radware

Threat Report

2026 Global Threat Analysis Report

Analysis of the Global Network and Application Attack Trends of 2025

Table of Contents

About the Cover Picture

The modern cybersecurity landscape has become a digital Garden of Eden for threat actors, characterized by an abundance of low-hanging fruit ranging from unpatched vulnerabilities and indirect prompt injection flaws to more sophisticated business logic attacks. This paradise is further cultivated by the democratization of DDoS as a service and the rise of generative AI-based attack frameworks, which significantly lower the barrier to entry for novice hackers. In this environment, the “cyber apple” is depicted as a classic temptation reimagined in a neon, cyberpunk aesthetic. It represents the ultimate prize that has become the focus of intense competition. And because the landscape is so ripe with opportunity, a growing number of criminals are now aggressively fighting to claim the same lucrative fruit.

Foreword

The year 2025 has proven to be a watershed moment in the history of cybersecurity, a period we depict on the cover of this report as a digital Garden of Eden for threat actors. While the metaphor suggests a paradise of opportunity for our adversaries, it represents a fundamental and violent paradigm shift for many defenders.

As you will read in the following pages, the democratization of cyber offense is no longer a theoretical concern; it is our current reality. The convergence of generative AI-based attack frameworks and the professionalization of DDoS-as-a-service offerings has effectively lowered the barrier to entry, allowing even novice hackers to wield the power once reserved for nation-states.

These defining trends shaped the 2025 landscape and will dictate our defensive posture in 2026:

1. **The Pincer Movement:** We are witnessing a dual assault. On one side, a violent revival of volumetric network DDoS attacks have reached record-breaking benchmarks of 29.7 Tbps. On the other, a surgical shift toward sophisticated application-layer strikes target the very business logic of modern enterprises.

2. **Time Compression:** The five-minute problem has become a critical challenge for traditional security. With the majority of record-level DDoS attacks now lasting less than 60 seconds, manual runbooks and human-in-the-loop interventions have officially become obsolete.

3. **The AI Identity Crisis:** The rapid deployment of AI agents has fractured our digital trust models. As platforms allow automated POST requests for "good" bots, malicious actors are spoofing these identities to bypass traditional mitigation, turning our most innovative tools into conduits for data exfiltration.

4. **Invisible IPI Vulnerabilities:** This crisis is deepened by zero-click indirect prompt injection (IPI) attacks. Vulnerabilities like ShadowLeak allow for service-side data exfiltration that bypasses traditional perimeter defenses entirely. Meanwhile, ZombieAgent represents a shift toward persistent compromise, embedding malicious logic into an agent's long-term memory to turn trusted AI assistants into persistent, silent insider threats.

Our analysis indicates that the "cyber apple," the lucrative prize of unpatched vulnerabilities and sensitive data, is now the focus of intense competition among a growing number of criminals. We have moved past the era of defending against individual threats and are now defending against automated ecosystems that evolve in real time.

The 2026 report is not just a collection of data points; it is a call to action. To survive this new reality, organizations must abandon the reactive assumptions of the last decade and embrace proactive, self-defending architectures built on the pillars of automation, scale and intelligence.

The digital apple is ripe, and the competition to claim it has never been more aggressive. It is time for defenders to change the rules of the game.

Pascal Geenens, Executive Editor, VP Cyber Threat Intelligence, Radware

Executive Summary

DDoS: End of an Era, Dawn of a New Threat Landscape

The year 2025 marks a fundamental paradigm shift in the DDoS threat landscape. The changes observed are not merely incremental; they represent a new reality defined by unprecedented scale, algorithmic speed and a complete transformation of the underlying attack infrastructure.

Terabit-scale distributed denial-of-service (DDoS) attacks, once a rare hundred-year-storm-level event, now happen regularly to major cloud and telecommunications providers. In 2025, massive new botnets emerged, including Aisuru and Kimwolf. Aisuru set a new global benchmark for volumetric carpet-bombing DDoS attacks, achieving a record-breaking 29.7 terabits per second (Tbps). These botnets are created and maintained by cybercriminals who provide DDoS-for-hire services, enabling even low-skilled actors to launch multi-Tbps attacks.

The first half of 2025 revealed a resurgence of brute-force network-layer attacks. After several years of being overshadowed by application-layer DDoS attack trends, network-layer DDoS[1] attacks roared back in the first half of 2025, increasing by 84% compared to the second half of 2024. This growth was not an isolated event but an accelerating trend, with the attack frequency escalating to alarming levels. In the second half of 2025, the average Radware customer faced more than 25,351 DDoS attacks, an average of 139 per day. The result is a staggering 168.2% increase in DDoS attacks in 2025 compared to 2024. The primary drivers of network DDoS attacks in 2025 were UDP floods, which accounted for half of the mitigated volume.

Geographically, North America was the primary target, accounting for 63.1% of all network DDoS attacks, followed by the Middle East (16.1%) and Europe (13.7%). In 2024, the finance sector was the primary target of network DDoS attacks, accounting for 41.5% of all attacks, while the technology sector was a relatively minor target at just 8.77%. By 2025, these positions essentially reversed: technology became the most attacked vertical by a wide margin, accounting for 45% of all network attacks, while finance dropped to 16.1%, tying with telecom for second place. The technology sector's surge represents a strategic pivot toward critical Internet infrastructure. By targeting hosting providers and cloud services, attackers can inflict widespread collateral damage that cascades down to thousands of downstream customers.

[1] Network-layer DDoS attacks target Layers 3 (Network) and 4 (Transport) of the OSI model. Layer 3 DDoS attacks, such as UDP floods, primarily target network infrastructure by overwhelming it with excessive amounts of network traffic. Layer 4 DDoS attacks, such as SYN floods, target the transport protocols to exhaust server connection states and resources.

While finance dropped to second place behind technology, it remains a critical target given its role as the cornerstone of online applications. Financial institutions face highly sophisticated, multi-vector campaigns designed to bypass standard mitigations.

This resurgence of brute-force network attacks did not occur in a vacuum. Rather, it formed one half of a pincer movement, complemented by an equally significant evolution in sophisticated Web DDoS attacks targeting the application layer. The tactics for Web DDoS attacks evolved significantly in 2025, featuring a marked shift away from large-scale attacks toward smaller, more frequent, and more persistent campaigns. Overall activity surged, with Web DDoS attacks increasing by 101.4% in 2025 compared to 2024. The key tactical shift lies in the attack size. Unlike previous years characterized by massive, high-intensity assaults, 2025 is defined by a higher frequency of smaller incidents. Our analysis shows that 94.4% of Web DDoS attacks were under 100,000 requests per second (RPS). This trend is in correlation with the democratization of DDoS capabilities, driven by the proliferation of open-source tools and generative AI (GenAI), which made Web DDoS the preferred attack method for a new wave of threat actors. Despite this trend, the most capable adversaries retain formidable power, with the largest Web DDoS attacks reaching beyond 10 million RPS. Geographically, EMEA remained the primary target region, accounting for 57% of the attacks, while the APAC region saw a 485% increase in attacks in 2025 compared to 2024.

The Five-Minute Problem

According to statistics reported in the media, the majority of volumetric attacks attributed to the Aisuru botnet lasted no more than five minutes, while most record-level Web DDoS attacks lasted less than 60 seconds. The duration of the largest DDoS campaigns has shrunk dramatically. For defenders, this time compression is a critical challenge. If detection and mitigation cannot respond at the network edge in under 60 seconds, then the only option for defense can be post-incident analysis. These are not simple, single-vector attacks; they are algorithmically orchestrated campaigns in which attack vectors cycle through a number of attacks—including carpet bombing, various types of flooding, direct-path and reflection/amplification attacks—faster than human operators can respond. This renders manual runbooks and human-in-the-loop interventions obsolete.

The average attack time of more “common” attacks, as we could call those that make up the majority of the network DDoS attacks observed by organizations protected by our cloud protection service in 2025, shows that attacks between 100 and 500Gbps last on average 10.2 hours, while multi-terabit attacks last on average 35 minutes. A strategy of just sitting out the five minutes of record-level aggression is not a good one either. Adequate mitigation solutions need to do more than respond within seconds to record-level flash attacks. They must also sustain their mitigation for hours of Tbps traffic volumes.

Application Layer: The Modern Center of Gravity

While DDoS attacks grab headlines with their sheer volume, the most sophisticated and potentially damaging attacks increasingly target business logic and data at the application and API layers. This is where organizations' most valuable assets reside, and threat actors firmly established it as their central battleground in 2024. In 2025, this arena was again defined by explosive growth in malicious transactions, a tactical shift toward vulnerability exploitation and a surge in automated bot attacks.

The focus on the application layer intensified dramatically in 2025. Radware's Cloud Application Protection Service detected a 128% increase in malicious transactions compared to 2024. The scale of this trend is stark: malicious activity recorded in the first six months of 2025 alone accounted for 87% of the total recorded throughout 2024, and the second half saw another 63% increase on top of the first.

An analysis of attack types reveals a tactical evolution by adversaries, with vulnerability exploitation (41.8%) as the leading category. Vulnerability exploitation became even more significant in Q4, accounting for nearly 58% of the attack activity during that quarter. Access violations and data leakage have decreased slightly compared to 2024, with rates of 9.4% and 4.3%, down from 10.8% and 5.6%. Most notably, traditional SQL injection attacks declined from a 2.2% share to a 1.1% share. This decline does not signal a more secure landscape; instead, it indicates that adversaries are abandoning commoditized, easily defended vectors in favor of more sophisticated, evasive attacks against business logic and unmanaged APIs.

The execution of application-layer attacks is increasingly automated. Bad bots are a primary tool for adversaries. In 2025, bad bot activity grew 91.8% compared to 2024. This volume was so significant that activity in the first six months of 2025 already accounted for 89.2% of the total for all of 2024. This exponential growth highlights the central role of automation in modern cybercrime. Regionally, North America was the most targeted region for bad bot activity, accounting for 40.7% of malicious transactions, followed by APAC (25%), EMEA (19.1%) and CALA (15.2%).

AI's Identity Dilemma: Malicious Bots in Disguise

The rapid deployment of AI assistants and agents in 2025 has created a significant security vulnerability by forcing platforms to allow automated POST requests, a privilege typically restricted to prevent malicious activity. This shift has fractured the digital trust model, as attackers can now spoof the identities of AI agents that lack robust verification standards, such as those from Anthropic or xAI, to bypass traditional bot mitigation and execute account takeovers or data scraping. While leaders like OpenAI and Google utilize cryptographic signatures or DNS lookups for validation, the reliance of other providers on easily manipulated User-Agent strings creates a massive blind spot, leaving security teams unable to distinguish between legitimate AI transactions and sophisticated malicious bots operating under the guise of trusted agents.

Hacktivism: A Persistent Threat

Beyond the technical evolution of attacks, a primary driver of DDoS activity remains geopolitical and ideological conflict. In 2025, hacktivist targeting remained sustained and evolved into a persistent, high-volume threat.

Hacktivist-led DDoS campaigns continued at a relentless pace. This sustained high frequency of attacks reflects a mature ecosystem in which hundreds of Telegram channels coordinate attacks and post about their campaigns, amplifying their visibility and psychological impact.

Hacktivist campaigns in 2025 were global in scope but concentrated on specific regions embroiled in political conflict. Europe was the most targeted region, bearing 48.4% of all claimed attacks, followed by the Middle East (17.7%) and Asia (17.5%). This geographic focus reflects the primary theaters of ongoing international tensions.

This regional activity was further concentrated against several nations. The top three most targeted countries were Israel (12.2%), the United States (9.4%) and Ukraine (8.9%). This distribution underscores the role of hacktivism as a proxy weapon in international conflicts, used to disrupt the digital infrastructure of perceived state adversaries.

Across all targeted nations, government services remained the primary focus for hacktivists, accounting for 38.8% of all attacks as groups sought to disrupt state functions and undermine public confidence. The next most targeted sectors were manufacturing (8%) and hospitality (6%), indicating a strategic effort to inflict both political and economic damage.

While the hacktivist landscape is comprised of numerous groups, one continues to dominate. An analysis of claimed attacks found that the pro-Russian group NoName057(16) was responsible for 4,692 attack claims, making it the most prolific hacktivist actor not only in 2025, but in the history of hacktivism.

Other highly active groups in 2025 included Keymous+, Hezi Rash, Mr Hamza, Anonymous VNLBN and RipperSec. The sustained activity of these groups demonstrates that ideologically motivated attacks remain a persistent and significant part of the threat landscape.

Regional Observations

Europe, the Middle East and Africa (EMEA)

EMEA remained the primary focus for Web DDoS attacks in 2025, maintaining a commanding 57% share of all Web DDoS attacks despite a relatively more moderate year-over-year growth of 47% compared to other regions. This dominance is largely a byproduct of its status as the most targeted region for hacktivist operations, accounting for 48.4% of all global DDoS attack claims recorded during the period. Geopolitical tensions significantly influenced this landscape, with Israel and Ukraine emerging as the most heavily targeted nations.

