5.8 Million Knocks:
A Three-Year Honeypot Record of SSH and Telnet Attacks

Rufo J. De Francisco
October 31, 2026

Introduction

Internet-facing servers are probed constantly, and almost none of that traffic ever reaches a human analyst. To see what is actually out there, this report draws on a Cowrie honeypot [1]: a deliberately exposed SSH and Telnet service that looks like a real Linux machine, accepts a share of the logins it receives and records everything that happens next, including every command typed and every file transferred. Because no legitimate user has any reason to connect, everything the sensor records is unsolicited, which makes it a clean window onto automated attacker behavior.

The report summarizes more than three years of that record, from July 26, 2023 to October 31, 2026. It asks practical questions. How much hostile traffic does a single exposed host attract? Who sends it, and from which networks? Which credentials, tools and malware families do the attackers rely on? Is the picture changing? And, most usefully for defenders, which controls would actually make a difference?

The sections move from the broad picture to the specifics. A summary of findings comes first, followed by the traffic profile, the attack funnel, credentials, malware, download infrastructure, protocols, source IPs, countries and networks, and trends over time. The report closes with a prioritized threat assessment, recommendations, a candid account of the data's limitations, and suggestions for further research. Readers short on time can read the summary and the recommendations alone; the sections in between provide the evidence. Most of the underlying data is publicly available on the author's website [2].

Executive summary

Between July 26, 2023 and October 31, 2026, a single honeypot sensor, running continuously for more than three years, logged 5,810,156 unsolicited/unwanted connections from 232,086 IP addresses in 215 countries. Nearly four in five of them (78.5%) were login attacks rather than simple scans, and the rest formed a steady background of scans that continued day after day. The traffic is automated, driven by guessed credentials and spread across a very large population of sources. That shape matters for defense: it is far better countered by shrinking exposure and detecting behavior than by blocking individual IP addresses or file hashes.

- **Attacks dominate.** The honeypot saw 4,562,740 attacks against 1,247,416 scans, which works out to roughly 3,825 attacks and 1,046 scans a day over the 1,193 days covered.
- **A steady background of scans runs beneath the attacks.** Scans hold a floor of about 820 a day (median 818) and move largely independently of attack volume. At least 147,597 IPs (63.6% of all sources) only ever scanned. Not all of this is necessarily hostile: a scan connects without trying to log in, which is also what internet-wide research projects, device search engines [13], [16] such as Shodan [20] and Censys [21], exposure-notification services such as Shadowserver [19] and

administrators' own port checks produce. The data cannot separate benign from malicious scanning, so this baseline should be read as a mixture (see Limitations).

- **Logins often succeed.** The honeypot accepted 39.5% of attacks (1,803,764), and 73.1% of those (1,319,254) went on to run commands or malware. This reflects the honeypot's login policy, so it should not be read as a measure of real-world password strength.
- **One credential family stands out.** Each attack presents exactly one username and password pair, and the pairs built on `345gs5662d34` or `3245gs5662d34` account for 693,195 attacks, exactly 15.2% of the total.
- **Malware is overwhelmingly unique.** Of the 1,838,290 distinct files observed, 1,837,681 were transferred in and only 609 were created inside sessions. Hash blocklists will therefore miss most of what arrives.
- **A few created files recur constantly.** One of them, SHA-256 `a846…f8f2`, appears in 347,073 attacks, or 26.3% of all active attacks, and VirusTotal rates it malicious.
- **SSH is the main vector; Telnet is mostly reconnaissance.** About 85% of connections used SSH, a ratio of 5.6 to 1 over Telnet. SSH carried 96.8% of all attacks (4,414,867), about 30 times as many as Telnet, and about nine in ten SSH connections were attacks rather than scans. Telnet is the reverse: it drew 59.1% of all scans (737,267, about 45% more than SSH), but its scans outnumber its attacks about 5 to 1, so it converts to attacks far less often.
- **No single IP drives the totals, with one possible exception.** The ten busiest IPs account for almost 13% of scans and 8% of attacks. The busiest, IP01 (AS51396, Pfcloud UG), made 66,742 scans (5.4%) and 84,295 attacks (1.8%). Those totals equal about 70% and 80% of the excess scan and attack volume in the two 2025 bursts (May 31 to June 1 and June 20 to 23), which is consistent with one source driving much of both, although the dashboard has no per-IP time series to confirm it.
- **Networks matter more than IPs.** DigitalOcean (AS14061) alone accounts for almost 14% of all attack volume, and two Chinese carriers (AS4134 and AS4837) together account for 23% of scan volume. The ten leading networks hold 50% of scans and 41% of attacks.
- **Attacks are elevated since late May 2026, but not unprecedented.** Weekly attacks stepped up in the week of May 25 and have stayed above the median in 20 of the 22 full weeks since, averaging about 35% above the whole-period average. Volume peaked at 6,226 attacks a day in August and eased to 4,783 in October. That rise followed a spring trough, when March and April 2026 were the quietest months since August 2023 (2,551 and 1,925 a day). The current level matches late 2023 and early 2024 (Q3 2026 averaged 5,458 a day against 5,510 in Q4 2023). Attacks show no long-term upward or downward trend. Scans held a steady base of about 28,000 a month, apart from two surges in June 2025 and May 2026. Total unsolicited traffic keeps a floor of about 100,000 connections a month.

Beyond the headline numbers, six insights deserve particular attention:

- **The honeypot is both a target and a weapon.** Attackers do not only try to plant malware on the sensor. The download sites ranked 1, 2 and 5, which make up 76.1% of the top ten's requests, were distributed denial-of-service (DDoS) targets, which means compromised hosts are also being recruited to flood third parties. A weakly protected server is therefore a potential launchpad for attacks on others as well as a victim in its own right.
- **One door, many tenants.** The ten most common files point to several distinct criminal ecosystems working through the same weakly protected SSH service: Outlaw and Shellbot cryptojacking, Mirai

and Gafgyt-style IoT botnets, and DDoS tooling. Access appears to be a commodity that different operators exploit in different ways.

- **Automation is the signature.** The recurring credential strings, the handful of constantly repeated scripts and the steady daily baseline all point to dominance of scripted tooling rather than hands-on intruders. That predictability is an advantage for defenders, because automated attackers follow the same steps every time: they test the login, fingerprint the host and then stage a payload.
- **Reputation scores lag behind behavior.** IP addresses IP01 and IP11 are each flagged by only one of 92 VirusTotal vendors, yet they carry 2,160 and 4,183 AbuseIPDB reports respectively, and eight of the 19 busiest addresses have not been reported for more than six months. A clean or quiet reputation today says little about how an address behaved when it attacked in the past.
- **Most of the traffic is invisible to blocklists.** The top ten download sites cover only about 1.4% of transfer events, and 1,837,681 of the 1,838,290 distinct files were transferred in, nearly all of them unique. Lists of URLs and file hashes will always trail the attacker, which is why the recommendations favor reducing exposure and detecting behavior.
- **One sensor, three years: time is the strength.** A single honeypot sees only what reaches its own address, but continuous operation for more than three years gives it what a snapshot cannot: a stable baseline (the steady scan floor and a median attack day of 2,883), a way to tell persistent campaigns from passing events, and a record of change. Only a long record puts the 2023 start-up step, the multi-week troughs, the 2025 bursts and the Q3 2026 surge in context, and it lets recurring signatures such as the 345gs5662d34 credential family (15.2% of attacks) be read as long-running campaigns rather than one-off events. Breadth is the trade-off (see Limitations); depth in time is what makes the patterns reliable.

Taken together, the findings describe a threat that is relentless and largely automated. The most effective response is also the least glamorous: turn off Telnet, require SSH keys, remove default credentials, limit who can reach SSH, and watch for the behaviors that follow a login, such as outbound downloads, files created by shell redirection and logins that disconnect immediately. The threat assessment and recommendations near the end of this report set out these steps in priority order.

Scope, data and method

This report analyzes findings from one Cowrie SSH and Telnet honeypot between July 26, 2023 and October 31, 2026, a span of about 1,193 days. Geolocation and autonomous system (ASN) data come from MaxMind GeoLite2 [3], and malware verdicts come from VirusTotal. IP address and domain reputation, shown under “Independent check,” comes from VirusTotal [4], AbuseIPDB [5], and URLhaus [6] lookups. Network owner names were obtained from public web sources, including the Hurricane Electric border gateway protocol (BGP) toolkit [8].

Several terms are used consistently throughout the report. A **scan** is a connection that does not try to log in. An **attack** presents a username and password. A **successful attack** is one the honeypot accepts and gives a shell. An **active attack** goes on to run commands or malware, whereas a **passive attack** disconnects right after login.

Traffic profile

The honeypot absorbed about 4,870 unsolicited connections a day, and roughly four in five were login attacks. Scans formed a steady background of about 1,046 a day.

Table 1. High-level indicators of unsolicited traffic

Measure	Value	Share or rate
Total unsolicited connections	5,810,156	About 4,870 per day
Scans	1,247,416	21.5% of connections, about 1,046 per day
Attacks	4,562,740	78.5% of connections, about 3,825 per day
Unique IP addresses (all traffic)	232,086	25.0 connections per IP on average
Unique attacking IPs	84,489	36.4% of all IPs, 54.0 attacks per IP on average
Unique ASNs	11,200	About 21 IPs per ASN
Countries of origin	215	Near-global coverage

Two points stand out. First, at least 147,597 IPs (63.6% of all IPs) never attacked: subtracting the attacking IPs from the total leaves that many sources that only scanned. Second, the scale is wide but shallow. With 232,086 sources averaging about 25 connections each, the honeypot was touched by a very large, mostly low-volume population rather than by a handful of persistent actors.

Attack funnel

Fewer than three in ten attacks (28.9%) went on to run commands or malware, but that still amounts to about 1,106 active attacks a day.

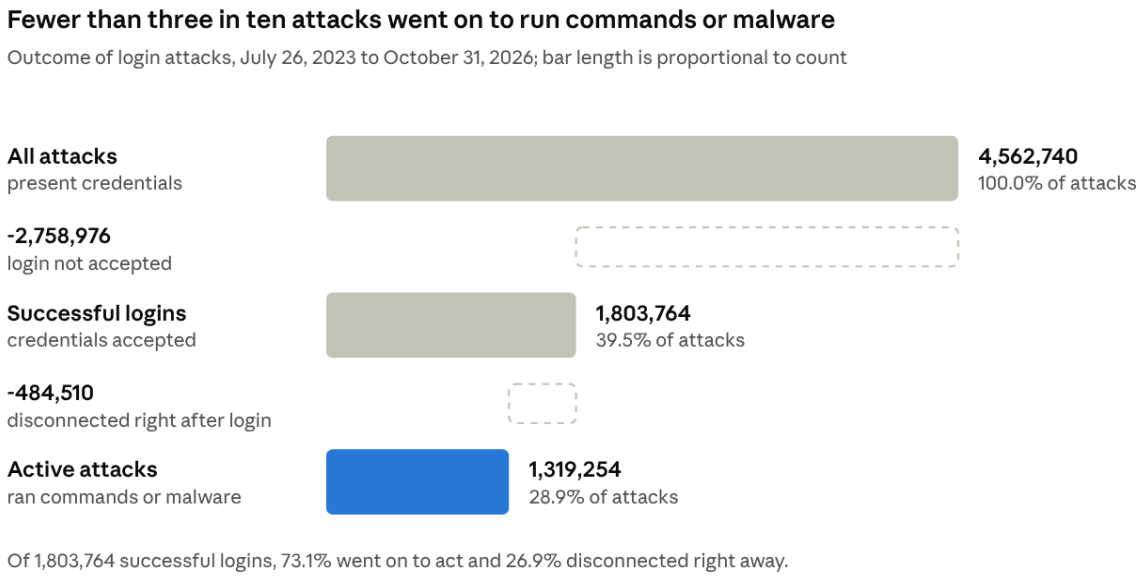


Figure 1. Attack funnel

As shown in Figure 1, attackers lose volume at two steps. The honeypot rejected 60.5% of attack sessions outright, and among the sessions that did get a shell, passive and active logins split roughly 1 to 2.7.

