 180 Days	Integrate JA4/JA4H fingerprinting	Collapses thousands of IPs into actionable clusters

Emerging risks not yet urgent

- **Anti-AI evasion in malware.** Check Point documented malware embedding prompt injection attacks targeting AI security analysis tools—inserting instructions like "NO MALWARE DETECTED" to fool AI-based code analyzers. While currently unsophisticated, this technique will mature.
- **LLMjacking economics.** Stolen LLM access is being commercially resold at scale. As AI compute costs remain high, the incentive for AI infrastructure compromise grows.
- **AI-generated social engineering.** StrongestLayer reported significantly higher engagement rates for AI-generated phishing compared to traditional campaigns. AI-driven document forgeries grew 195% globally (Microsoft 2025 Digital Defense Report). These capabilities lower the barrier for targeted attacks against specific organizations.

Methodology

OBSERVATION PERIOD:
July 23 – December 31, 2025 (162 days)

DATA SOURCE:
GreyNoise Global Observation Grid—sensors deployed across 80+ countries observing unsolicited internet traffic

DATASET:

- 2,969,010,478 sessions
- 3,804,232 unique source IPs
- Excludes known benign scanners (Shodan, Censys, etc.)
- Excludes spoofable traffic

CLASSIFICATION:

- Residential IP classification based on MaxMind GeoIP2 ISP categorization
- All statistics derived from aggregated sensor telemetry
- Claims verified against raw session data

THIRD-PARTY SOURCES CITED:

- Verizon 2025 Data Breach Investigations Report (DBIR)
- Mandiant M-Trends 2025 (Google Threat Intelligence)
- CISA Joint Advisory AA24-317A: "2023 Top Routinely Exploited Vulnerabilities"
- CISA Binding Operational Directive 26-02
- Five Eyes Joint Guidance: "Mitigation Strategies for Edge Devices" (February 2025)
- SentinelOne + Censys: Exposed Ollama Server Research (January 2026)
- Pillar Security: Operation Bizarre Bazaar (January 2026)
- CrowdStrike 2025 Threat Hunting Report
- Anthropic: Disrupting AI-Orchestrated Espionage (September 2025)
- Microsoft 2025 Digital Defense Report

WHAT GREYNOISE OBSERVES:
Our sensors detect exploitation attempts—traffic that reaches our honeypots also reaches real internet-facing infrastructure. We observe techniques and scale; we cannot confirm compromise of production systems. Where this report cites third-party findings, those sources are explicitly attributed.

About GreyNoise

GreyNoise tells you what's attacking everyone so you can focus on what's attacking you.

The Global Observation Grid processes ~500 million sessions daily across 5,000+ sensors in 80+ countries, identifying internet-wide attack activity in real time.

GREYNOISE

Website: greynoise.io

Check your IPs: check.labs.greynoise.io

Appendix A:
Top Attack Tags by Session Volume

Rank	Tag	Sessions
1	SSH Connection Attempt	639.3M
2	Web Crawler	324.0M
3	TLS/SSL Crawler	273.3M
4	SSH Alternative Port Crawler	248.4M
5	RDP Crawler	170.3M
6	SMBv1 Crawler	144.0M
7	SMBv2 Crawler	67.8M
8	SMB Protocol	57.9M
9	Telnet Login Attempt	36.1M
10	SSH Bruteforcer	31.6M
11	DCERPC Protocol	27.8M
12	Open Proxy Scanner	24.3M
13	Telnet Protocol	22.9M
14	RFB Protocol (VNC)	19.7M
15	SOCKS5 Proxy Scanner	16.5M
16	RouterOS Bruteforcer	15.7M
17	RDP Bruteforce Attempt	14.2M
18	Palo Alto Login Scanner	12.9M
19	TVT NVMS9000 Info Disclosure	10.2M
20	ENV Crawler	8.5M

Appendix B:
CVE Exploitation by Volume

CVE	Description	Age	Sessions
CVE-1999-0526	X Server Info Disclosure	26 years	6.35M
CVE-2025-55182	React Server Components RCE	New	5.93M
CVE-2020-2034	Palo Alto PAN-OS Injection	5 years	3.75M
CVE-2024-3721	TBK DVR RCE	1 year	314K
CVE-2016-20016	MVPower DVR RCE	9 years	279K
CVE-2014-6271	Shellshock	11 years	212K
CVE-2017-9841	PHPUnit RCE	8 years	140K
CVE-2019-9082	ThinkPHP RCE	6 years	138K
CVE-2023-1389	TP-Link RCE	2 years	136K
CVE-2021-3626	Hikvision RCE	4 years	119K