The scale of activity in EMEA is further illustrated by the concentration of highly visible cyber campaigns, such as #OPISRAEL and #OPUKRAINE. These efforts often targeted critical state infrastructure, including the Ukrainian government's rada.gov.ua, which faced the highest number of individual targeted attacks globally. While the region's share of network DDoS attacks declined slightly from 34.3% to 30.5%[2], its online application and API attack share declined from 25.5 to 15.1%, and its bad bot transaction share remained relatively stable at roughly 19%. Note that shares are relative measures in a sharply increasing trend, so even though the online application and API attack share declined by 10% from 2024 to 2025, the number of attacks in the region increased by 4.7% in 2025 compared to 2024. EMEA's position as a central theater for application-layer DDoS attacks remains its defining feature of 2025.

Asia-Pacific (APAC)

The APAC region emerged as the most volatile region in 2025, characterized by a staggering 485% year-over-year increase in Web DDoS attacks. By the end of the year, it had captured nearly a quarter of the global share for these attacks, rising to 24.2%. APAC's share of bad bot transactions jumped to 25% in 2025, up from just 16.6% the previous year, while its share in online application and API attacks slightly increased from 7.5% to 8.8%. This rapid escalation in Web DDoS attacks suggests a major strategic pivot by threat actors toward the region's digital infrastructure, supported by a significant increase in automated malicious activity, indicating an increasingly automated and persistent threat environment.

Hacktivism also played a major role in this regional surge, with APAC accounting for 17.5% of global attack claims. This activity was driven by prolific regional groups such as Keymous+ and Mr Hamza. These groups frequently targeted government institutions, with domains in Taiwan, Thailand and Vietnam appearing prominently on the list of most-targeted domains. The presence of large-scale operations like #OPINDIA further underscores the growing link between regional politics and disruptive cyber events.

North America

North America remained the global leader in network DDoS activity, with its share of global attacks rising from approximately 50% in 2024 to 63.1% in 2025. This concentration is closely tied to the region's high density of technology, telecom and financial services providers, which were the top three targeted verticals globally during the year. The United States specifically experienced a sharp 223% increase in Web DDoS attacks as illustrated by the DieNet campaign targeting U.S. organizations in March.

[2] Network DDoS attack shares in 2024: Europe (18.5%), Africa (12.7%), Middle East (3.1%)
Network DDoS attack shares in 2025: Europe (13.7%), Africa (0.7%), Middle East (16.1%)

In addition to network floods, the region remained the primary destination for malicious web application and API transactions, accounting for 73.7% of global transactions, and the most important target for automated traffic, accounting for 40.7% of all detected bad bot transactions. North America was also a major secondary target for hacktivist groups; the United States was the second most attacked country, with 1,447 attack claims against its infrastructure throughout 2025.

Central and Latin America (CALA)

The year 2025 marked a fundamental paradigm shift in the threat landscape across Central and Latin America (CALA), as the region saw significant growth in automated and application-layer attack activity. While North America and EMEA continued to handle the highest overall volumes of global traffic, CALA experienced a substantial 146% year-over-year increase in Web DDoS attacks. This surge indicates that threat actors are increasingly targeting the region's digital infrastructure with more frequent and persistent campaigns, mirroring the pincer movement of surgical application strikes observed globally.

On the online application and API front, CALA accounted for 2.4% of global malicious transactions in 2025, up from 1.4% in 2024. The most notable shift in the region, however, was the proliferation of automated threats, with CALA's share of global bad bot transactions rising from 13.4% in 2024 to 15.2% in 2025. As bad bot activity nearly doubled worldwide, the region has become a very active theater for coordinated ecosystems performing large-scale scraping and account takeover operations.

Finally, the relative decline of network DDoS attack shares in the region from 5.59% in 2024 to 1.33% in 2025 reinforces the overall strategic pivot toward the application layer in CALA. This trend suggests that attackers are less focused on commoditized network-layer vectors and are instead pursuing more sophisticated, evasive attacks against business logic and APIs.

Artificial Intelligence: A Vastly Increasing Threat Surface

The rapid adoption of generative AI is fundamentally reshaping the cyberthreat landscape, creating a dual reality in which innovation and risk advance in parallel. AI has dramatically lowered the barrier to entry for cybercrime, democratizing offensive capabilities that were once reserved for highly skilled or well-funded actors. Through natural language prompts alone, minimally technical individuals can now orchestrate sophisticated attacks, accelerating both the speed and scale of cyber aggression. This shift has fueled a thriving as-a-service underground economy and given rise to "vibe hacking," which leverages AI to build and execute attack tools, compressing timelines from days or weeks to hours.

Simultaneously, offensive and malicious service providers are professionalizing their operations through autonomous, agentic attack frameworks. Ecosystems such as Xanthorox AI exemplify this evolution, automating several steps in the attack kill chain and introducing self-healing logic that dynamically bypasses model safeguards and routes tasks to other, less restrictive models. These systems are increasingly operating less as assistants and more as independent actors capable of real-time decision-making, marking a strategic inflection point in cyber-offense tools readily available to the public.

Along with being used as a malicious tool, AI is also a target of these new threats. Enterprises now face a new class of stealthy threats targeting AI assistants. Zero-click indirect prompt injection (IPI) attacks, such as ShadowLeak and ZombieAgent, exploit trusted content channels to exfiltrate sensitive data directly from the AI provider's infrastructure, bypassing traditional perimeter defenses and endpoint controls. The emergence of persistent compromises like ZombieAgent further escalates risk by embedding malicious logic into an agent's persistent memory, effectively turning AI assistants into silent insiders that act against organizational interests across sessions.

Looking ahead, the transition to an Internet of Agents will magnify these risks. As autonomous agents interconnect via standardized protocols and APIs, the attack surface expands from individual models to entire agentic ecosystems. The primary risk is no longer what AI systems say, but what they do: compromised agents can trigger cascading failures across interconnected services. And with APIs serving as the backbone of many agentic AI services and protocols, they will become even more central in the 2026 threat landscape.

Conclusion: A New Defensive Posture for a New Reality

The threat landscape of 2026 demands that defenders abandon the assumptions that have served for the last decade. The speed, scale and sophistication of modern attacks render traditional, manually driven defenses insufficient.

To survive in this new reality, a modern defensive posture must be built on three core pillars.

1. **Automation** – Humans cannot match the speed of algorithmic attacks that execute and pivot in minutes; defense must be automated to detect and mitigate threats in real-time without intervention.
2. **Massive Scale** – Terabit-class floods must be absorbed at the network edge, requiring architectures with the capacity to handle previously unimaginable traffic volumes.
3. **Integrated Intelligence** – Sophisticated behavioral analysis is needed to distinguish malicious traffic originating from legitimate-looking residential IPs and, increasingly, from AI assistants and AI browsers.

Attackers have already transitioned to this new paradigm of automated, intelligent and massively scaled warfare. The critical question for 2026 is how many defenders will evolve from reactive, manual processes to the proactive, self-defending architectures required to survive.

DDoS Threat Landscape

Web DDoS Attack Activity

Web DDoS activity escalated significantly in 2025, marking a decisive turning point in both scale and frequency. By the end of the first six months of 2025, malicious web activity already accounted for 87% of the total recorded in 2024, firmly establishing the application layer as a central battleground in the threat landscape. Over the past year, Web DDoS activity increased by 101.4% compared to 2024. This growth was not linear; the first half of 2025 saw a 38.5% increase over the previous half year, and this momentum intensified as the year progressed. The activity in the second half accelerated significantly with a 54.5% surge compared to the first half. The significant surge in Web DDoS activity is being fueled by the use of generative AI tools that democratize attack capabilities, lowering the barrier to entry by enabling less-experienced actors to launch campaigns without coding skills. This surge also correlates directly with heightened geopolitical friction across EMEA and APAC, where the expectation of a digital disruption follows any major political announcement or escalation in a conflict.

Figure 1:

Evolution of Web DDoS attacks mitigated per year (source: Radware)

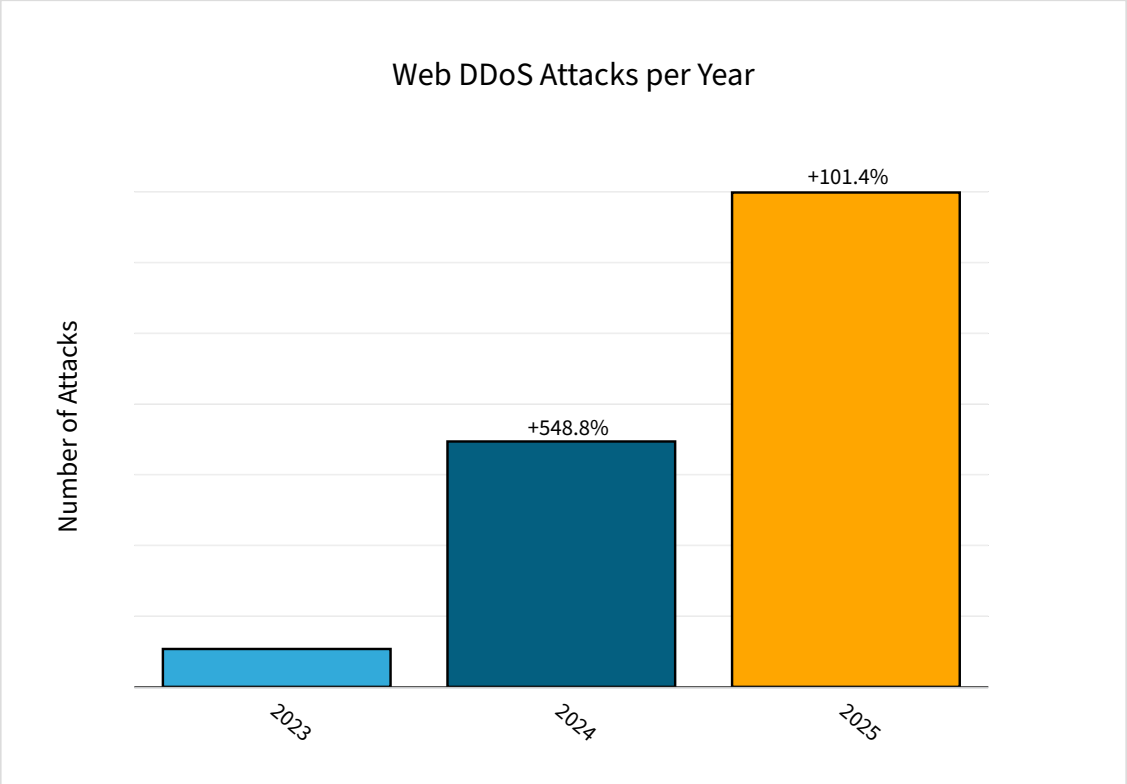
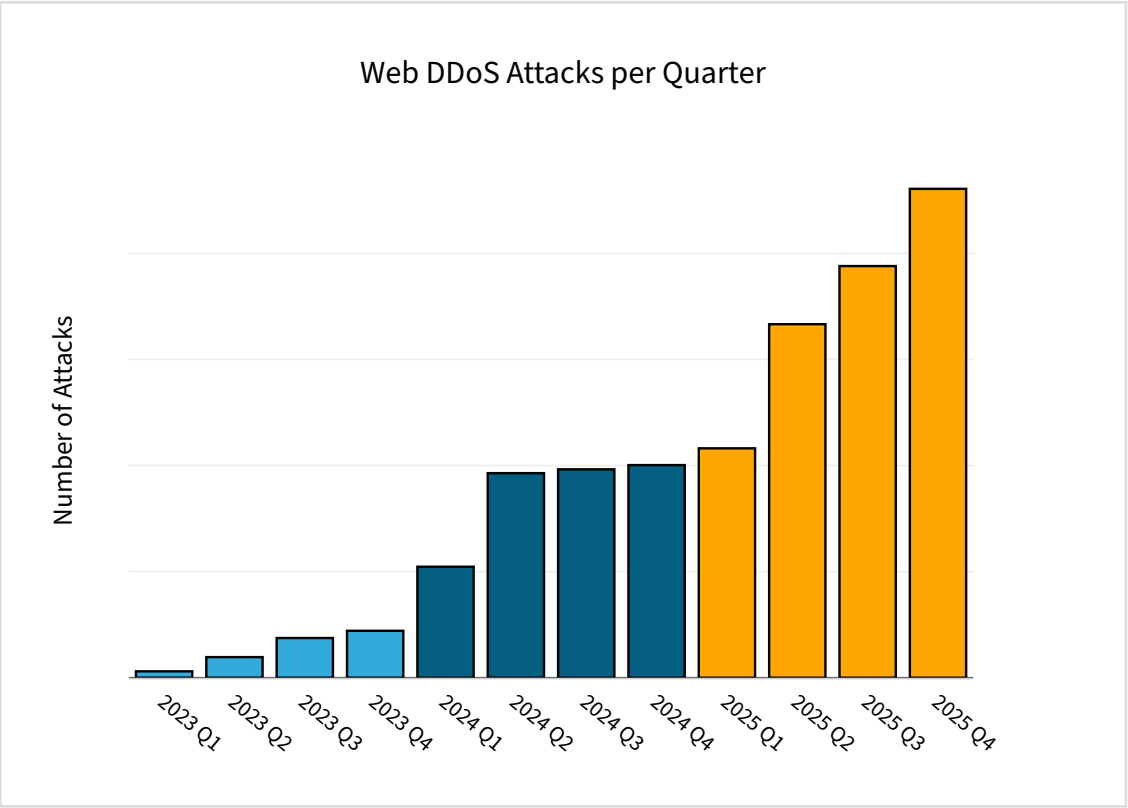