- **Logging in is routine for a source.** Of 84,489 attacking IPs, 67,704 (80.1%) logged in at least once, averaging about 26.6 successful attacks each.
- **Active outnumbers passive.** The 1,319,254 active attacks are 2.7 times the 484,510 passive ones, which means most logins are put to use immediately.
- **Passive logins are a distinct pattern.** A login that disconnects at once fits credential validation or access checking, common in reconnaissance.

The success rate depends on the honeypot's acceptance policy, so these proportions describe the honeypot construction choices rather than real systems.

Credential analysis

Two odd strings, not classic weak passwords, lead the credential list, and together they account for 14.8% of all attacks. Each attack presents exactly one username and password pair, so every count below is a number of attacks: the ten pairs cover 870,690 attacks, exactly 19.1% of the 4,562,740 total.

Top-10 credential pairs by number of attacks they appear in

Table 2. Login credentials

Rank	Username	Password	Attacks	Share of all attacks	Share of top 10
1	`345gs5662d34`	`345gs5662d34`	350,630	7.68%	40.3%
2	`root`	`3245gs5662d34`	324,965	7.12%	37.3%
3	`root`	`admin`	54,169	1.19%	6.2%
4	`root`	`(blank)`	52,271	1.15%	6.0%
5	`admin`	`admin`	24,699	0.54%	2.8%
6	`admin`	`3245gs5662d34`	17,600	0.39%	2.0%
7	`root`	`root`	15,428	0.34%	1.8%
8	`root`	`admin123`	12,782	0.28%	1.5%
9	`root`	`qwerty`	9,917	0.22%	1.1%
10	`root`	`123456`	8,229	0.18%	0.9%

Four observations follow from the table:

1. **One string family dominates.** Ranks 1, 2 and 6 all use `345gs5662d34` or the near-identical `3245gs5662d34`, totaling 693,195 attacks: 15.2% of all attacks and 79.6% of the top ten. These credentials are part of the Outlaw Shellbot SSH botnet's routine. This was confirmed after observing a strong correlation between the IP addresses that used a brute-force approach with the `345gs5662d34` credential family and the IP addresses with indicators of compromise associated with

the Shellbot attacks (e.g., presence of the `mdrfckr` string). It is not entirely established why these particular credentials entered brute-force dictionaries. A few theories exist:

a. **Polycom CX600 default password.** This is the most repeated explanation, and it traces back to one sentence in a 2023 cybersecurity study [9] developed by researchers from Aalborg University and the Technical University of Denmark. The authors say that, upon researching the password, they found it is possibly the default credential for a Polycom CX600 IP telephone for both the username and password. They give no citation or vendor documentation. Later blog posts and write-ups repeat it, some more firmly than the original.
b. **"Is this host already ours?" check.** One cyber-research team speculates that the unique strings were either chosen or synthetically constructed to implement a check for whether the server is already compromised. That would require the attack to change the SSH username and password after a successful penetration to `3245gs5662d34`.
c. **Poorly coded bots.** One analysis [10] notes that bots sometimes enter passwords into the wrong fields, which could explain why the same string, perhaps intended as a password, shows up as both username and password. This clumsy behavior may explain the slight variation—an extra character—observed in the field (`32`**`2`**`45gs5662d34` vs `345gs5662d34`). That is a comment on the quirk, not a confirmed cause.

In short: `345gs5662d34` is a real default credential from Polycom hardware that entered the common credential pools years ago and remains heavily used by the Outlaw cryptojacking botnet and generic SSH sprayers. That combination produces the volume honeypot operators continue to observe to this day.

2. **Root is the main target.** The `root` username appears in seven of the top-ten pairs, in at least 477,761 attacks (10.5% of all attacks). `admin` appears in two pairs, in at least 42,299 attacks (0.9%). Because these totals add up only the top-ten pairs that use each name, they are lower bounds. The same applies to the password `3245gs5662d34` (at least 342,565 attacks, 7.5%), `admin` (at least 78,868, 1.7%) and the blank password (at least 52,271, 1.1%).
3. **Classic weak pairs are a small share.** Setting the string family aside, the other seven pairs cover 177,495 attacks, 3.9% of all attacks. They are the blank password, `admin`, `root`, `admin123`, `qwerty`, and `123456`.
4. **The tail is wide.** Exactly 3,692,050 attacks (80.9%) used pairs outside the top ten. That suggests large and varied dictionaries, something confirmed by the analysis of the data: in all, 427,345 distinct credential pairs were involved.

One caution applies. The honeypot's acceptance policy decides which pairs succeed, so this list shows what attackers try, not what real systems would accept.

The string family also makes a useful detection signal: a login attempt with either string on a production system is a strong sign of automated tooling.

Malware analysis

Malware activity is huge in volume yet almost entirely unique. Some 5,575,924 file interactions produced 1,838,290 distinct files, while a handful of scripts recur in hundreds of thousands of attacks.

Table 3. File interactions during attacks

Measure	Value	Note
File interactions	5,575,924	About 4.2 per active attack
File creations (command redirection)	447,720	8.0% of interactions
File transfers (uploads and downloads)	5,128,204	92.0% of interactions
Unique files	1,838,290	About 1.4 per active attack
Unique files created	609	Each recreated about 735 times on average
Unique files transferred	1,837,681	Each seen about 2.8 times on average
Distinct download hosts	766	

The per-attack figures assume that file activity occurs in active attacks (1,319,254), which matches the honeypot's definitions.

Top 10 files by number of active attacks they appear in

Table 4. Hashes of files involved in attacks

Rank	SHA-256 (abbreviated)	Attacks	Share of attacks	Type	Activity	VirusTotal verdict
1	`a8460f44…f8f2`	347,073	26.3%	Created	Attack prep	Malicious
2	`01ba4719…546b`	48,091	3.7%	Created	Attack prep	Undetected*
3	`1b20a210…2057`	37,189	2.8%	Created	Attack recon	Undetected
4	`207e9bcc…ed2a`	8,867	0.7%	Created	Attack recon	Undetected
5	`6468ac9f…34bf`	5,918	0.5%	Download	DDoS	Undetected*
6	`d9c5bd8d…acf4`	5,104	0.4%	Download	Attack dropper	Malicious
7	`d46555af…a98e`	1,219	0.1%	Upload	Attack (crypto)	Malicious
8	`16174ef4…a080`	1,130	<0.1%	Download	Attack dropper	Malicious
9	`80c15d81…762d`	972	<0.1%	Created	Attack dropper	Undetected
10	`32f2fa94…c864`	743	<0.1%	Download	Attack download	Undetected*

Undetected identifies those samples that were not reported by either the official VirusTotal vendors or members of the wider VirusTotal community, which includes individual researchers and honeypot operators. **Undetected*** identifies samples that were not reported by the VirusTotal vendors but were identified as suspect or malicious by the VirusTotal community.

List 1. Full SHA-256 hashes, for blocklists and lookups

```
a8460f446be540410004b1a8db4083773fa46f7fe76fa84219c93daa1669f8f2
01ba4719c80b6fe911b091a7c05124b64eeece964e09c058ef8f9805daca546b
```

1b20a210fe96e5a8abc347dfb91d7befecb4b5f9b7ed40d856410fac15952057
207e9bcc282cfb15a4e5b5b3b9e5adc0622ae4af0346366a5ee917b85651ed2a
6468ac9f9bca964f3910fc967b80781c1c8634300e36f95ae49056d91a2734bf
d9c5bd8dc94485e3d286637b6b97d54a4225cf23a7f2f59a4c6c92e47d16acf4
d46555af1173d22f07c37ef9c1e0e74fd68db022f2b6fb3ab5388d2c5bc6a98e
16174ef4d82f50ebc62b573c05f71b9db34660b456b9558ef9be06201bfda080
80c15d8108ea5828bdbdf8a5eff145358ca69532c748ca8ef44b1b82d962762d
32f2fa940d4b4fe19aca1e53a24e5aac29c57b7c5ee78588325b87f1b649c864

With the help of a number of cyber analysis tools and services [4], [6], [7], and Artificial Intelligence (AI) large language models (LLMs) [17], [18], a few meaningful observations and insights can be deduced from the data:

1. **Created files are few and heavily reused.** Five redirection-type files (ranks 1 to 4 and 9) appear in 442,192 attacks. Because each use involves at least one creation event, they account for at least 98.8% of the 447,720 creations, while the other 604 created files share at most 5,528 events.
2. **The rank 1 and 2 samples point to cryptojacking preparation, and both lead by a wide margin.** In particular, they are associated with a long-lived Linux botnet built around Shellbot, a Perl-based Internet Relay Chat (IRC) bot, operated by the hacking group Outlaw, believed to be Romanian. The campaign is a cryptojacking operation (cryptomining done on someone else's machine) that has remained active since at least mid-2018. The group originally spread the bot by exploiting a common command injection vulnerability on IoT devices and Linux servers, and later campaigns shifted toward SSH brute-force attacks against servers with weak credentials. Once a machine is infected, it takes commands over an IRC channel, including port scans, DDoS attacks and file downloads. The main goal is cryptojacking: the toolkit combines the IRC bot for command and control, Monero mining for revenue, and the `haiduc` scan and brute-force tool, and newer analysis says it also deploys modified XMRig miners and spreads in a worm-like way. XMRig is a popular open-source cryptocurrency mining software, primarily used to mine Monero (XMR) and other coins that use the RandomX algorithm. The rank 1 sample appears in 26.3% of active attacks and 19.2% of successful attacks, making it about seven times as common as the next sample. Together, the rank 1 and rank 2 samples appear in 30.0% of all active attacks, about ten times as often as the next sample.
   a. The rank 1 sample is the hash of a malicious SSH `authorized_keys` file used as a persistence backdoor, commonly labeled `trojan.shell/malkey`. Attackers brute-force or exploit weak or default SSH credentials on internet-facing Linux servers, IoT devices, containers and poorly secured hosts. Once they have a shell, they append (or replace) this specific public key in the victim's `authorized_keys` file. The key carries the comment string `mdrfckr`, and with the matching private key the attackers gain permanent passwordless SSH access, even if passwords are changed later. From that foothold they deploy coinminers (often XMRig-based) and IRC and Shellbot components, remove competing miners, and sometimes add further backdoors or scanners to spread. Related names and tooling include Outlaw, Shellbot and various XMRig droppers that plant this exact key.

b. The rank 2 sample is the hash of a single newline character. When the honeypot receives a request for a file, such as those Shellbot makes, it often answers with an HTTP 200 OK status whose body is exactly one newline.

3. **The rank 3 and 4 samples are indicators of cryptojacking reconnaissance.** Cybercriminals often run automated reconnaissance to decide whether a host is worth further infection and which architecture-specific payload to fetch. They do this with a multi-stage fingerprinting script that tries several ways to get `uname` output; determines the target architecture (x86_64, aarch64, armv7l and so on) with similar fallbacks; collects uptime, CPU count, CPU model, GPU information and recent logins; runs a shell-behavior test that creates a temporary script that echoes a string, makes it executable, runs it and deletes it; and prints everything with labeled prefixes. The rank 3 and rank 4 samples are the hashes of that temporary script or of its output, and they show the same labeled hardware-fingerprinting style documented in SSH bots that triage systems for mining suitability. Together they appear in about 3.5% of all active attacks.
4. **The rank 5 sample is the hash of a "The service is unavailable" response.** It was seen in a DDoS attack against the website of the National Library of Israel.
5. **The rank 6 sample is the hash of a Mirai dropper.** It is associated with the file `ssh.sh`, a well-known Mirai-related shell script that downloads architecture-specific payloads after a successful SSH login. Mirai is an IoT botnet that first appeared in 2016. It infects internet-connected devices such as cameras, DVRs and home routers, then uses them to launch large DDoS attacks. Infected devices scan the internet for others with open telnet (and later SSH) ports and try a built-in list of common default credentials, such as admin/admin, so devices whose owners never changed the factory password are easy prey. Bots connect to a command-and-control server and wait for attack orders. Mirai lives in memory and has no persistence, so a reboot clears the infection, but an exposed device with default credentials is usually reinfected within minutes, since its IP address was already captured when it was first infected.
6. **The rank 7 sample is the hash of a cryptojacking script.** It is associated with `clean.sh`, a short cleanup or uninstall-style script used by cryptojacking botnets. It targets the C3Pool miner (a legitimate-looking XMRig setup that is frequently abused), typically to remove competing miners or to clean up after an installation.
7. **The rank 8 sample is the hash of another Mirai dropper.** In this case the file name is `gay.sh`, which fits the pattern of Mirai-family botnets that deliberately use provocative or meme-style filenames.
8. **The rank 9 sample is the hash of a non-Mirai dropper.** Specifically, it is a classic Gafgyt/Bashlite-style IoT botnet dropper, tried over SSH or Telnet against routers and embedded devices. Gafgyt is a long-running botnet, first seen in 2014, that, like Mirai, targets IoT devices in order to build a botnet for DDoS attacks. The observed tiny-dropper, multi-fallback-download and BusyBox pattern fits that lineage.
9. **The rank 10 sample is the hash of the response to a file download request.** Although this is the hash that uniquely identifies the default Nginx "403 Forbidden" HTML error page, automated analysis tools, such as the Hatching Triage Sandbox, list this hash when malware scripts attempt to retrieve secondary payloads or reach a Command & Control (C2) server.
10. **Transfers are a long tail.** The five most-used transferred files appear in only 14,114 attacks, set against 5.13 million transfer events and 1.84 million unique transferred files. Payloads are rarely repeated.

11. **Verdict coverage is weak.** By attack count, 77.7% of top-ten activity from 4 samples involves files rated malicious by VirusTotal, 12.0% activity from 3 samples as suspect (i.e., reported by the community, but not by the primary security vendors), and 10.3% activity from 3 samples are unreported.
12. **Unreported (or even a harmless rate) does not mean benign.** The rank 3, 4, and 9 samples appear as not reported in VirusTotal, yet they were created in 47,028 attack sessions. The extended VirusTotal community, but not the main vendors, reported the rank 2, 5, and 10 samples, which combined appeared in 54,752 attacks. These samples represent files that are not malicious per se (e.g., new line, HTTP status codes or messages), but are artifacts and indicators of attack activity. As such, they should be reviewed, not dismissed.

Because transferred files are almost never repeated, attackers likely vary their payloads per delivery. Hash-based blocking will therefore catch only a sliver of transfers, and behavioral and network indicators are more durable.

Download activity

The honeypot recorded a significant number of download attempts. In general, download requests fall into one of two broad categories:

- **Attempts to obtain malware.** The aim is to download malware onto the target system, in this case the honeypot, for later execution.
- **DDoS attacks.** Here the download host is itself the target of the attack, rather than infrastructure used to carry out cyberattacks.

The top-ten hosts together cover only a small slice of all download requests. The honeypot captured 766 distinct download hosts in total.

Table 5. Download hosts and number of download requests

Rank	Download host	Host type	Download requests	Share of top 10
1	`postv.media`	DDoS target	43,914	63.2%
2	`www.nli.org.il`	DDoS target	6,354	9.1%
3	`cloudsecure.live`	Malware host	6,080	8.8%
4	`213.209.143.44`	Malware host	3,257	4.7%
5	`170.75.166.213`	DDoS target	2,640	3.8%
6	`107.150.0.18`	Malware host	2,072	3.0%
7	`kraken16.buzz`	Malware host	1,888	2.7%
8	`golden-orda.biz`	Malware host	1,376	2.0%
9	`5.255.121.213`	Malware host	1,215	1.7%
10	`93.123.85.79`	Malware host	663	1.0%

A few findings stand out:

1. **Concentration at the top.** Rank 1 is 6.9 times larger than rank 2, so taking down or blocking that one host would remove the single biggest identified source.
2. **DDoS attacks dominate.** Hosts ranked 1, 2 and 5, which together make up 76.1% of all download requests from the top-ten hosts, were all recipients of DDoS attacks. Rank 1, `postv.media`, is the website of a pro-government Georgian television and digital media platform based in Tbilisi. Rank 2 is the website of the National Library of Israel (`nli.org.il`). Rank 5 is a web server hosted in Canada by LunaNode, a budget-focused cloud hosting provider that has carved out a niche in the privacy and cryptocurrency space. That server was the target of an application-layer (Layer 7) DDoS and fuzzing campaign designed to bypass web application firewalls (WAFs) and cache layers.
    a. **What a Layer 7 flood looks like.** A sample capture of the commands aimed at the rank 5 server, `170.75.166.213`, contains 588 `curl` requests to its root URL over plain HTTP, only 182 of them distinct. Each distinct request is deliberately different from the others: nine HTTP methods (GET, POST, PUT, PATCH, DELETE, HEAD, OPTIONS, TRACE and CONNECT), 55 different `User-Agent` strings spanning desktop browsers and legacy feature-phone browsers, a unique random cookie, varied `Cache-Control` and `Accept` headers, and a spoofed distinct `Referer` in 421 of the 588 requests. Several of those referers point to pages named for public "f5" DDoS tools, which suggests a widely shared flood script. Varying every request this way is a standard technique for defeating caching, rate limiting and simple WAF signatures, and it fits a campaign whose goal is to overwhelm the server rather than to steal anything. In this exchange, the honeypot is the launchpad and `170.75.166.213` is the victim.
3. **Mixed infrastructure.** Five of the top ten hosts are bare IP addresses and the other five are named domains, which fits the known use and targeting of both purpose-registered domains and raw servers by cybercriminals.
4. **A large visibility gap.** The top ten sites account for 69,459 download attempts, about 1.4% of the 5,128,204 transfer events (uploads and downloads). The rest came from the other 756 download sources and from uploads, where the source URL is not captured. With 1.84 million unique transferred files but only 766 URLs, most unique files probably arrived by upload or from servers that vary their content. URL and domain blocking therefore helps but cannot be the main control, since it covers an unknown minority of transfers. The gap also presents an opportunity for further research.

Independent check of the sites

Reputation lookups of the seven hosts that were not attack targets were run on October 31, 2026. The checks corroborate the malicious nature of all seven, with varying degrees of certainty.

Table 6. Reputation of suspect hosts

Host	Flagged as malicious by	Notes
`cloudsecure.live`	Meta [12]	Phishing domain controlled by CyberRoot, an Indian hack-for-hire group.
`213.209.143.44`	VirusTotal, URLhaus, AbuseIPDB	Hosts malicious payloads such as `netgear.sh`, `ssh.sh`, and Mirai-variant ELF binaries (`UnHAnaAW.arm5`) designed for SSH scanning and IoT exploitation.

Host	Flagged as malicious by	Notes
`107.150.0.18`	VirusTotal, URLhaus, AbuseIPDB	Serves both Mirai and Gafgyt payloads from the United States.
`kraken16.buzz`	VirusTotal (suspicious)	Phishing domain routed through AWS and hosted in the Netherlands.
`golden-orda.biz`	VirusTotal	Russian-language domain associated with phishing and investment scams. Registrar in China, registrant in the Netherlands, reverse-proxied through Cloudflare.
`5.255.121.213`	VirusTotal, URLhaus, AbuseIPDB	Serves Mirai, ELF binaries, and `ua-wget` payloads from the Netherlands.
`93.123.85.79`	VirusTotal, URLhaus, AbuseIPDB	Hosts Linux malware and C2 infrastructure linked to MooBot (DDoS botnet) and payloads like `GuruITDDoS3.sh`; based in Bulgaria.

Protocols

SSH carried 84.8% of all connections (4,925,016), a ratio of 5.6 to 1 over Telnet (885,140, or 15.2%). Table 7 below breaks each protocol's traffic down into scans and attacks.

Telnet's share is still large for a protocol with no encryption, which shows that attackers keep probing for it.

The practical reading is that both protocols need attention, and that Telnet, the weaker of the two in terms of security, should not be exposed at all.

Table 7. Unsolicited traffic by protocol

Protocol	Activity	Unsolicited connections count	Share
SSH	Scans	510,149	40.9%
SSH	Attacks	4,414,867	96.8%
Telnet	Scans	737,267	59.1%
Telnet	Attacks	147,873	3.2%

The following insights can be gleaned from the data:

- **SSH: overwhelmingly attack-driven.** Attacks outnumber scans nearly 9 to 1 (4.41M vs 510K), so about 90% of SSH traffic went beyond probing. This fits what honeypots typically see: credential-stuffing and brute-force botnets hammering port 22 around the clock. A scan is cheap and often a one-off, while an attack involves many repeated login attempts per source, so the attack count is inflated by volume per attacker.
- **Telnet: mostly reconnaissance.** The ratio flips: scans outnumber attacks about 5 to 1 (737K vs 148K). Telnet scanning is dominated by IoT botnets (Mirai and its descendants) that sweep the internet for open port 23 and only escalate to credential attempts against hosts that respond in a way they like.

A honeypot may not look like the device types they target, so many probes stop at the scan stage. Telnet is also a legacy protocol, so there are fewer real targets and the sweeps are more speculative.

- **Cross-protocol comparison.** SSH traffic is about 85% of all activity recorded, while traffic received over Telnet is 15% of the total. SSH attacks exceed Telnet attacks by roughly 30x, but Telnet scans are about 45% higher than SSH scans. In other words, Telnet gets plenty of attention, but it converts to attacks far less often. Attackers appear to prioritize SSH once they find a listener, since it usually fronts higher-value servers (cloud virtual machines, Linux boxes) that are useful for cryptomining, proxying, and DDoS attacks.

Source concentration: the busiest IPs

The ten busiest IPs generated only 7.4% of all connections (431,540 of 5,810,156). This means that the problem is spread across a very large population rather than concentrated in a few sources.