Figure 2:

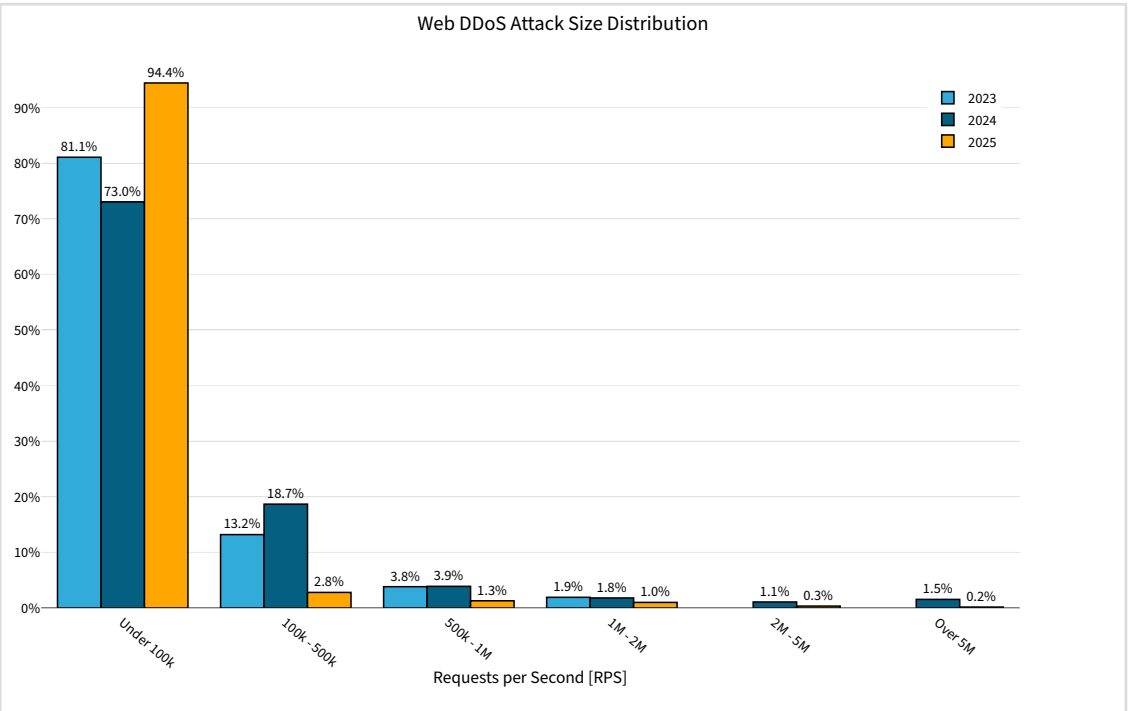
Evolution of Web DDoS attacks mitigated per quarter (source: Radware)



Advanced threat actors are increasingly favoring smaller, sustained attacks to evade traditional volumetric detection. Attacks measuring under 100,000 RPS accounted for 94.4% of all events, a significant jump from 73% in 2024. Conversely, high-intensity attacks exceeding 5 million RPS accounted for only 0.2% of all attacks, compared to 1.5% in 2024. Note that this does not mean that the number of high-intensity attacks decreased; rather, their share declined. Web DDoS attacks with more than 10 million RPS have not disappeared from the threat landscape.

Figure 3:

Web DDoS attack size (RPS) distribution per year (source: Radware)

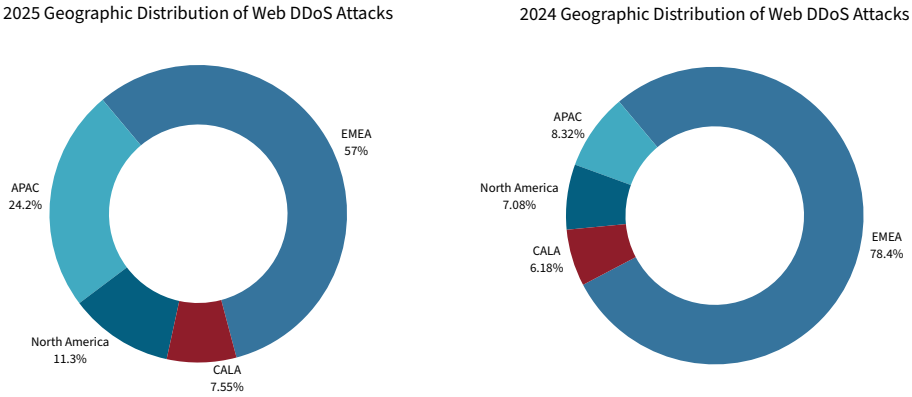


Geographical Activity

While EMEA remained the most targeted region for Web DDoS (57% share), the APAC region saw the most aggressive year-over-year growth at 485%, followed by the U.S. at 223%.

Figure 4:

Geographic distribution of Web DDoS attack activity (source: Radware)



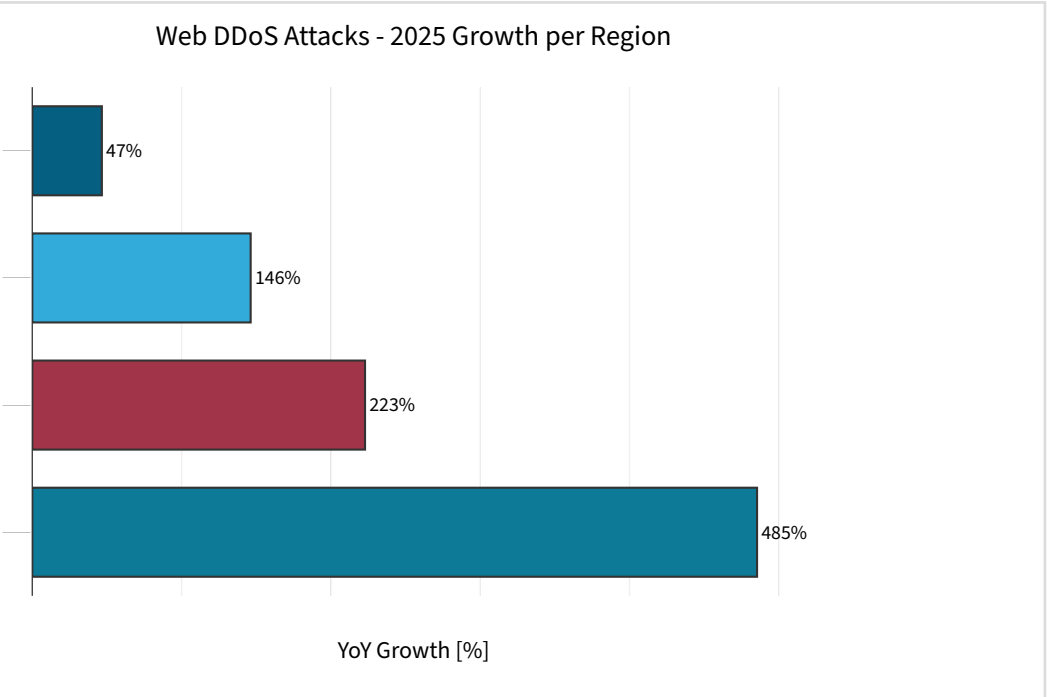
The global distribution of Web DDoS attack activity shows a world in digital turmoil with clear regional hot spots. EMEA remains the primary theater for Web DDoS, largely fueled by the ongoing Russia-Ukraine and Middle Eastern conflicts. While its 47% growth is the lowest relatively, it represents the highest absolute volume of attacks.

The near-fivefold growth in APAC is the story of 2025 for that region. This explosion is tied to burgeoning tensions in the area and a massive uptick in localized hacktivism targeting the Indian subcontinent and Southeast Asian governments and financial hubs.

North America remains a lucrative target, particularly for DDoS-as-a-service providers and state-aligned actors targeting critical infrastructure and the financial sector.

Figure 5:

Growth of Web DDoS activity per region in 2025 compared to 2024 (source: Radware)

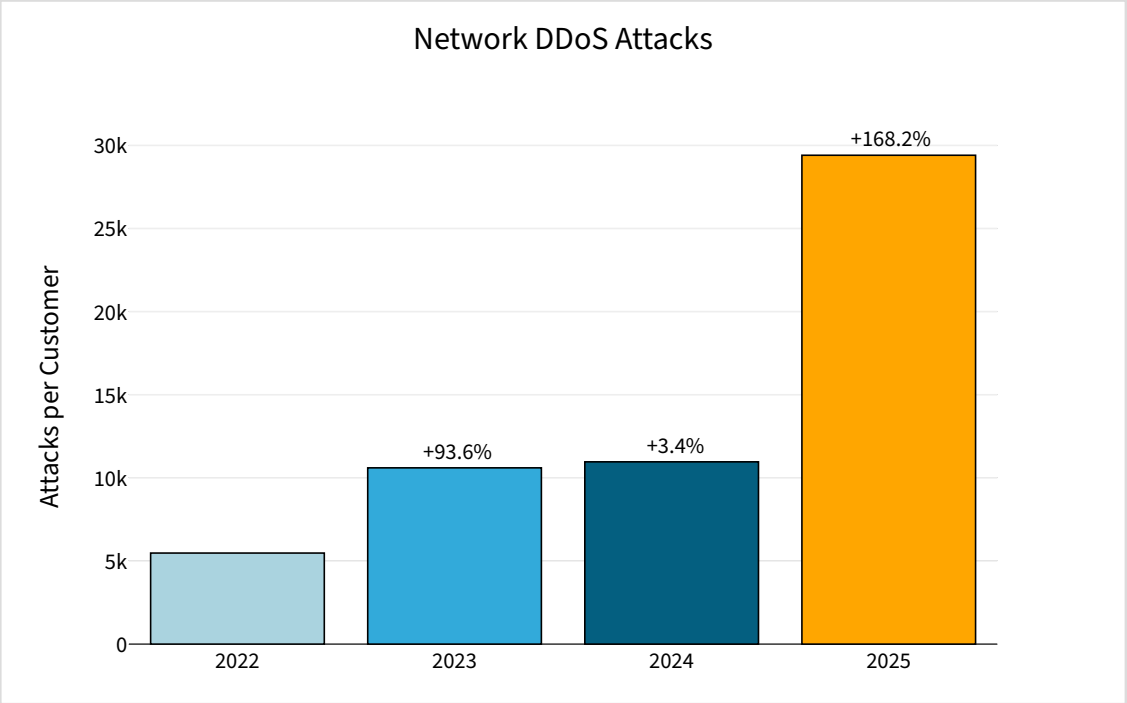


Network DDoS Attack Activity

In 2025 network DDoS attacks combined with a surge in Web DDoS activity to form a “pincer movement” in the threat landscape. Although the previous section of this report details the rise of surgical, encrypted Web DDoS strikes, our data confirms that threat actors have not abandoned their big guns. On the contrary, 2025 marked a violent revival of volumetric network DDoS, where brute force met unprecedented scale and technical sophistication.

Figure 6:

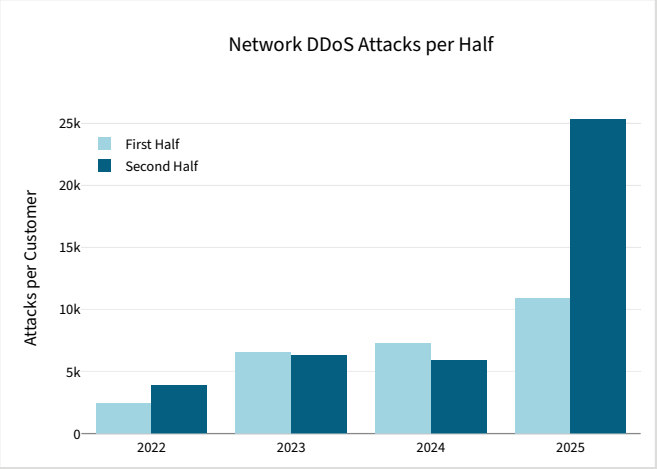
Evolution of network DDoS attacks normalized per customer (source: Radware)



Cybercriminals’ own version of “big blaster cannons” returned with a vengeance in 2025, as network DDoS attacks per customer surged by 168.2% compared to 2024, with momentum accelerating throughout the year.