Table 8. Origin (IP addresses) of all unsolicited traffic

IP address	ASN	Country	All traffic	Share of traffic	Profile
IP01	AS51396	Netherlands	151,037	2.60%	scan-and-attack
IP02	AS197540	Germany	33,764	0.58%	attack-only
IP03	AS265268	Brazil	33,154	0.57%	attack-only
IP04	AS4808	China	32,608	0.56%	nearly attack-only
IP05	AS8151	Mexico	31,802	0.55%	attack-only
IP06	AS49505	Russia	31,769	0.55%	attack-only
IP07	AS197071	Germany	30,748	0.53%	scan-only
IP08	AS149722	Indonesia	29,325	0.50%	attack-only
IP09	AS22612	United States	28,736	0.49%	attack-only
IP10	AS4134	China	28,597	0.49%	nearly attack-only

Table 9. Origin (IP addresses) of scans

IP address	ASN	Country	Scans	Share of scans	Profile
IP01	AS51396	Netherlands	66,742	5.35%	scan-and-attack
IP07	AS197071	Germany	30,748	2.46%	scan-only
IP12	AS9009	United States	12,365	0.99%	scan-only
IP13	AS215174	Germany	8,511	0.68%	scan-only
IP14	AS32475	Netherlands	8,081	0.65%	scan-only
IP15	AS137688	China	7,611	0.61%	scan-only
IP16	AS200593	Russia	6,510	0.52%	scan-only
IP17	AS133296	India	6,317	0.51%	scan-only

IP address	ASN	Country	Scans	Share of scans	Profile
IP18	AS58224	Iran	5,647	0.45%	scan-only
IP19	AS400341	United States	5,284	0.42%	scan-only

Table 10. Origin (IP addresses) of attacks

IP address	ASN	Country	Attacks	Share of attacks	Profile
IP01	AS51396	Netherlands	84,295	1.85%	scan-and-attack
IP02	AS197540	Germany	33,763	0.74%	attack-only
IP03	AS265268	Brazil	33,153	0.73%	attack-only
IP04	AS4808	China	32,506	0.71%	nearly attack-only
IP05	AS8151	Mexico	31,801	0.70%	attack-only
IP06	AS49505	Russia	31,768	0.70%	attack-only
IP08	AS149722	Indonesia	29,324	0.64%	attack-only
IP09	AS22612	United States	28,735	0.63%	attack-only
IP10	AS4134	China	28,585	0.63%	nearly attack-only
IP11	AS22552	United States	26,599	0.58%	nearly attack-only

List 2. Clear IPs, for blocklists and lookups

IP01	176.65.151.53
IP02	159.195.53.121
IP03	168.90.66.80
IP04	106.75.86.153
IP05	189.168.152.115
IP06	45.131.41.43
IP07	109.230.239.178
IP08	103.187.91.2
IP09	159.198.46.132
IP10	121.12.76.28
IP12	37.120.214.30
IP13	45.135.232.24
IP14	107.6.151.170
IP15	59.47.5.35
IP16	91.215.85.88
IP17	43.241.37.250
IP18	2.183.82.226

IP19	165.140.216.140

What the data show:

1. **IP01 is an outlier.** It made 151,037 connections (2.6% of all traffic), 4.5 times as many as the runner-up. Its 66,742 scans are 5.4% of all scans and its 84,295 attacks are 1.8% of all attacks.
2. **Most heavy hitters specialize.** Eight of the top ten IPs are attack-only or nearly so, with few scans each. IP07 is the reverse: all 30,748 of its connections are scans. Only IP01 exhibits a true mixed operation profile, as it is both a prolific scanner and attacker.
3. **Attackers cluster at 26,000 to 34,000 connections.** After IP01, the leading attackers fall in a narrow band of 26,599 to 33,763 attacks.
4. **Scanning is more concentrated than attacking.** The top ten scanners made 12.7% of all scans (157,816 of 1,247,416), compared with 7.9% of attacks (360,529 of 4,562,740) for the top ten attackers.
5. **The long tail dominates.** Outside the top ten attackers, the other 84,479 attacking IPs averaged about 50 attacks each, close to the overall average of 54.
6. **Geography of the heaviest sources is diverse.** The heaviest IPs sit in the Netherlands, Germany, Brazil, China, Mexico, Russia, Indonesia, India, Iran and the United States, with China and Germany each appearing twice among the top ten.

Beyond IP01 and IP07, the top scanners are considerably smaller: IP12 (AS9009, United States) made 12,365 scans, followed by IP13 (AS215174, Germany) at 8,511, IP14 (AS32475, Netherlands) at 8,081, IP15 (AS137688, China) at 7,611, IP16 (AS200593, Russia) at 6,510, IP17 (AS133296, India) at 6,317, IP18 (AS58224, Iran) at 5,647, and IP19 (AS400341, United States) at 5,284. For comparison, IP11 (AS22552, United States), the tenth-ranked attacker, made 26,599 attacks.

Similarly, beyond IP01, the top attackers drop significantly in traffic: IP02 (AS197540, Germany) made 33,763 attacks, followed by IP03 (AS265268, Brazil) at 33,153, IP04 (AS4808, China) at 32,506, IP05 (AS8151, Mexico) at 31,801, IP06 (AS49505, Russia) at 31,768, IP08 (AS149722, Indonesia) at 29,324, IP09 (AS22612, United States) at 28,735, IP10 (AS4134, China) at 28,585, and IP11 (AS22552, United States) at 26,599.

The implication is that per-IP blocking removes little. Banning all ten of the busiest attackers would cut attacks by under 8%, because the rest of the attacking population is large and flat.

Independent check of the addresses

Reputation lookups run on October 31, 2026 corroborate the honeypot evidence for all 19 addresses, with varying strength. On VirusTotal, 13 addresses are flagged as malicious by 1 to 9 of 92 security vendors (median 5; IP01, IP09 and IP17 have only 1), and IP11 has 10 or more detected files communicating with it. The other five (IP02, IP07, IP10, IP15 and IP18) were not among the VirusTotal results supplied, but all five are reported on AbuseIPDB, with 8 to 1,146 reports, the most recent 10 months to 2 years old.

Table 11. Reputation of suspect IP addresses

IP address	VirusTotal verdict	AbuseIPDB verdict
IP01	1/92 vendors flag as malicious	2,160 reports; latest 1 year ago
IP02	Not supplied	12 reports; latest 10 months ago
IP03	2/92 vendors flag as malicious	306 reports; latest 6 months ago
IP04	6/92 vendors flag as malicious	859 reports; latest today
IP05	6/92 vendors flag as malicious	10 reports; latest 2 months ago
IP06	3/92 vendors flag as malicious	206 reports; latest yesterday
IP07	Not supplied	107 reports; latest 1 year ago
IP08	5/92 vendors flag as malicious	1,052 reports; latest today
IP09	1/92 vendors flag as malicious	38 reports; latest 2 weeks ago
IP10	Not supplied	105 reports; latest 10 months ago
IP11	1/92 vendors flag as malicious	4,183 reports; latest 4 months ago
IP12	2/92 vendors flag as malicious	626 reports; latest 4 months ago
IP13	7/92 vendors flag as malicious	4,458 reports; latest 1 year ago
IP14	5/92 vendors flag as malicious	3,257 reports; latest 2 weeks ago
IP15	Not supplied	1,146 reports; latest 2 years ago
IP16	9/92 vendors flag as malicious	2,386 reports; latest today
IP17	1/92 vendors flag as malicious	734 reports; latest yesterday
IP18	Not supplied	8 reports; latest 1 year ago
IP19	5/92 vendors flag as malicious	1,570 reports; latest 2 months ago

Based on the AbuseIPDB information, 5 of the suspect IP addresses were still active when this report was written, 8 are stale (last reported more than six months ago), and the remaining 6 were last seen between 2 weeks and 4 months ago.

A word of caution on the information above: Reputation lookups show an address's standing at the time of consultation, but the dataset spans more than three years (July 2023 to October 2026). Addresses are cleaned up, reassigned or abandoned, so a low or absent current score does not show how an address behaved when it attacked, and an old report does not show that it is still abusive. One or two flags out of 92 vendors should be treated as weak corroboration; the honeypot's own logs remain the primary evidence. Before blocking, check each address's current score and last-report date.

Countries and networks of origin

Attacks come mainly from cloud and hosting networks, while scans come mainly from national carriers. China, the United States and the Netherlands lead the country rankings for both scans and attacks. Chinanet, China Unicom and Pfcloud lead the network rankings for scans, while DigitalOcean, Tencent and TechTies lead the rankings for attacks. Each share below is measured against the total shown at the head of its column.

Table 12. Unsolicited traffic by country

Rank	All traffic (5,810,156)	Scans (1,247,416)	Attacks (4,562,740)
1	China 17.5%	China 35.6%	United States 13.2%
2	United States 13.0%	United States 12.4%	China 12.5%
3	Netherlands 10.3%	Netherlands 8.7%	Netherlands 10.7%
4	Singapore 5.1%	India 5.7%	Singapore 5.9%
5	India 4.7%	Germany 5.0%	India 4.5%
6	Germany 4.4%	Pakistan 3.7%	Germany 4.2%
7	Bulgaria 3.4%	Iran 3.3%	Bulgaria 4.1%
8	Russia 3.0%	Russia 2.2%	Russia 3.1%
9	South Korea 2.6%	Singapore 2.1%	Indonesia 3.1%
10	Hong Kong 2.6%	South Korea 1.9%	Hong Kong 2.9%

Table 13. Unsolicited traffic by autonomous system

Rank	All traffic (5,810,156)	Scans (1,247,416)	Attacks (4,562,740)
1	AS14061 DigitalOcean 12.0%	AS4134 Chinanet 15.2%	AS14061 DigitalOcean 13.8%
2	AS4134 Chinanet 5.3%	AS4837 China Unicom 7.7%	AS132203 Tencent 6.1%
3	AS132203 Tencent 4.8%	AS51396 Pfcloud 5.9%	AS197170 TechTies 6.0%
4	AS197170 TechTies 4.7%	AS14061 DigitalOcean 5.8%	AS45090 Tencent 2.8%
5	AS51396 Pfcloud 2.9%	AS9829 BSNL 3.4%	AS4134 Chinanet 2.6%
6	AS45090 Tencent 2.6%	AS58224 Iran Telecom 3.1%	AS51396 Pfcloud 2.1%
7	AS4837 China Unicom 2.0%	AS396982 Google Cloud 2.7%	AS198364 BanatSync 2.1%
8	No ASN mapped 1.9%	AS197071 Active 1 2.5%	No ASN mapped 1.9%
9	AS198364 BanatSync 1.6%	AS45090 Tencent 2.0%	AS16276 OVH 1.7%
10	AS4766 KT Corporation 1.5%	AS9541 Cyber Internet Services 1.8%	AS47890 Unmanaged Ltd 1.6%

What the rankings and their underlying data show:

1. **Chinese scanning runs through national carriers.** China accounts for 35.6% of all scan traffic, nearly three times the United States (12.4%). Its two backbone carriers, AS4134 and AS4837, hold 22.9% of all detected scans.
2. **Attacks run through cloud and hosting providers.** DigitalOcean (AS14061) accounts for 13.8% of global attack traffic, more than twice the next network. DigitalOcean operates in several of the countries on the list, which plausibly explains part of the traffic attributed to Singapore, the

Netherlands, India and Germany. The two Tencent ASNs and the Chinanet ASN combined account for 11.5%.

3. **The networks are concentrated.** IP01 made 66,742 scans and 84,295 attacks from Pfcloud (AS51396), amounting to 5.4% of all scans and 1.8% of all attacks. The ten leading scan networks carry 50% of all scans (624,261 of 1,247,416), and the ten leading attack networks carry 40.7% of all attacks (1,856,899 of 4,562,740). By contrast, the ten busiest IPs account for only 12.7% of scans and 7.9% of attacks, so network-level controls reach much further than IP-level ones.
4. **The Netherlands is third everywhere.** It holds 8.7% of all scans and 10.8% of all attacks. One Dutch-labelled address, IP01, made 5.4% of all scans and 1.8% of all attacks.
5. **City-scope jurisdictions made the top list.** Singapore is a particularly relevant case: although the country operates 0.75% of all IPv4 addresses, the honeypot detected that 2.1% of all scan traffic and 5.9% of all attack traffic originated there. Hong Kong shows a similar pattern, with 2.9% of all attacks coming from IPs operating there.
6. **Scan and attack networks are largely separate.** AS132203, AS197170, AS198364, OVH (AS16276) and Unmanaged Ltd (AS47890) appear only on the attack side, whereas BSNL (AS9829), Iran Telecom (AS58224), Google Cloud (AS396982), Active 1 GmbH (AS197071) and Cyber Internet Services (AS9541) appear only on the scan side. The Google Cloud share may include benign research scanners, which the data cannot separate.
7. **Traffic with no attributed ASN was detected.** At the time it was consulted, the MaxMind GeoLite2 database failed to attribute an ASN to a number of IP addresses responsible for a significant volume of attacks: 86,481 of 4,562,740, or 1.9%.