Figure 7:

Evolution of network DDoS attacks normalized per customer, per half year (source: Radware)



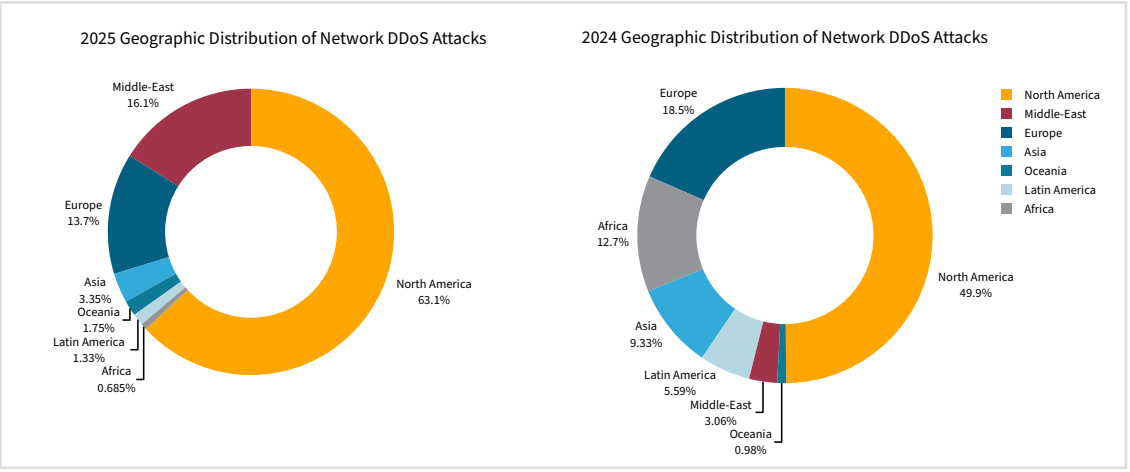
The second half of 2025 alone rose 132.5% when compared to the first half of 2025 and a staggering 328% when compared year over year with the same period in 2024. The data signals not just incremental growth but a clear phase shift in which attackers are once again exploiting massive new botnets and amplification vectors to unleash flood-based attacks that were previously believed to be in decline in favor of application-layer techniques.

Geographic and Industrial Targets

The “where” and “who” of network DDoS attacks shifted significantly toward high-value infrastructure and Western targets in 2025.

Figure 8:

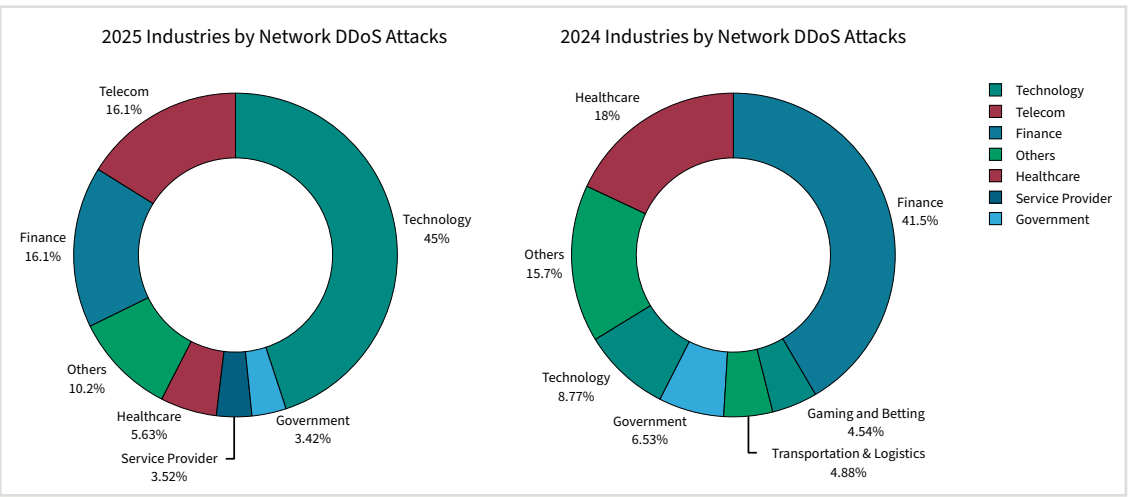
Network DDoS attack distribution by regions (source: Radware)



North America emerged as the primary theater for network aggression, accounting for 63.1% of global activity, up from 49.9% in 2024. The Middle East also saw a dramatic escalation, with its share more than quintupling to 16.1% from 3.1% the previous year. This underscores the emerging prominence of large-scale volumetric floods in regional geopolitical conflicts in 2025. Europe remained a major target, accounting for 13.7%, but its relative share declined from 18.5%.

Figure 9:

Network DDoS attack distribution by industry (source: Radware)



At the same time, attackers pivoted toward critical infrastructure sectors. The technology industry accounted for nearly half of all network DDoS attacks (45%), a substantial increase from 8.8% in 2024, making it the most targeted vertical. Telecommunications and financial services each accounted for 16.1% of attacks, underscoring their continued importance as targets for disruption.

When normalized to the customer base, the automotive sector stands out as an unexpectedly high-intensity target. Its 19.9% share indicates growing attacker interest in network-level disruption of industrial and supply-chain-linked environments.

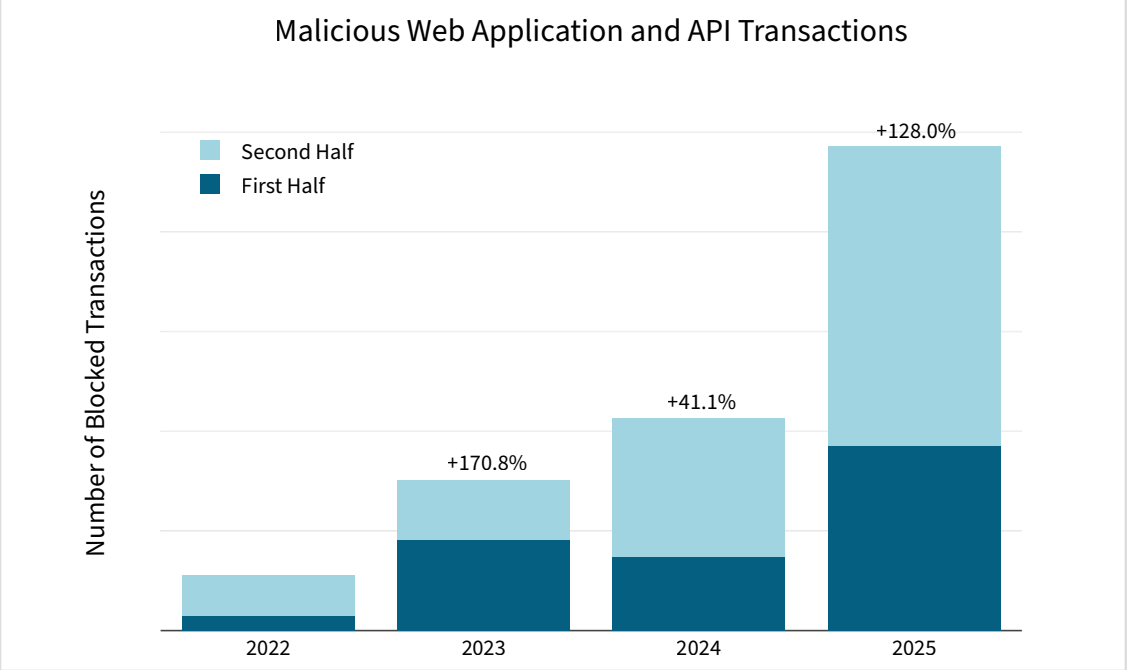
Web Application and API Threat Landscape

Web Application and API Attack Activity

The explosive growth in malicious web application and API transactions throughout 2025 represents a critical shift in the threat landscape, signaling that attackers have fully industrialized their exploitation of the application layer. Our data indicates that the total volume of malicious transactions in 2025 increased by 128% year over year compared with 2024.

Figure 10:

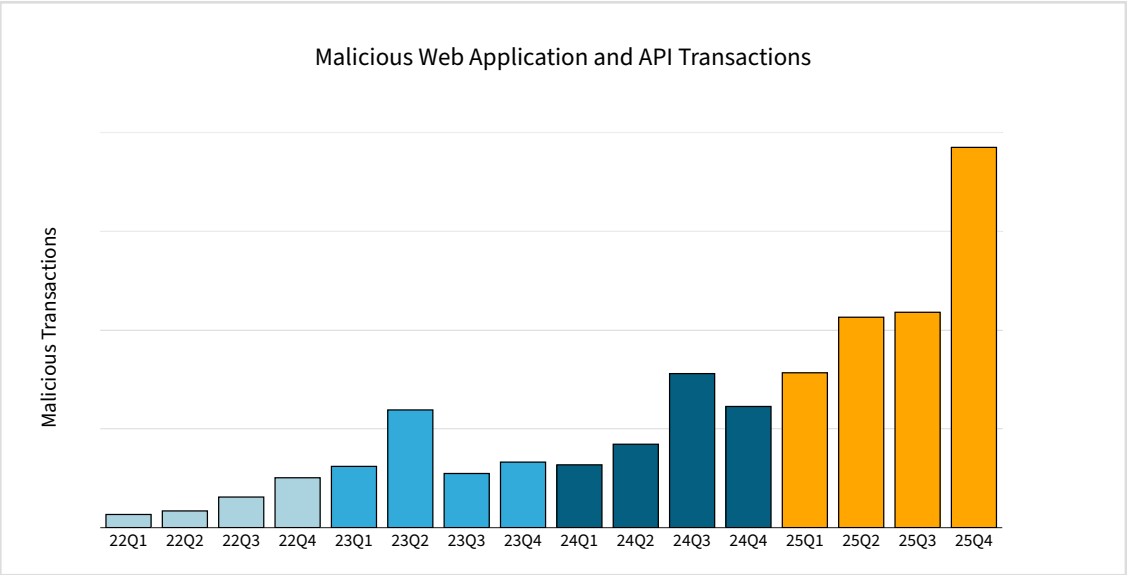
Malicious web application and API transactions per year (source: Radware)



This surge was characterized by accelerating momentum as the year progressed, with the second half of 2025 seeing a 63% increase in blocked transactions relative to the first half. The upward trajectory was particularly pronounced in the fourth quarter, which saw a massive 213.8% growth in malicious activity compared to the same period in 2024. This indicates that in Q4 2025, more than three times as many attacks were mitigated as in Q4 2024. These figures suggest that while technical sophistication is increasing, the sheer volume of noise generated by automated probes, bot-driven credential stuffing and vulnerability scans is scaling at a rate that legacy security perimeters will struggle to contain.

Figure 11:

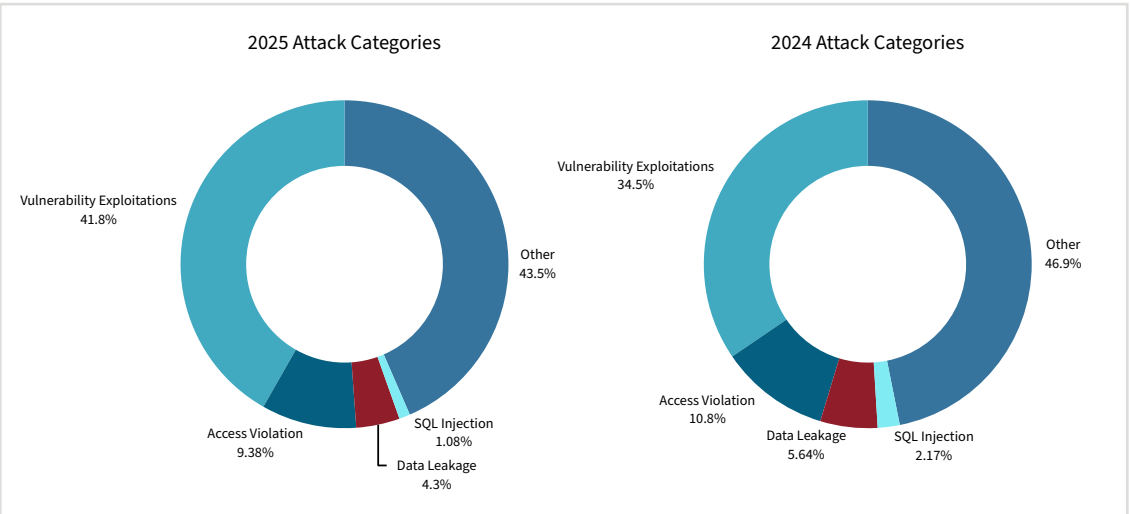
Malicious web application and API transactions per quarter (source: Radware)



The composition of these attacks also reveals a strategic pivot towards the exploitation of software flaws rather than generic noise or access violations typical for web crawling. In 2025, vulnerability exploitations emerged as the dominant threat category, accounting for 41.8% of all malicious activity—up from a 34.5% share in the previous year. This increase suggests that threat actors are more effectively weaponizing newly disclosed common vulnerabilities and exposures (CVEs) and targeting unpatched APIs as primary entry points into corporate environments. Meanwhile, the share of access violations declined slightly from 10.8% to 9.38% and SQL injection attempts fell from 2.17% to 1.08%, indicating that attackers are shifting away from traditional SQL injection patterns toward more complex vulnerability probes. The category labeled "Other," which encompasses sophisticated attack activity such as business logic abuse, remains a significant portion of the attack surface at 43.5%, underscoring the need for behavior-based detection next to simple signature matching.