Provider-level reporting and rate limiting would reach furthest, since a handful of networks carry a large share of the volume. The cost of network-wide blocks is collateral damage to legitimate users on the same networks, which cannot be quantified from this data.

Geolocation map

The geolocation maps cover every inhabited continent. 215 jurisdictions, almost all, appear in the data record as origin of unsolicited traffic. A notable exception is tightly-controlled North Korea. The densest clusters are in the eastern and central United States, Western and Central Europe, southeastern Brazil, India, eastern China, Japan and Korea, and Southeast Asia. Marker size reflects connection volume, so a large marker shows a location from which many connections originated.

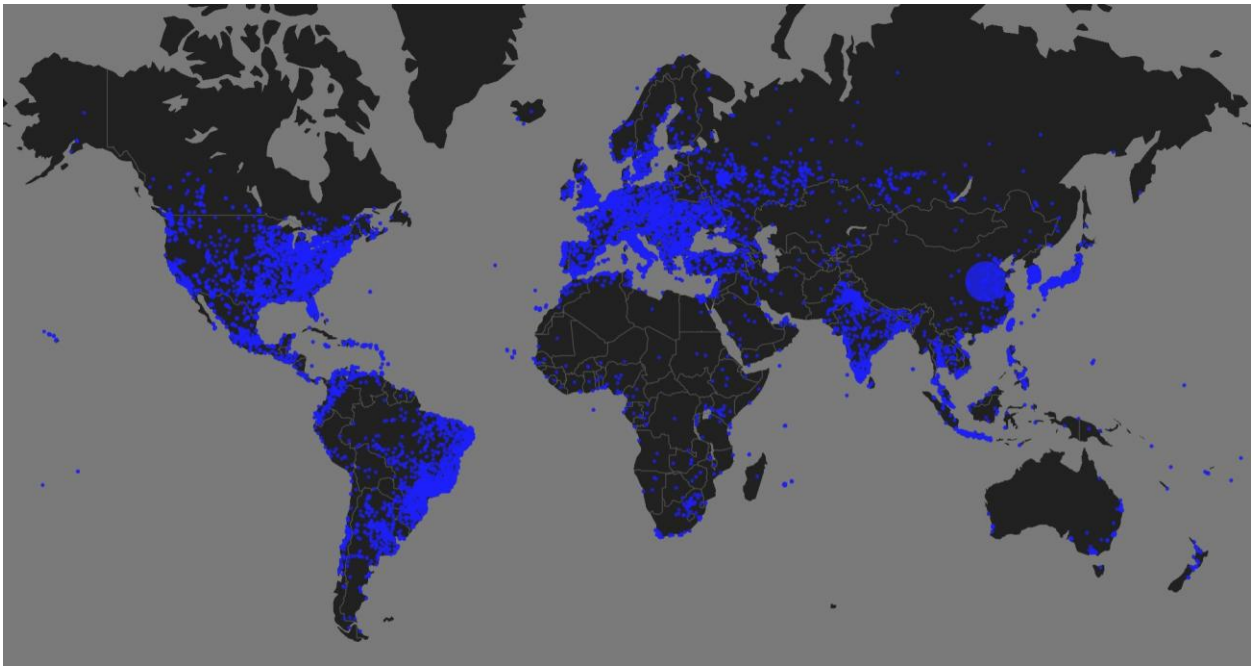


Figure 2. Geolocation of scan origins

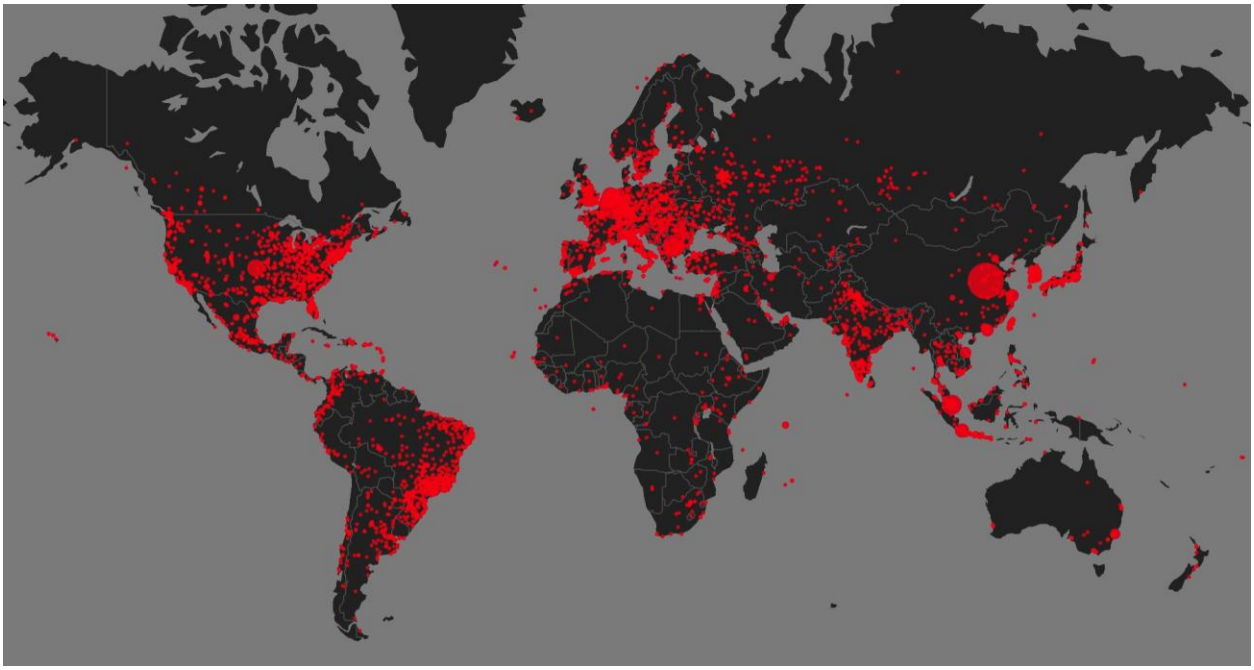


Figure 3. Geolocation of attack origins

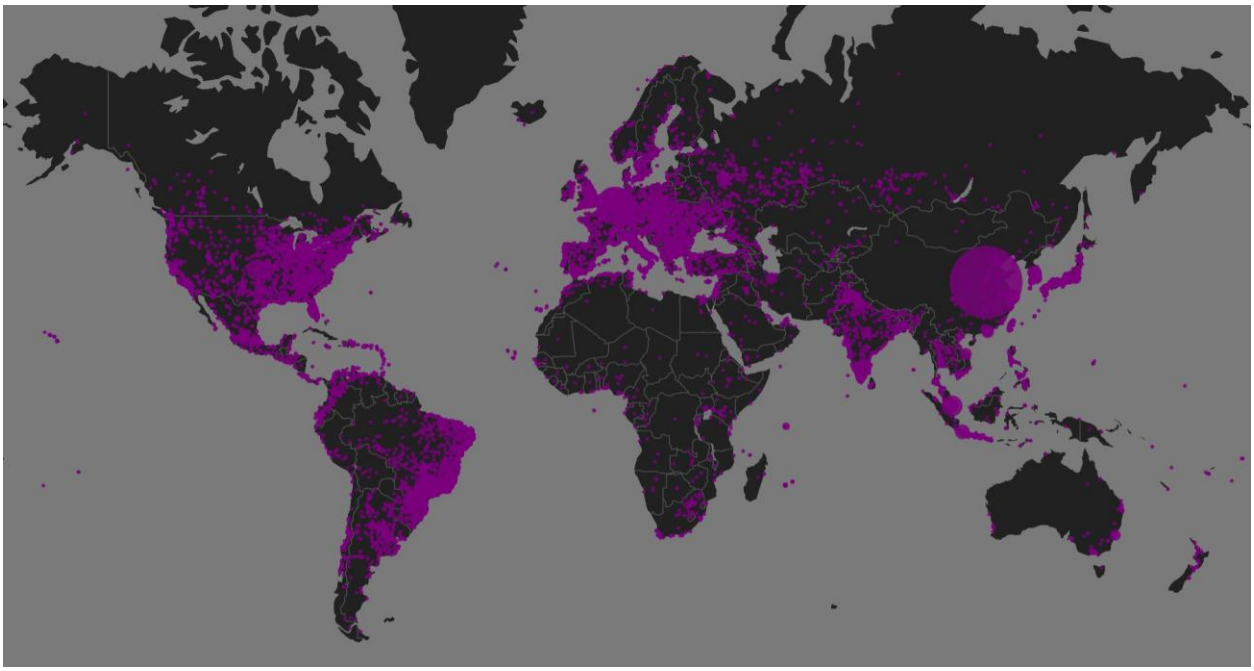


Figure 4. Geolocation of all unsolicited traffic origins

The three maps differ in ways that match the rankings above:

- **Scans (Figure 2, blue markers)** are spread thinly across a very wide area, including Russia and central Asia, with one dominant marker in eastern China.
- **Attacks (Figure 3, red markers)** are more concentrated. Large markers sit in eastern China, western Europe near the Netherlands and Germany, the Balkans (consistent with Bulgaria's place in the attack-country ranking), the central United States, Indonesia, and Korea and Japan. Russia and central Asia are sparse in red. Brazilian attacks cluster in the southeast, while Brazilian scans cover the whole country.
- **All traffic (Figure 4, purple markers)** combines both, with its largest marker in eastern China.

The large United States and China markers sit near the middle of each country, which fits GeoLite2 placing addresses it can locate only at country level on a single default point, which sometimes is the country centroid. Such markers therefore aggregate many addresses and do not show one busy host. Geolocation is approximate, and VPN or proxy use can hide the true origin, so country attribution should be treated as indicative.

Trends over time

Unsolicited traffic runs on a steady base with spikes on top of it. In a quiet month, the honeypot receives about 25,000 scans and 75,000 to 80,000 attacks (together, about 100,000 connections), and a typical month brings about 28,000 scans and 105,000 attacks, as shown in Figure 7. Scans are the steadier of the two, apart from two surges, in June 2025 and May 2026. Attacks swing more and have been elevated since late May 2026, at levels comparable to late 2023 and early 2024. There is no discernible long-term trend in scans, attacks or their total.