Figure 12:

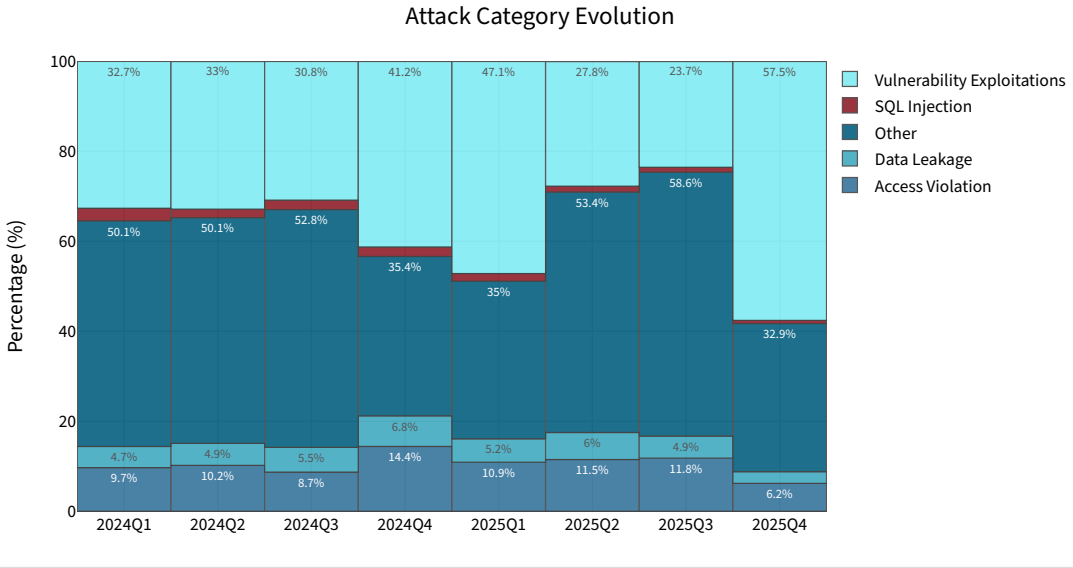
Web application and API attacks by category (source: Radware)



The massive growth in malicious transactions in Q4 was driven by the exploitation of vulnerabilities, which correlated directly with the high volume of critical, unauthenticated remote code execution flaws discovered in web application components and internet-connected devices, catalyzed by the weaponization of high-impact flaws such as React2Shell (CVE-2025-55182).

Figure 13:

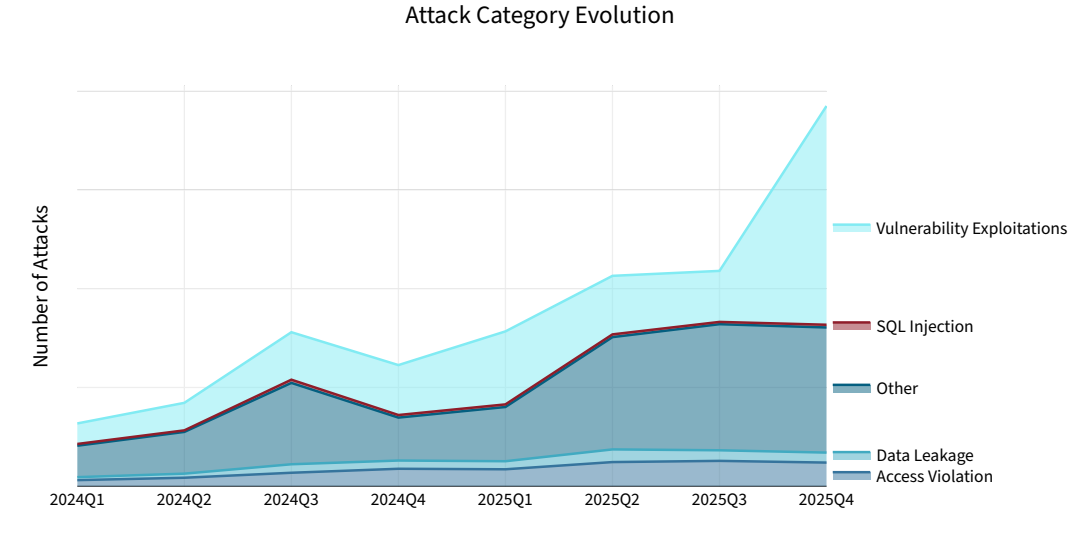
Evolution of web application and API attack categories – normalized (source: Radware)



It is important to recognize that the increased prevalence of vulnerability exploitations in the final quarter of 2025 was an additive surge. While categories such as Access Violations saw their share drop to 6.2%, their absolute transaction counts did not decline; rather, they were simply eclipsed by the sheer volume of new exploitation attempts. This trend is visible in the Web Attack Categories chart in Figure 14, where the baseline layers remain consistent while the Vulnerability Exploitation and Other segments expand to drive the total volume toward new record highs.

Figure 14:

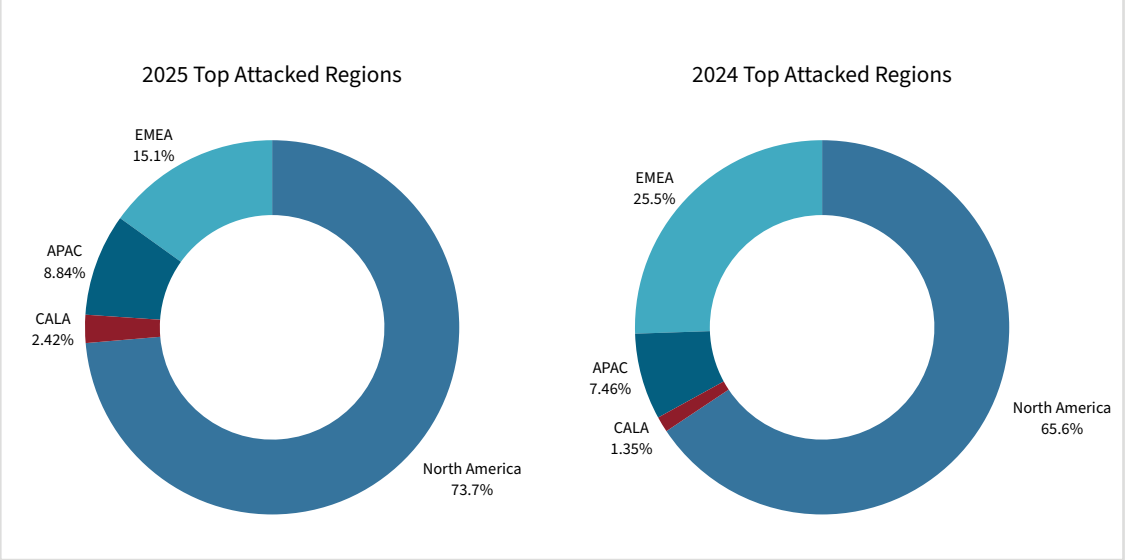
Evolution of web application and API attack categories (source: Radware)



Geographically, the focus of web application and API attacks has shifted significantly toward North America, which accounted for 73.7% of global attack transactions in 2025. This represents a notable increase from the already significant 65.6% share recorded in 2024. Conversely, the share of attacks targeting the EMEA region dropped significantly from 25.5% in 2024 to just 14.6% in 2025. While APAC saw a slight increase in its share from 7.5% to 8.8%, the overwhelming concentration in North America suggests a deliberate prioritization of high-value financial, technological and government targets within that region.

Figure 15:

Web application and API attacks by region (source: Radware)



Cashout Attack Kill Chain: Five Overlooked Indicators

From a threat intelligence perspective, a successful account takeover (ATO) is rarely a single, isolated event; it is a carefully orchestrated campaign that manifests as distinct, often overlooked traces across an organization's infrastructure. By deconstructing the credential stuffing threat landscape, we have identified a unified kill chain comprising five key events that connect pre-attack reconnaissance to post-attack financial fallout.

Events 1 & 2: Reconnaissance and Enumeration

The chain begins with a bypass, where an attack script developer probes login pages and APIs to map defenses. This low-and-slow traffic aims to identify rate limits and detection mechanisms, including which devices or APIs lack robust CAPTCHA protection. Once defenses are mapped, the account cracker initiates the "list" phase. Rather than conspicuously launching direct attacks on login pages, attackers utilize registration and password reset flows as litmus tests to validate user existence, building a valid victim list before the primary attack launches.

Events 3 & 4: The Breach and The Sale

The primary breach typically targets weaker entry points identified during reconnaissance, such as legacy mobile APIs, which often lack the strict security policies of web-based logins. Following a successful breach, the account cracker rarely monetizes the account directly. Instead, a specialization of labor occurs: the validated credentials are sold on dark web marketplaces to an account buyer persona. This actor meticulously logs in to cash out assets via untraceable methods like digital gift cards or shipping address manipulation.

Event 5: The Fallout

The final trace is a lagging indicator observed by finance departments: a spike in chargebacks. By the time legitimate users report unauthorized transactions, the reputational damage and financial loss, including the potential for higher transaction fees from payment processors, have already solidified.

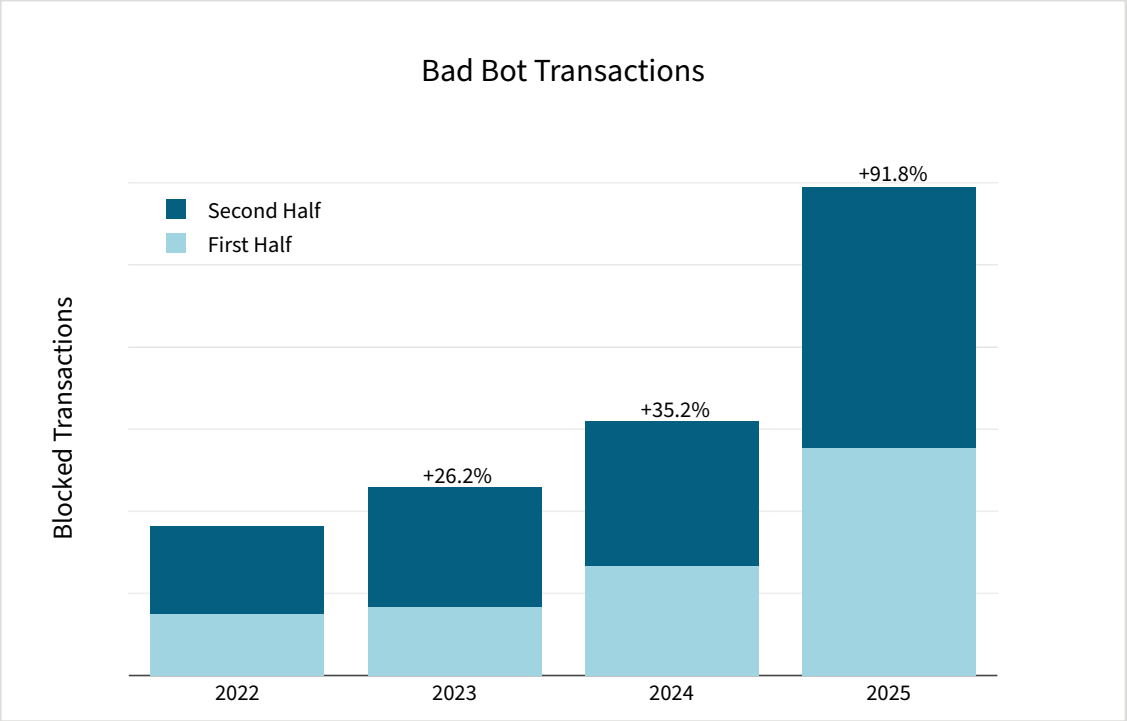
Effective mitigation requires correlating these seemingly disparate events to break the chain at the reconnaissance phase.