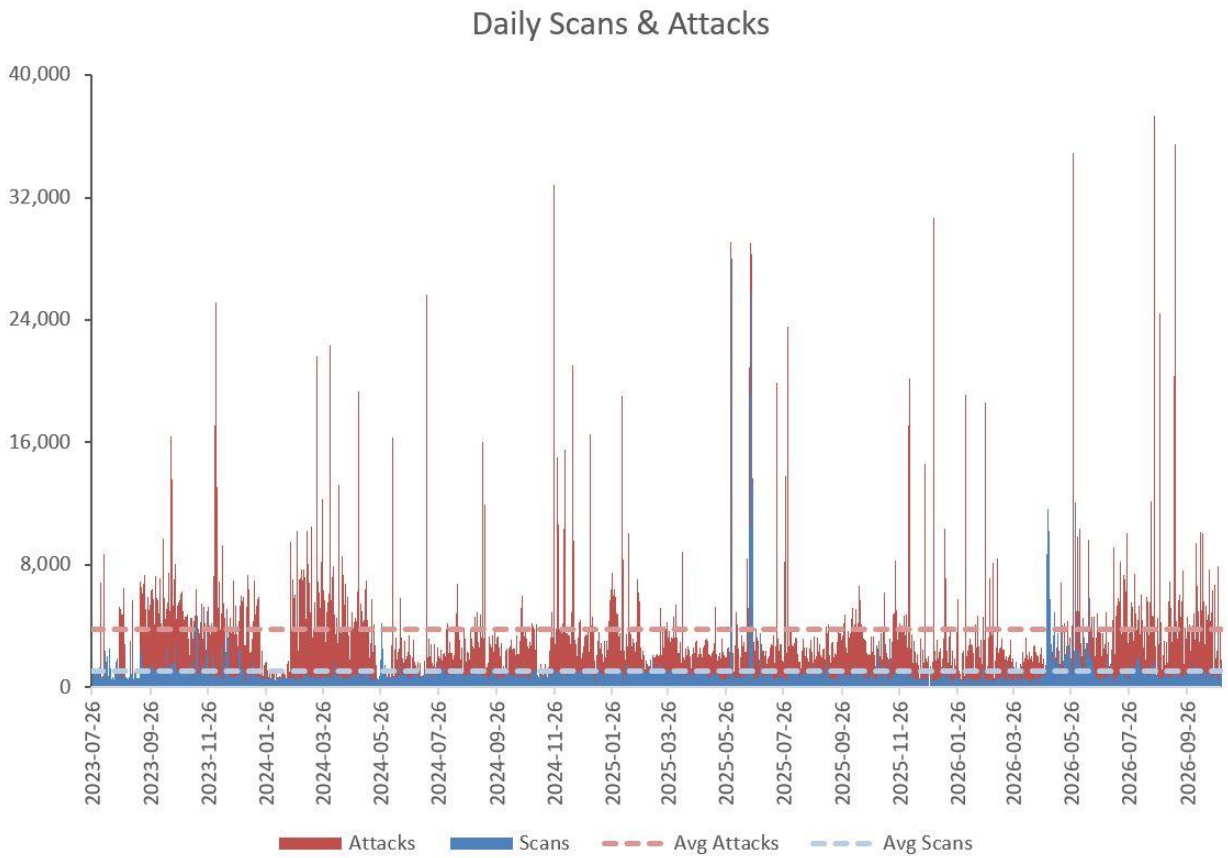


Figure 5. Daily trend of scans and attacks

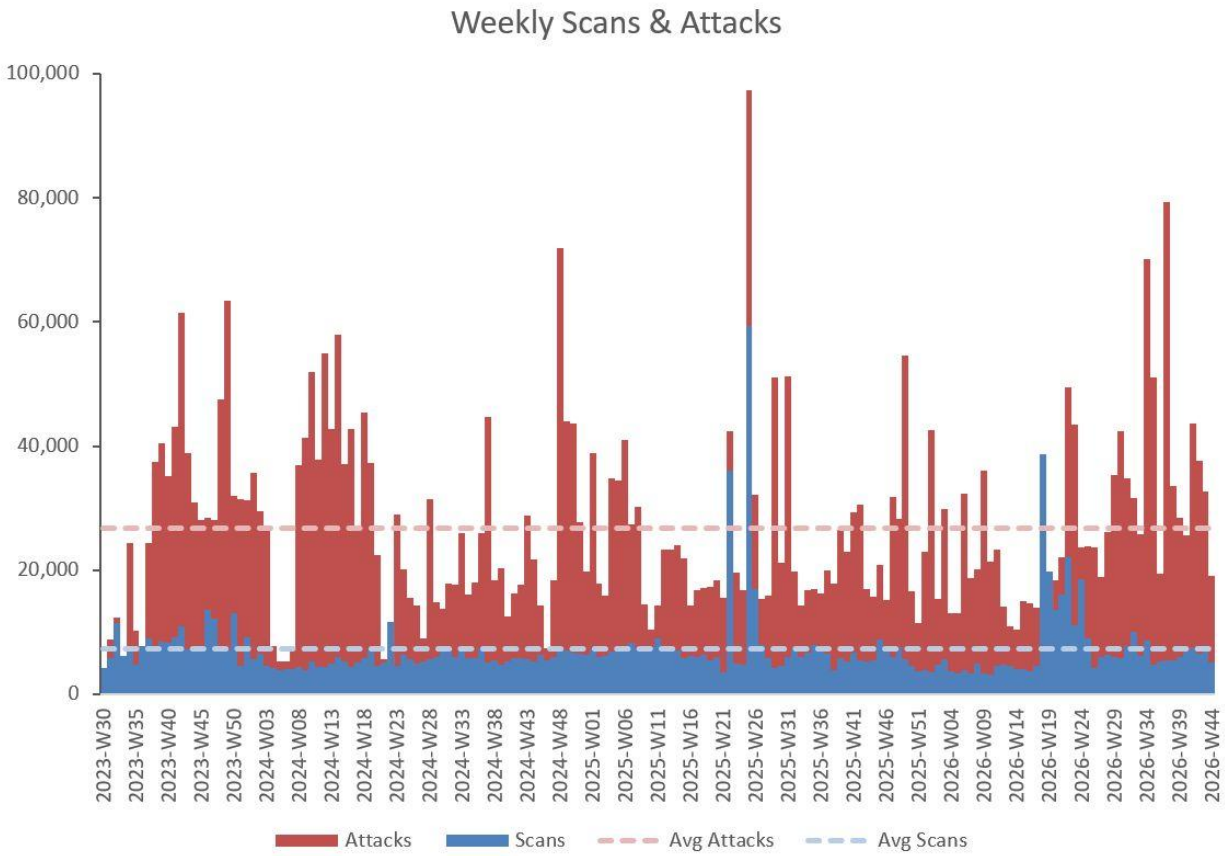


Figure 6. Weekly trend of scans and attacks

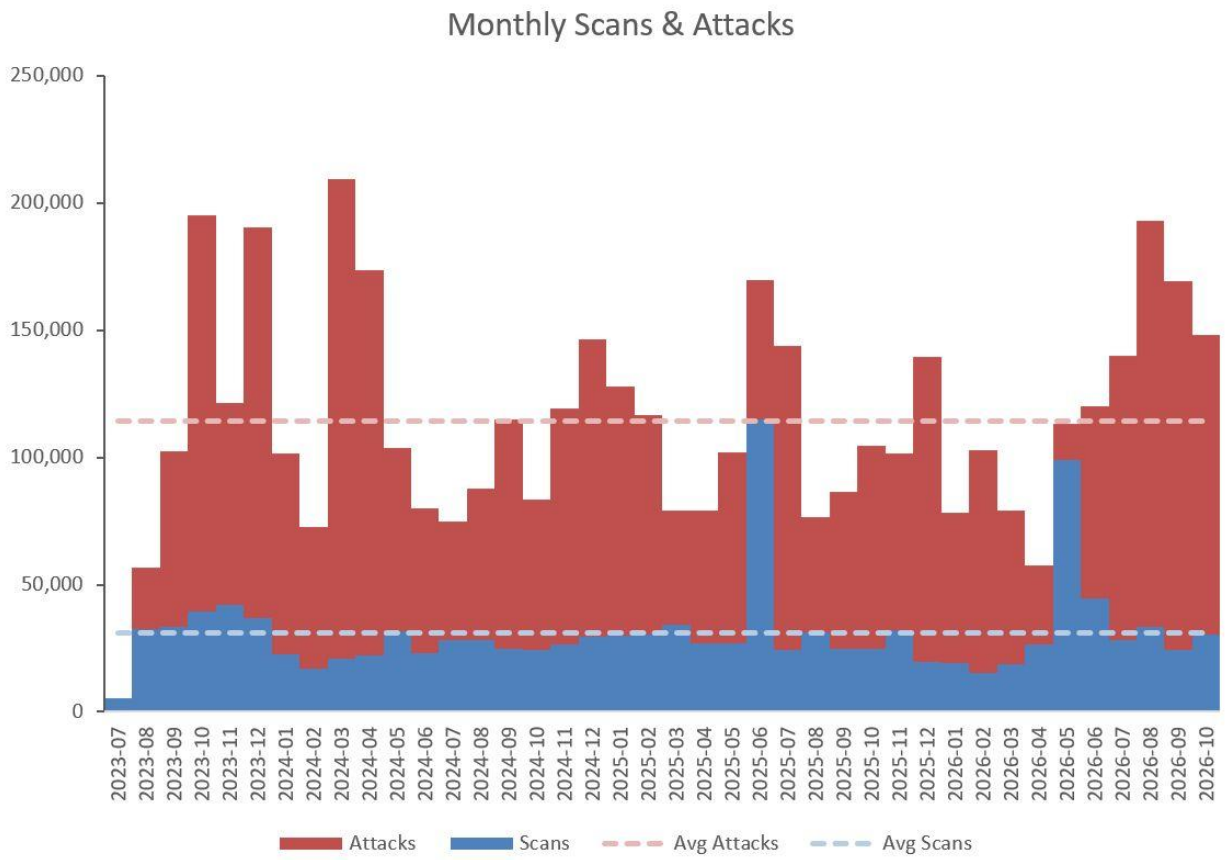


Figure 7. Monthly trend of scans and attacks

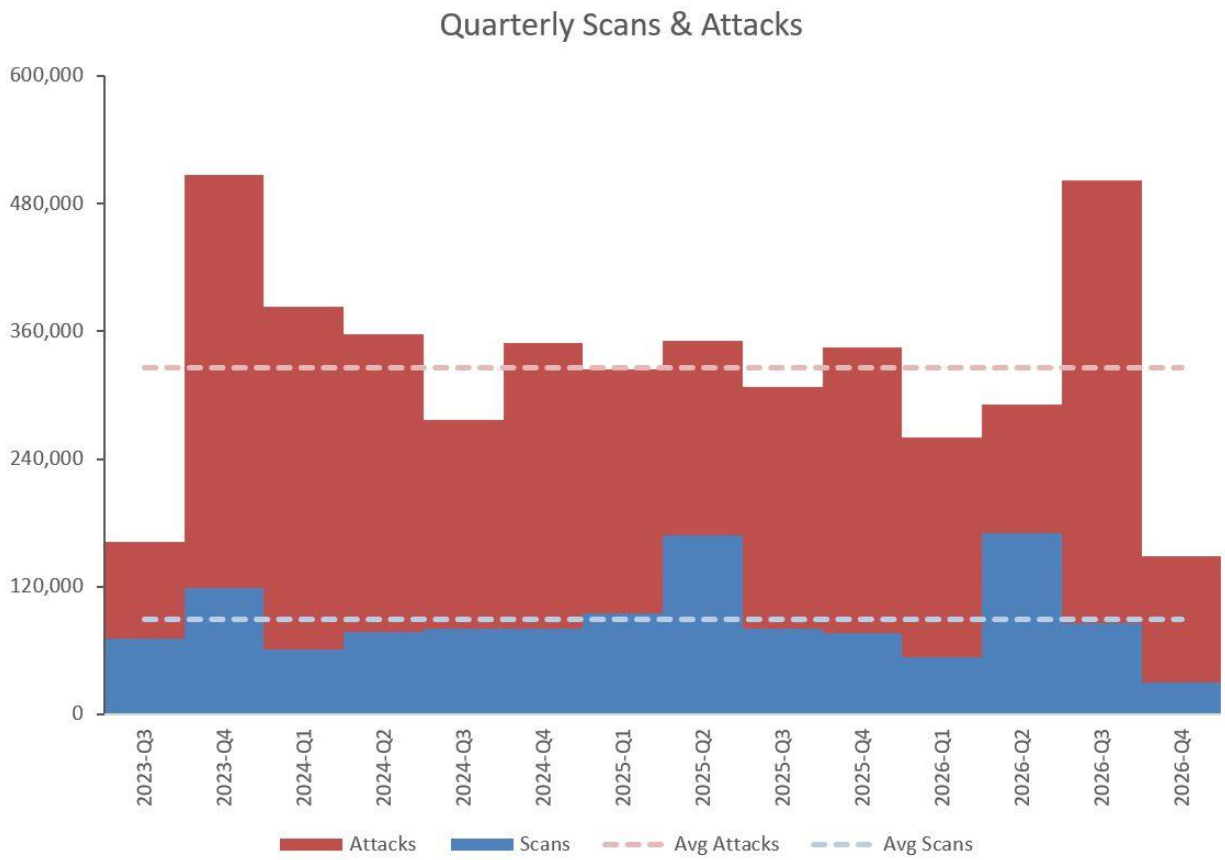


Figure 8. Quarterly trend of scans and attacks

Table 14. Timing of unsolicited traffic events

Date	Event	Size
Late May to Oct 2026	Elevated period in attacks, peaking in August	Since May 25, 36,407 attacks a week against 26,883 overall (+35%); August 6,226 a day, third-highest month
Sep 13, 2026	Second-largest attack day of the period	35,515 attacks
Aug 22, 2026	Largest attack day of the period	37,375 attacks
Apr 30 to Jun 14, 2026	Scan surge, with peaks on Apr 30 to May 3 and May 30; back to normal by late June	137,532 scans (2,990 a day); peak 12,098 scans on May 30
Mar to Apr 2026	Spring trough: lowest attack rate since August 2023	2,551 (Mar) and 1,925 (Apr) a day; weekly attacks 10,500 to 15,000 from Mar 16 to May 10
Jun 20 to 23, 2025	Second burst, four days	91,880 attacks and 68,537 scans; attacks peaked Jun 21 (28,991), scans Jun 22 (25,890)

Date	Event	Size
May 31 to Jun 1, 2025	First burst: attacks on May 31, scans the next day	29,058 attacks on May 31; 28,029 scans on Jun 1 (largest scan day)
Nov 25, 2024	Largest attack day before 2026	32,824 attacks
Nov 8 to 19, 2024	Shallow attack trough, 12 days	941 a day
May 21 to Jun 5, 2024	Deepest sustained attack trough, 16 days	539 a day
Jan 27 to Feb 16, 2024	Attack trough, 21 days; 7-day average above 4,000 again by Feb 23	784 a day
Sep 14, 2023	Step up in attacks, 50 days after the start	1,357 a day before, 5,516 a day after (4.1 times)
Aug 8 to Sep 13, 2023	Two early troughs (Aug 8 to 22 and Sep 7 to 13)	392 and 196 a day

Scans hold a floor of about 820 a day (median 818 a day), and 80% of days fall between roughly 530 and 1,370 scans, as shown in Figure 5. For completeness, weekly and quarterly trends are shown as Figure 6 and Figure 8, respectively.

What the trends show:

1. **Scans form a steady base with two surges.** Of the 39 full months, 37 fall between 15,000 and 45,000 scans (median 28,000; the quietest quarter of months had 24,300 or fewer). The exceptions are June 2025 (114,908) and May 2026 (99,029), about 4.1 and 3.5 times the median. The surges have different shapes. June 2025 is six days in two bursts (May 31 to Jun 1 and Jun 20 to 23, 100,776 scans). The 2026 surge is a six-week plateau (Apr 30 to Jun 14, 137,532 scans, 2,990 a day) that peaked on Apr 30 to May 3 and on May 30 and ended abruptly in late June. Eleven of the twelve largest scan days fall in these two episodes.
2. **Attacks vary far more, around a stable floor.** Monthly attacks range from 56,745 to 209,289. The quietest tenth of months had 76,000 or fewer, the median month 104,555 and the busiest tenth 177,000 or more. Month-to-month variation is about 1.5 times that of scans outside their two surges (coefficient of variation 0.35 against 0.24). The median day holds 2,883 attacks, the ten largest days carry only 6.8% of all attacks, and 32 of the 1,194 days topped 15,000.
3. **There is no long-term trend.** Monthly scans, attacks and their total show no significant drift in either direction (Kendall tau between -0.04 and 0.06, p-values above 0.58). Attacks averaged 3,733 a day in 2024, 3,637 in 2025 and 3,953 in January to October 2026, all within 9% of one another.
4. **Attacks are not clearly higher now than in the past.** Q3 2026 averaged 5,458 attacks a day, level with Q4 2023 (5,510) and above every quarter in between (2,889 to 4,212). August 2026 (6,226 a day) is the third-highest month, behind March 2024 (6,751) and October 2023 (6,299). The highest 28-day average (7,857, ending Sep 13, 2026) is only 9% above the earlier peak (7,201 in Q2 2024), and a third of that window's attacks came on two days. Two of the five largest weeks fall in 2026 (September 7 with 79,284 attacks and August 17 with 70,149). Against the same months of earlier years, the 2026 level is far higher (August to October: 5,549 a day against 3,106 in 2024 and 2,911 in

2025), and against the long-run average of 3,823 a day Q3 2026 is 43% above. The current level is high, but it is within the range the series has already shown.

5. **An elevated period began in late May 2026.** From March 16 to May 10, weekly attacks held between 10,500 and 15,000, the longest low stretch since the 2023 start-up (eight weeks; the lowest 8-week average after the start-up, 13,554), and April's 57,742 attacks was the lowest month since August 2023. The level stepped up in the week of May 25 (49,494 attacks). Since then, 20 of 22 full weeks have been at or above the median week (23,336 attacks), averaging 36,407 against 26,883 overall (+35%). Monthly attacks were 113,399 in May, 120,087 in June, 139,915 in July, 192,995 in August, 169,220 in September and 148,281 in October. Two qualifications apply. June weeks sat only just above the median (about 23,700), so the clearly elevated stretch begins in July (15 of 16 full weeks at or above the median, averaging 38,617). And earlier episodes lasted longer: 19 straight weeks at or above the median from Sep 11, 2023 to Jan 15, 2024 and 12 from Feb 19 to May 6, 2024, against blocks of 8 and 7 weeks in 2026, split by a single week of 19,519 attacks (Aug 31).
6. **Total traffic keeps a floor near 100,000 connections a month.** Monthly totals (scans plus attacks) never fell below 84,311 (April 2026) and were under 100,000 in only 5 of 39 months; the median is 136,120. Attacks are 78.5% of the total, so the total follows them, and it is only slightly steadier than attacks alone (coefficient of variation 0.32 against 0.35; scans alone 0.59). Scans and attacks barely move together (correlation 0.22 by month and 0.20 by day, falling to about zero once the six burst days of 2025 are removed), so their spikes arrive independently.