Bad Bot Activity

Automated threats surged throughout 2025, with bad bot transactions increasing by 91.8% compared to the previous year. This nearly doubling of the number of bad bot transactions represents a significant acceleration in bot activity when compared to the 35.2% growth observed in 2024 and the 26.2% increase in 2023. The data suggests that bots and automated attacks have reached a new maturity level, with attackers no longer running simple scripts but deploying coordinated bot networks to conduct scraping, account takeover and inventory hoarding at scale.

Figure 16:

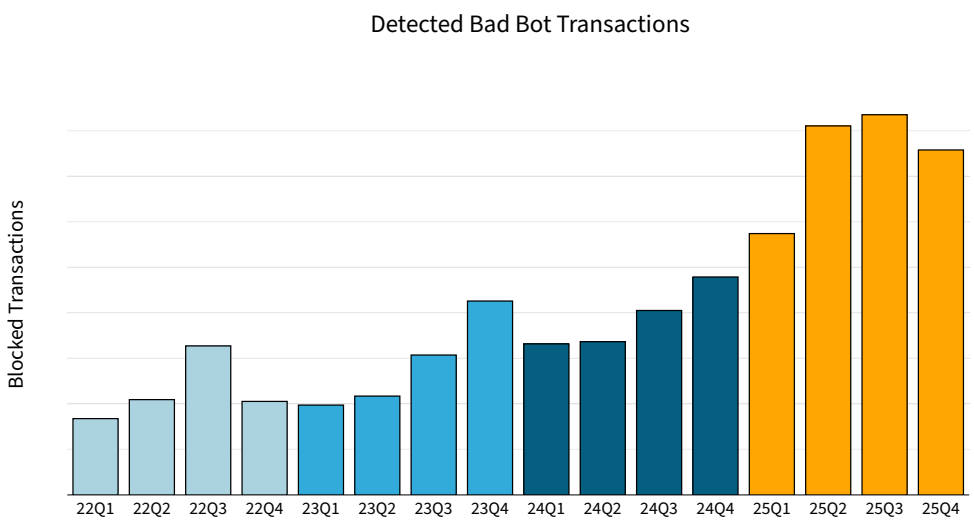
Bad bot transactions per year (source: Radware)



Examining the temporal distribution of these threats reveals a year of high-intensity momentum that began with a substantial 56.7% increase in the first half of 2025 relative to the latter half of 2024. While the growth rate moderated slightly in the second half of the year to an increase of 15%, the baseline of activity remained at historic highs. The quarterly breakdown of detected bad bot transactions shows a steady increase through the first three quarters of 2025, followed by a peak in transactions in Q3 and then a marginal correction in Q4. This sustained high volume indicates that bad bots have become a permanent, always-on component of the threat landscape, serving as the automated engine behind the high volumes of web application attacks and reconnaissance probes discussed in previous sections of this report.

Figure 17:

Bad bot transactions per quarter (source: Radware)

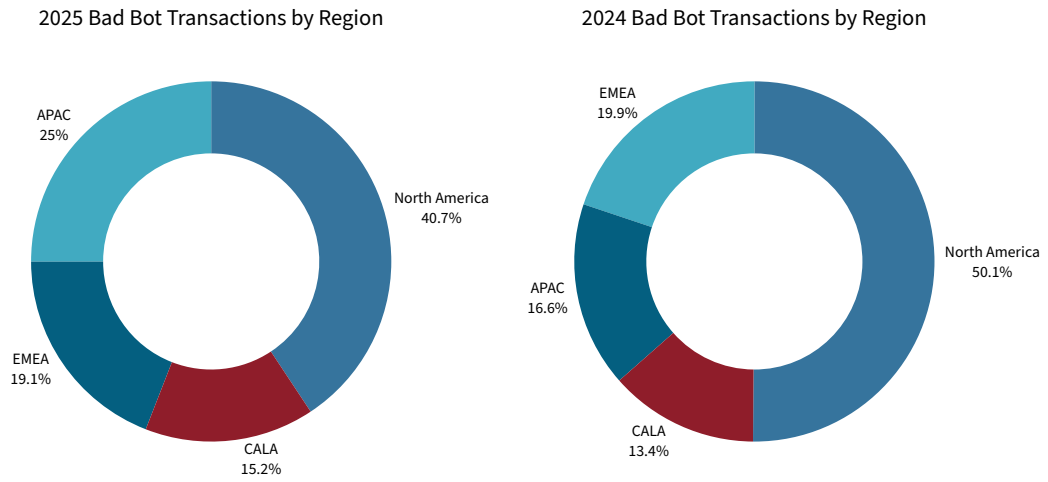


Geographical Activity

Geographically, the distribution of bad bot traffic has undergone a notable shift, signaling that threat actors are diversifying their infrastructure and targeting new markets. While North America remains the primary target, its share of global bad bot transactions decreased from 50.1% in 2024 to 40.7% in 2025. Conversely, the APAC region has emerged as a major focal point for bot activity, with its share increasing from 16.6% to 25% of the global total. This shift toward APAC, combined with a rise in the CALA region's share to 15.2%, highlights the global nature of the bot problem.

Figure 18:

Bad bot transactions per region (source: Radware)



The doubling of bot volume in a single year suggests that automation is the primary force multiplier for modern cybercriminals. We are no longer defending against individual bad bots, but against automated ecosystems that evolve in real time. Organizations must transition from static bot signatures to intent-based detection to protect their business logic from this relentless automated pressure.

MFA and OTP Bots

A critical development in the application threat landscape is the industrialization of social engineering via (OTP) bots. As organizations increasingly adopt multi-factor authentication (MFA), threat actors have pivoted to automated solutions to bypass these controls. Radware's analysis of dark web forums identified 38 distinct OTP bot services, with mentions increasing by 31% year over year.

Automation of Social Engineering

These bots operate primarily via Telegram and represent a sophisticated evolution in fraud. The attack flow typically begins after a credential stuffing attempt hits a MFA wall. The attacker inputs the victim's details into the OTP bot service, which then initiates an automated call or SMS to the victim. Leveraging AI-generated voice scripts, the bot impersonates a trusted entity, such as a bank's fraud department, and urgently requests the one-time password under the guise of account verification.

Scalability and Accessibility

Once the victim discloses the code, the attacker takes full control of the account, often changing passwords and MFA settings to lock out the legitimate owner. This model has democratized high-level fraud; services are available for as little as $10 to $50 per attack, allowing virtually anyone with access to Telegram to bypass MFA protections that were previously considered robust barriers against automated threats.

AI's Identity Dilemma: Malicious Bots in Disguise

The rapid deployment of interactive AI agents has introduced a fractured trust model that malicious actors are actively exploiting. In 2025, major platforms like OpenAI, Google and Anthropic deployed agents capable of real-time web interaction and transactions. To function, these "good" bots require permissions that break traditional security assumptions: they must be allowed to make POST requests to execute tasks like booking tickets or making purchases.

The Security Gap: POST Request Allowlisting

Historically, bot mitigation relied on restricting automated traffic to GET requests only. However, creating exceptions for AI agents has opened a critical vulnerability. Malicious bots can now spoof the identities of legitimate AI agents to gain POST access, thereby bypassing detection systems and executing large-scale ATO or scraping attacks. The risk is compounded by the fact that legitimate AI agents can render dynamic JavaScript, making their traffic patterns nearly indistinguishable from sophisticated malicious browsers.

A Fractured Verification Landscape

The primary enabler of this threat is the inconsistency in how AI providers verify their agents. While OpenAI has established a gold standard using cryptographic HTTP message signatures (RFC 9421) that are resistant to spoofing, other major providers have left significant gaps. Also, Google utilizes reverse DNS lookups and publishes IP ranges, offering a robust method for validation. However, agents from providers like Anthropic (ClaudeBot) and xAI (Grok) rely solely on user-agent strings and do not publish official IP ranges. This makes them trivial to spoof, as attackers need only modify a header string to impersonate these trusted agents.

Security teams now face a dilemma. Accommodating the surge of legitimate AI traffic creates a blind spot where malicious bots, disguised as agents with weak verification standards, can operate undetected.

Hacktivist Threat Landscape

Hacktivism is a complex phenomenon that can be motivated by various factors, including religious and political beliefs. While hacktivists may have different motivations and methods, they all share a desire to use technology to advance their cause and to challenge those they believe are acting against it.

Hacktivists employ a variety of tactics to achieve their goals, and the specific tactics they employ depend on their motivations and the resources available to them. Their methods are constantly evolving as new technologies and platforms emerge. While some tactics may be illegal or unethical, hacktivists argue that they use their skills to promote social or political change and hold powerful organizations and governments accountable for their actions.

Some common tactics used by hacktivists include denial-of-service attacks, website defacements and data breaches.

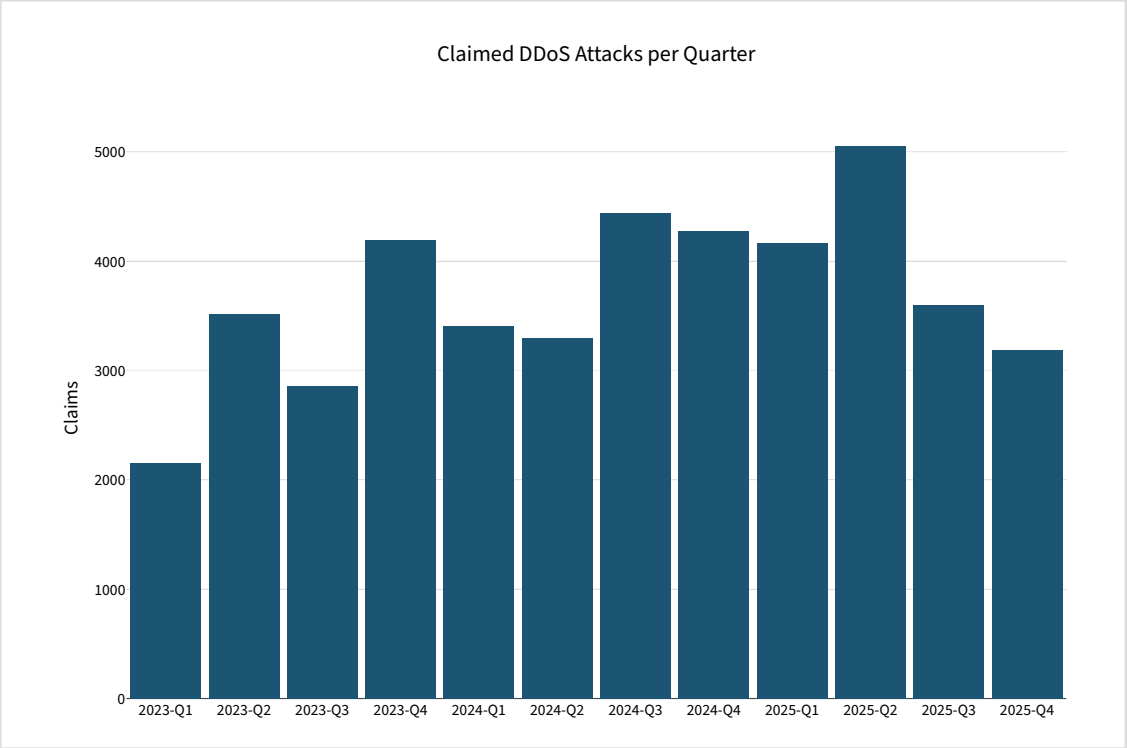
Hacktivist DDoS Claims

Hacktivist groups frequently post claims of their DDoS attacks on Telegram, often providing evidence by sharing snapshots of website availability through check-host links. These links enable verification of the claimed target as well as the date and time of the attack. By focusing on messages containing valid check-host links, it is possible to monitor claimed attacks on Telegram with greater reliability. It is important, however, to note that check-host reports are not infallible and not all attack claims result in downtime of the target. From a cyberthreat intelligence perspective, attack claims provide insight into the threat actors' intent and enable analysis of their tactics and evolution.

Our analysis of messages and claimed DDoS activity reveals a highly organized and ideologically driven environment in 2025. In the first half of 2025 alone, we tracked 7,488 unique attack claims, maintaining the high-intensity baseline established during the previous year's surges. Quarterly data shows that hacktivist aggression reached a fever pitch in Q2 2025, with over 5,000 claimed attacks.

Figure 19:

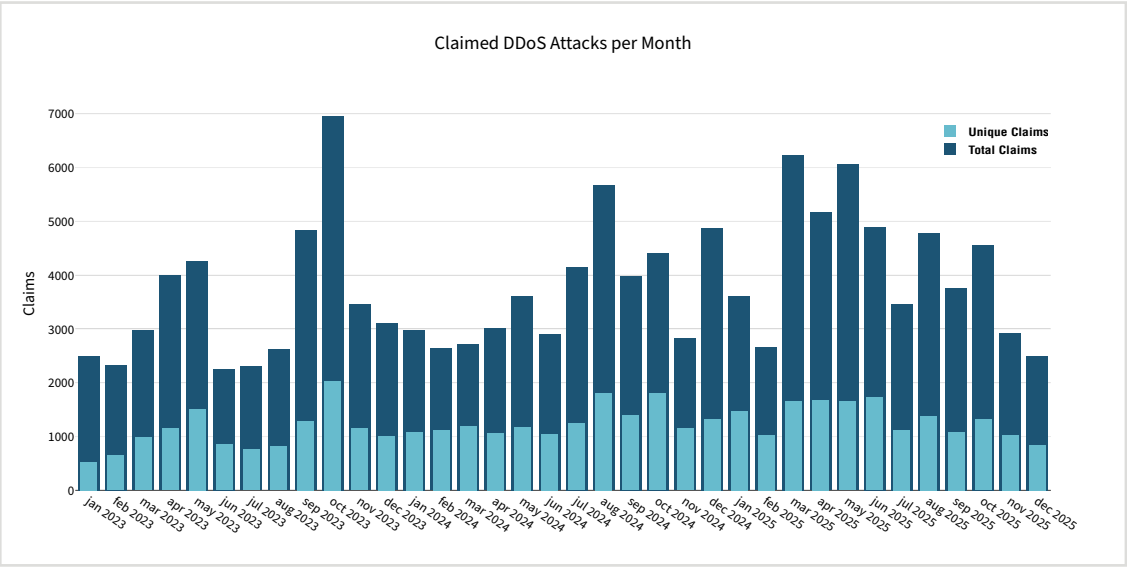
DDoS attack claims per quarter on Telegram (source: Radware)



Attack claims posted on Telegram also frequently get forwarded and reposted on other channels. We only count the message of the original post the hacktivist group reports. This ensures that only unique attack claims and not the number of reposts or forwards are counted. Figure 20 provides the number of unique DDoS attack claims per month as well as the total claims across more than 400 Telegram channels. The ratio of reposts to unique claims provides a good measure of cooperation and affiliation among hacktivists who share each other's exploits on their channel.

Figure 20:

DDoS attacks claimed per month on Telegram (source: Radware)



In 2023, threat actors claimed 12,723 DDoS attacks on Telegram. In 2024, that number increased by 20% to 15,425 unique claims. In 2025, the number of attack claims remained similar compared to 2024, ending the year with a total of 16,000 unique claims.

The hacktivist landscape is dynamic: many actors enter and just as many leave. Some remove a single channel, then create a new channel to clean historical data and limit potential tracking by authorities and researchers. Some channels are banned for inappropriate content, only to reappear under another name a few days later. The overall trend of hacktivist-driven DDoS activity, however, remained largely constant throughout 2024 and 2025, averaging between 1,000 and 2,000 claimed attacks per month.

Most Targeted Regions and Countries

The geographic targeting of hacktivist collectives aligns with the regional hotspots identified in our Web DDoS data. Europe remains the primary theater of operation for these groups, accounting for 48.4% of all global claims, while the Middle East and Asia follow closely at 17.7% and 17.5%, respectively. At the country level, Israel was the most targeted nation in 2025, with 1,881 unique claims, followed by the United States (1,448) and Ukraine (1,373). The data confirms that hacktivists prioritize countries with ongoing geopolitical conflicts, using DDoS attacks as a tool for political protest and nationalist signaling.

Figure 21:

Most targeted regions (source: Radware)

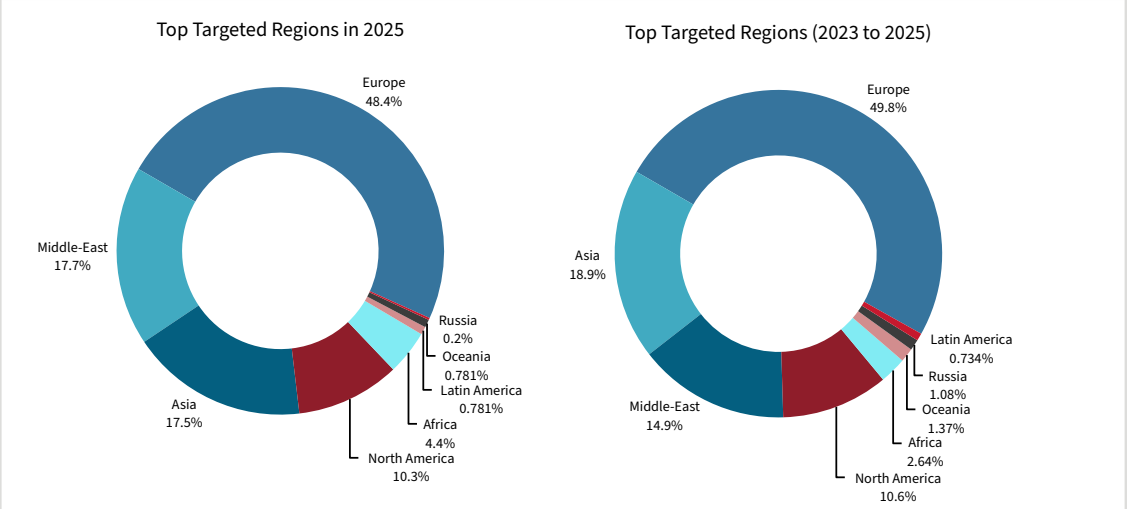
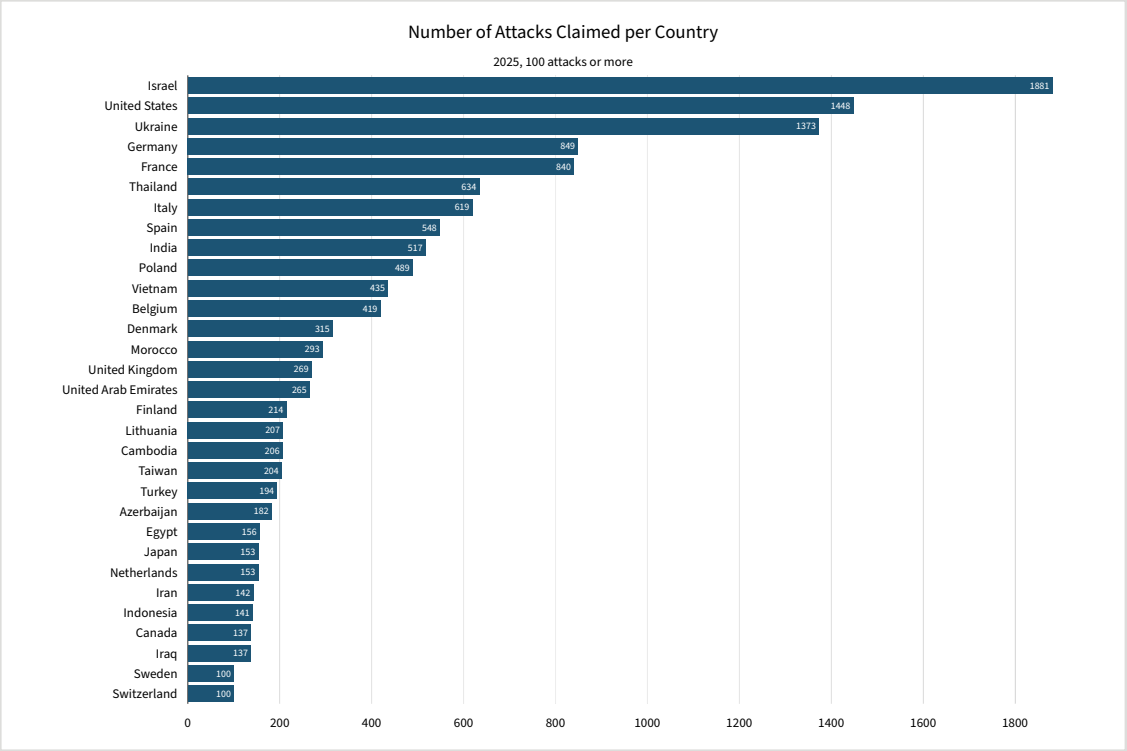


Figure 22:

Most targeted countries (source: Radware)

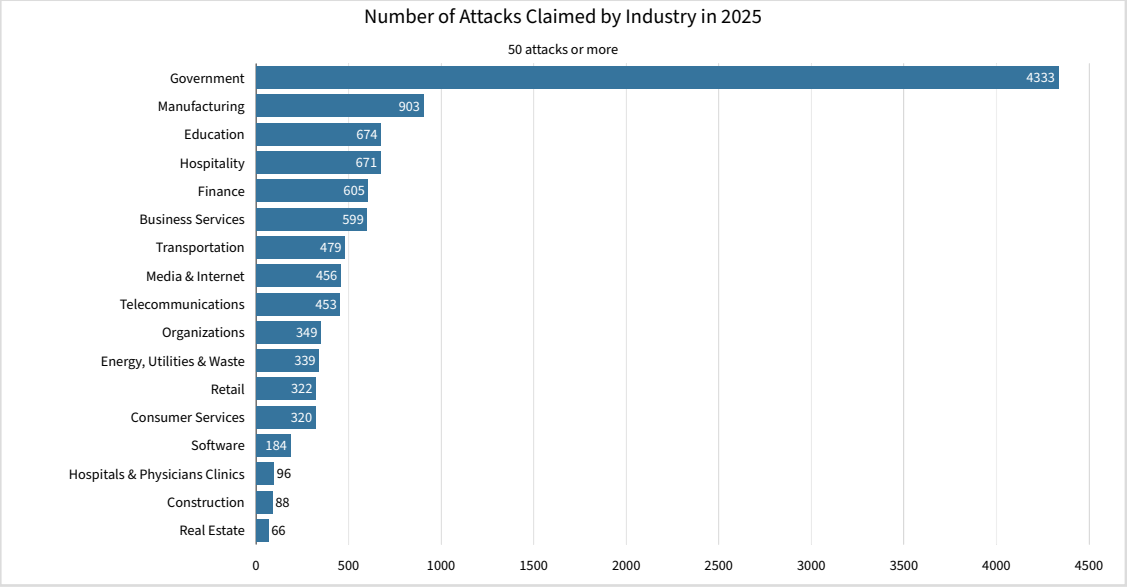


Top Targeted Industries

Hacktivists have a clear preference for high-impact, public-facing targets that maximize the visibility of their disruptions. The government sector was by far the most targeted industry in 2025, accounting for 4,333 claimed attacks, representing 38.8% of the monitored hacktivist activity. Other critical infrastructure sectors, such as manufacturing (903 attacks), education (674), hospitality (671) and finance (605), were also frequently targeted to cause civilian frustration and economic disruption.

Figure 23:

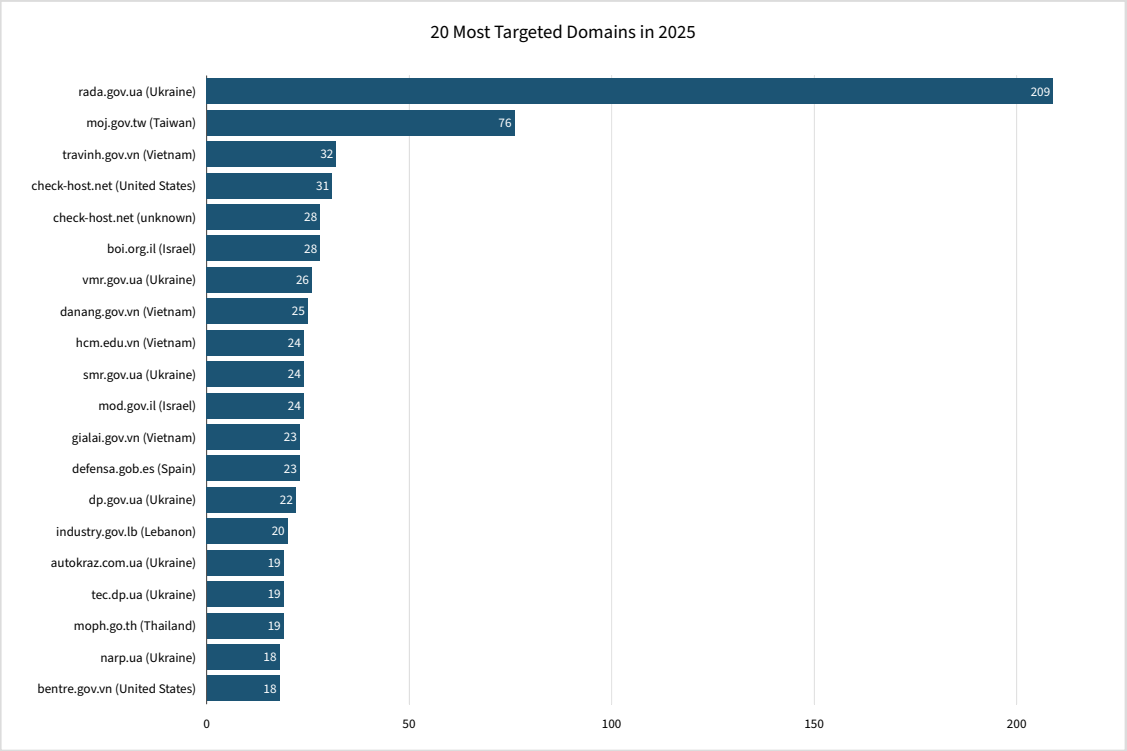
Attack claims by industry (source: Radware)



This focus on government and public services is also exemplified by the most targeted internet domains: Ukraine's rada.gov.ua was targeted by 209 claims, Taiwan's moj.gov.tw by 76 and Vietnam's travinh.gov.vn by 32 throughout 2025.