7. **Attack troughs do not follow scans.** Attacks stayed below 1,500 a day for a week or more in the 2023 start-up, from January 27 to February 16, 2024 (21 days, 784 a day), from May 21 to June 5, 2024 (16 days, 539 a day) and from March 26 to April 1, 2026 (7 days). Scans averaged 568 a day in the first of those 2024 troughs and 1,150 in the second. Recovery was fast after February 2024: the 7-day average was back above 4,000 by Feb 23, and March 2024 became the highest month of the period (6,751 a day). After June 2024 the rebound faded, and the 7-day average was back near 950 by July 11. A shallow trough from November 8 to 19, 2024 (941 a day) was followed by the largest attack day before 2026 (32,824 on Nov 25).
8. **The honeypot start-up took seven weeks.** The first 50 days (July 26 to September 13, 2023) averaged 1,357 attacks a day, with a median of only 333. On September 14, 2023 the level jumped, and from then to the end of the year attacks averaged 5,516 a day (median 4,977), about four times the start-up average. Averages that include these weeks understate the later baseline.
9. **One of the two 2025 bursts involves IP01.** The second 2025 burst covers four days (June 20 to 23) and holds 160,417 connections: 68,537 scans and 91,880 attacks. IP01 was the origin of 149,198 connections (93.0% of all): 65,989 scans (96.3%) and 83,269 attacks (90.6%).

Possible explanations (hypotheses):

The trend data show when things happened but not why. The ideas below are hypotheses that fit the pattern, not findings, and the daily counts alone cannot choose between them. Each ends with a test the honeypot or dashboard data could run.

Why scans are steady

1. **Fixed-rate background scanning.** Many actors sweep the whole IPv4 space on schedules that change slowly: research and exposure-mapping projects, device search engines such as Shodan [20] and Censys [21], notification services such as Shadowserver [19], and the propagation code of botnets.

Each address then receives a roughly constant stream of probes, which would explain the floor of about 820 scans a day, the narrow monthly band and the absence of a trend. It would also explain why scans are unrelated to attacks: at least 147,597 IPs (63.6% of all sources) only ever scanned. *Test: split scan sources by network owner and by whether they ever attack; fixed-schedule scanners should repeat weekly or monthly.*

2. **Time-boxed campaigns cause the surges.** The June 2025 bursts (one to four days each) and the six-week 2026 plateau, which decays and stops abruptly, look like finite jobs: a research or commercial scan project, or a botnet recruitment drive. The 2025 bursts may be one source (IP01). In 2026 the first peak (April 30 to May 3) fell on the days around CISA's May 1 listing of the Linux "Copy Fail" flaw (CVE-2026-31431) as exploited [15], and a separate honeypot described in a SANS Internet Storm Center guest diary logged its own spike on May 2 [11], which hints at an internet-wide event, not something specific to this sensor. The link is coincidence in time only: Copy Fail is a local privilege escalation, which would not by itself draw scans of the SSH port. *Test: per-IP daily series for April 30 to Jun 14, 2026; whether those scanners share a network or SSH client fingerprint.*

Why attacks vary without trending

3. **Many campaigns switching on and off around a stable pool of targets.** The ten busiest IPs account for only 7.9% of attacks, so volume is the sum of many bots, not a few heavy hitters. If the pool of guessable SSH hosts and the cost of attacking them change slowly (APNIC's honeynet write-up notes that old, cheap-to-scan threats never disappear and that botnets dating from 2017 and 2020 still keep tens of thousands of machines attacking every week [14]), attack volume should hover around a long-run level with episodes above and below it, as the series does, instead of trending. *Test: weekly count of distinct attacking IPs; spikes from more IPs, not more attempts per IP, point to botnet size.*
4. **Controller-driven tasking produces one-day spikes.** The SANS guest diary describes synchronized bursts from hosts in different countries that use the same SSH client fingerprint and throttled, uniform per-node rates, which its author reads as a controller assigning quotas to botnet nodes [11]. Days such as August 22 and September 13, 2026 (37,375 and 35,515 attacks) are the kind of event such tasking could create. *Test: SSH client fingerprints and credential lists on spike days against ordinary days.*
5. **Target lists, not organic growth, set this sensor's level.** Attackers choose targets from scan data and shared IP lists. The 4.1-fold step on September 14, 2023 looks like the sensor entering such lists, not gradual discovery, and the troughs (January to February 2024, May to June 2024, March to Apr 2026) could be the reverse: the address dropping out of lists, an operator pausing, or a major source network being cleaned up. The one outside honeypot found for the same months reported sustained high volume from March 9 to April 14, 2026, when this sensor was at its quietest [11], which fits sensor-specific effects, although a single home honeypot is thin evidence. *Test: compare with a second sensor on a different address; shared steps and troughs would be internet-wide, unshared ones would be sensor-specific.*

Why attacks are elevated since May 2026

6. **Part rebound, part new activity.** Some of the "elevated" impression is contrast with the spring trough: January to October 2026 as a whole averages 3,953 attacks a day, about 10% above the same months of 2024 (3,610) and 2025 (3,574). The rise from late May could be a lull ending (operators returning, infrastructure rebuilt, new credential lists) or a different actor adding volume. *Test: weekly*

share of the 345gs5662d34 credential family and of the top file hash before and after May 2026; a new family or new fingerprint points to a new actor, an unchanged mix to the old botnet returning.

7. **Source networks shift.** Attacks come mainly from cloud and hosting networks (DigitalOcean alone accounts for 13.8%). A provider cleaning up or a new one being abused would move volume without any change in attacker behavior. *Test: the dashboard's month-by-month network view for May to October 2026 against earlier months.*

Taken together, hypotheses 1 and 3 would produce the steady floor in the total: a stable scanner population and a stable population of persistent botnets, each roughly constant and largely independent, with discretionary campaigns (2, 4, 6) adding the spikes on top. What would help most is a per-IP daily series, network and credential-family counts by week, SSH client fingerprints by week and a second sensor on a different address.

Threat assessment and recommendations

The dominant threat is automated credential guessing followed by scripted post-login activity. It is broad-based, so the best controls reduce exposure and detect behavior. Attack volume has been elevated since late May 2026. Q3 2026 averaged 5,458 attacks a day, 43% above the long-run average of 3,823, and October was still about 25% above it at 4,783 a day. This is high but within the range the honeypot has already seen. It matches the level of late 2023 and early 2024, and the series shows no long-term trend, so it should be read as a spike on a stable base, not as a permanent increase. Earlier elevated periods lasted longer than this one has so far, so this period may well end without warning.

Beneath the attacks runs a constant, steady background of scans. A typical month brings about 28,000 scans, with a floor near 820 a day, interrupted only by surges in June 2025 and May 2026. It plausibly includes benign research and search-engine-style scanners as well as hostile reconnaissance, and the data cannot yet separate them.

By protocol, SSH carries 96.8% of attacks and is the principal target of credential guessing, while Telnet draws 59.1% of scans but only 3.2% of attacks, which makes it chiefly a reconnaissance target. The table below ranks the main threats by the strength of the evidence in the data.

Table 15. Threat assessment

Priority	Threat	Evidence from the data	Confidence
1	Credential guessing against SSH and Telnet	4,562,740 attacks; `root` in 54.9% of top-10 credential volume; `345gs5662d34` family in 693,195 attacks (15.2%); SSH carries 96.8% of attacks (4,414,867).	High on prevalence; medium on success rate (policy-dependent)
2	Scripted post-login activity and malware staging	1,319,254 active attacks (73.1% of successful logins); 5,575,924 file interactions.	High

Priority	Threat	Evidence from the data	Confidence
3	Varied payload delivery that evades hash lists	1,837,681 unique transferred files; top five transferred files in only 14,114 attacks.	Medium
4	Concentrated download infrastructure	766 URLs seen; top ten cover about 1.4% of file transfers.	Medium (large visibility gap)
5	Constant background scanning and reconnaissance	1,247,416 scans; steady floor of about 820 a day (median 818); at least 147,597 scan-only IPs; plausibly includes benign research and search-engine-style scanners	High on volume; benign versus malicious mix unknown
6	Burst attackers	The two 2025 bursts (May 31 to Jun 1 and Jun 20 to 23) exceeded normal levels by about 95,900 scans and 105,800 attacks; IP01's totals equal about 70% and 80% of those excesses, consistent with one source but unconfirmed; ten leading IPs at 26,599 to 84,295 attacks each	Medium
7	Exposed Telnet	737,267 scans (59.1% of all scans) and 147,873 attacks (3.2% of attacks); 15.2% of all connections	High

Recommendations

Reduce exposure first

1. Turn Telnet off on anything internet-facing.
2. Use SSH keys only, and disable root login and password authentication.
3. Remove default, blank and common passwords everywhere, and alert on any attempt using the `345gs5662d34` or `3245gs5662d34` strings, because they signal automated tooling.
4. Put any remaining SSH exposure behind a VPN, bastion host or IP allowlist. Changing the port is not a substitute.

Detect behavior, not just indicators

5. Alert on servers making outbound HTTP or HTTPS downloads, and block egress by default. Add the malware-hosting entries in the download-site table to blocklists. The DDoS targets in the same table are victims rather than attacker infrastructure, so verify each one (for example, www.nli.org.il) before blocking it.
6. Search for the ten file hashes, especially `a846…f8f2`, but do not rely on hash matching alone, since transferred files are almost never repeated.
7. Watch for files created by shell redirection right after login, and for logins that disconnect immediately. Passive logins were 26.9% of successful attacks and fit credential-validation behavior.

Be realistic about blocking

8. Per-IP bans give limited relief, since blocking the ten busiest attackers would cut attacks by under 8%, and the reputation evidence for these addresses ranges from strong to weak (see the per-IP table). Use rate limiting and reputation scoring instead, and report abuse to the hosting providers: DigitalOcean (AS14061) alone accounts for about 14% of all attack volume, and the ten leading networks hold 41% of attacks. Block whole networks only where the collateral cost is acceptable.

Improve the research dashboard

9. Conduct further analysis to improve data quality and the reliability of derived insights (see Further research).

Limitations and data quality notes

The main limits of the Cowrie-collected data are:

1. **Login success is policy-dependent.** The login acceptance rate depends on how the honeypot accepts credentials, an implementation choice, so it does not measure real-world password weakness.
2. **Scans are mixed.** Scans include benign research and search-engine traffic alongside malicious reconnaissance, and the data does not currently separate them.
3. **IP-level reputation is partial and time-dependent.** Although the AbuseIPDB database shows that all of the top-19 IP addresses were reported as suspect, VirusTotal vendors flag only 14 of them as malicious. The VirusTotal and AbuseIPDB figures come from manual, separate lookups on October 31, 2026. Because the honeypot dataset spans more than three years, today's reputation may not match an address's behavior at the time of the attack. IP addresses are routinely cleaned up and, in some cases, change hands.
4. **Malware verdicts are incomplete.** VirusTotal has no data for several of the top ten malware samples, and an unreported or "harmless" verdict says nothing about whether a file is used maliciously in context.
5. **There is uncertainty in geolocation.** IP geolocation is approximate, and VPNs or proxies can hide the real origin of unsolicited internet traffic. Large markers at country centers on the geolocation maps are most likely country-level defaults.
6. **One sensor.** A single honeypot sees only what targets its own IP address and configuration, so the results may not generalize to other networks and locations.

Further research

The unanswered questions and limitations noted throughout the report provide a blueprint for future follow-up work along the following lines of research:

- Segregate malicious probing/reconnaissance from harmless scans (i.e., crawlers and other "good bots").
- Cluster the 1.84 million files observed in unsolicited traffic with fuzzy hashing to find the families behind them.
- Identify proxied access to the honeypot to add context to the origin of unsolicited traffic.
- Resolve gaps in IP reputation and malware sample verdicts.
- Track changes over time in IP reputation scores, malware sample verdicts, IP assignments, and geolocation.

- Investigate clean-up efforts by network operators, for example by measuring how long abusive IP addresses remain visible to the honeypot.
- Rank network operators and jurisdictions based on meaningful normalizations (e.g., number of assigned IPv4 space).
- Automate AI-assisted analysis of malware samples and captured command activity.

Sources and References

[1] **Cowrie** medium to high interaction honeypot.
`https://cowrie.org`

[2] **Rufo de Francisco's Cybersecurity** page.
`https://defrancisco.us/cyber`

[3] **MaxMind GeoLite2**, a free, public IP geolocation dataset.
`https://www.maxmind.com/en/geolite-free-ip-geolocation-data`

[4] **VirusTotal**, analysis of suspicious files, URLs, domains, and IP addresses to detect malware, viruses, and other security breaches.
`https://www.virustotal.com`

[5] **AbuseIPDB**, repository of IP addresses that have been associated with malicious activity.
`https://www.abuseipdb.com`

[6] **URLhaus**, platform for sharing malicious URLs used for malware distribution.
`https://urlhaus.abuse.ch`

[7] **Joe Sandbox**, advanced malware and phishing analysis.
`https://www.joesecurity.org`

[8] **Hurricane Electric BGP Toolkit**, tools to research ASNs, BGP-announced prefixes, peers, routing relationships, and route propagation.
`https://bgp.he.net`

[9] Zetong Zhao, Shreyas Srinivasa, Emmanouil Vasilomanolakis. **SweetCam: an IP Camera Honeypot**. Aalborg University, Technical University of Denmark; 2023.
`https://vbn.aau.dk/ws/files/573748244/sweetcam_honeypot_paper_1_.pdf`

[10] Johannes Ullrich. **The Top 10 Not So Common SSH Usernames and Passwords**. SANS Internet Storm Center; 2024.
`https://isc.sans.edu/diary/31360`

[11] Adam Nason. **The Behavior of Coordinated SSH Brute Force Attacks over the last three months**. SANS Internet Storm Center; 2026.
`https://isc.sans.edu/diary.html?date=2026-06-17`

[12] David Agranovich, Michael Dvilyanski, Margarita Franklin. **Threat Report on the Surveillance-for-Hire Industry**. Meta Platforms; 2022.
`https://about.fb.com/wp-content/uploads/2022/12/Threat-Report-on-the-Surveillance-for-Hire-Industry.pdf`

[13] Zakir Durumeric, David Adrian, Ariana Mirian, Michael Bailey, J. Alex Halderman. **A Search Engine Backed by Internet-Wide Scanning**. ACM Conference on Computer and Communications Security (CCS); 2015.
https://faculty.cc.gatech.edu/~mbailey/publications/ccs15_censys.pdf

[14] Adli Wahid. **One week on the APNIC honeynet: 99,000 IP addresses and 173 million events**. APNIC; 2026.
https://blog.apnic.net/2026/10/26/one-week-on-the-apnic-honeynet-99000-ip-addresses-and-173-million-events

[15] Etiido Uko. **CISA flags actively exploited Copy Fail Linux kernel flaw (CVE-2026-31431)**. Tom's Hardware; 2026.
https://www.tomshardware.com/software/linux/cisa-flags-actively-exploited-copy-fail-linux-kernel-flaw-enabling-root-takeover-across-major-distros-unpatched-systems-may-remain-vulnerable-to-attack

[16] Camilo Gutiérrez Amaya. **Top 5 search engines for internet-connected devices and services**. WeLiveSecurity (by ESET); 2023.
https://www.welivesecurity.com/2023/05/18/5-search-engines-internet-connected-devices-services

[17] **Anthropic Claude** Sonnet 5.5 AI LLM.
https://claude.ai

[18] **SpaceXSI Grok** 4.7 AI LLM, Expert mode.
https://grok.com

[19] **Shadowserver Foundation**, daily Internet-wide scans and remediation reports sent to network owners.
https://www.shadowserver.org/what-we-do/network-reporting

[20] **Shodan**, search engine for Internet-connected devices.
https://www.shodan.io

[21] **Censys Search**, search engine to find exposed databases, IoT devices, vulnerable servers, and SSL/TLS certificates.
https://platform.censys.io