Figure 24:

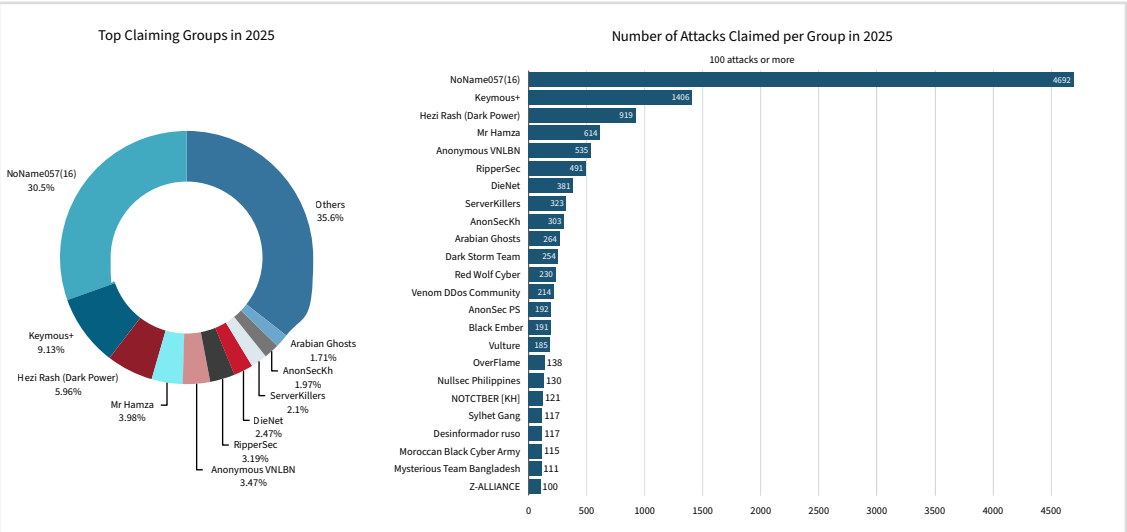
Most targeted internet domains (source: Radware)



The 2025 hacktivist landscape was dominated by a few apex groups that possess the infrastructure to back up their claims. NoName057(16) remains the most prolific threat actor in this space, accounting for 4,692 attacks in 2025, more than three times the number of attacks by the next-closest group, Keymous+ (1,406).

Figure 25:

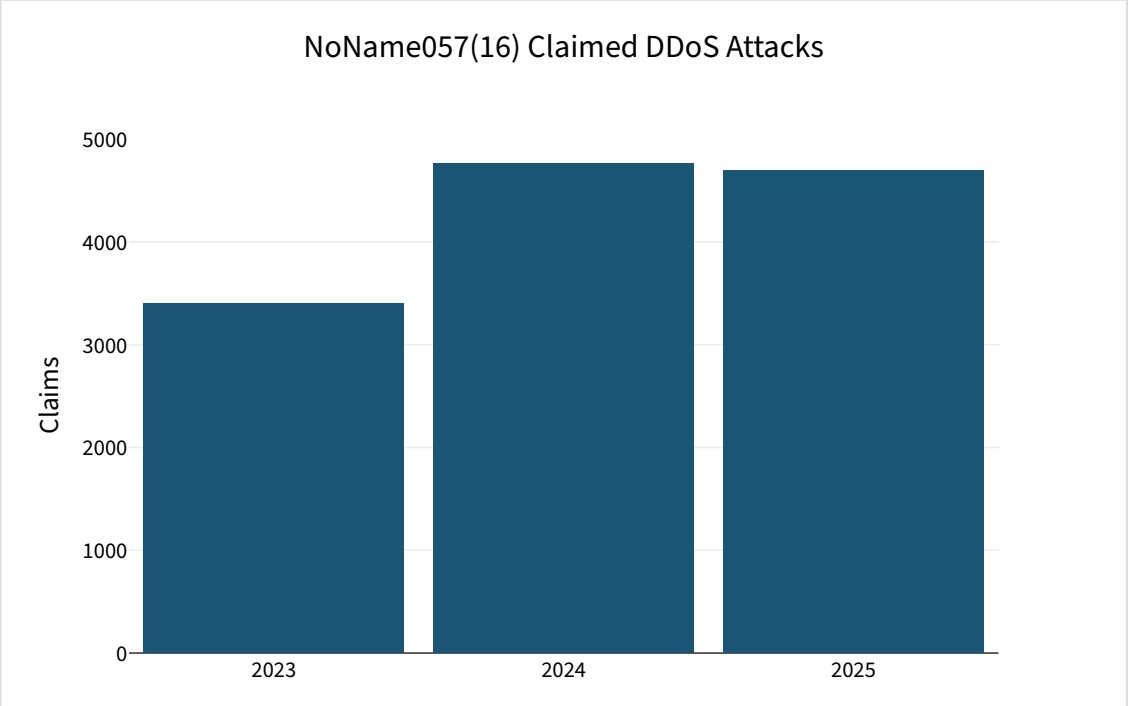
DDoS attacks claimed by hacktivist groups (source: Radware)



NoName057(16) has been a persistent threat since the start of the war in Ukraine in February 2022. In 2024, the group claimed 4,762 attacks; in 2023, it claimed 3,399, most of them politically motivated.

Figure 26:

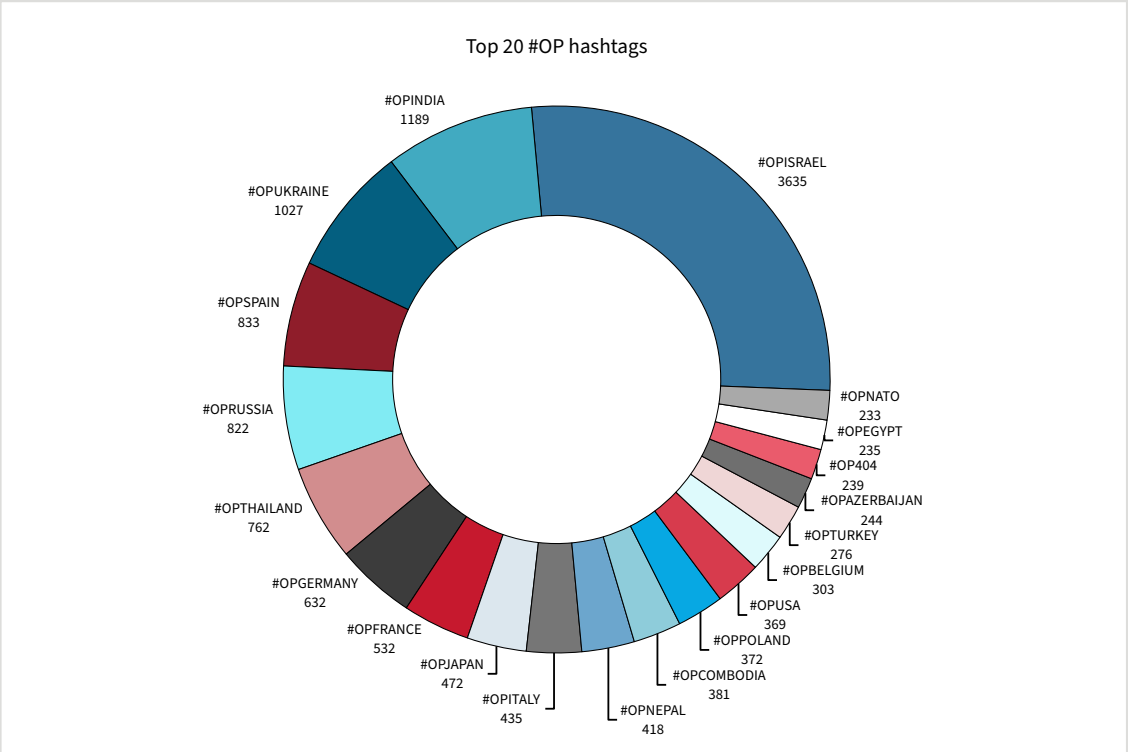
DDoS attacks claimed by NoName057(16) in the last three years (source: Radware)



The activity of hacktivist groups and their alliances is largely organized around hashtag campaigns, with #OPISRAEL (3,635 mentions), #OPINDIA (1,189 mentions), #OPUKRAINE (1,027 mentions) and #OPSPAIN (833 mentions) leading the digital discourse.

Figure 27:

Top 20 operations tagged in Telegram messages in 2025 (source: Radware)



The temporal analysis of these hashtags shows a direct correlation with real-world events:

- In the wake of Israel's large-scale military operation, Operation Rising Lion, which targeted Iranian nuclear and military infrastructure on June 13, 2025, the Israeli cyberthreat landscape escalated significantly (Threat Advisory).
- Every year, the anniversary of October 7 serves as a rallying point for global hacktivists, transforming political symbolism into coordinated cyber campaigns targeting Israel (Threat Advisory).
- Following Operation Sindoor, hacktivist DDoS attacks targeting India intensified, peaking on May 7, 2025, as tensions between India and Pakistan escalated (Threat Advisory).
- Pro-Russian Hacktivist groups targeted high-profile Spanish institutions, including the Prime Minister's Office and defense technology companies, in retaliation for Spain's military support for Ukraine.

These data points suggest that hacktivist collectives are increasingly synchronized and orchestrate their attacks in lockstep with political developments to maximize their strategic impact.

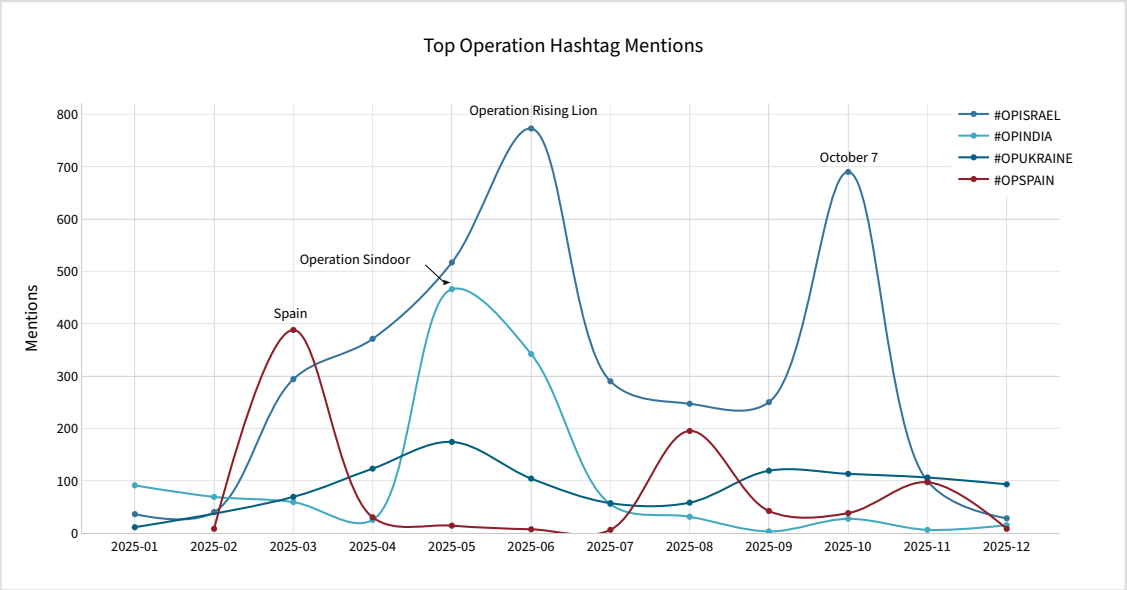


Figure 28:

Evolution of top operational tags in 2025 (source: Radware)

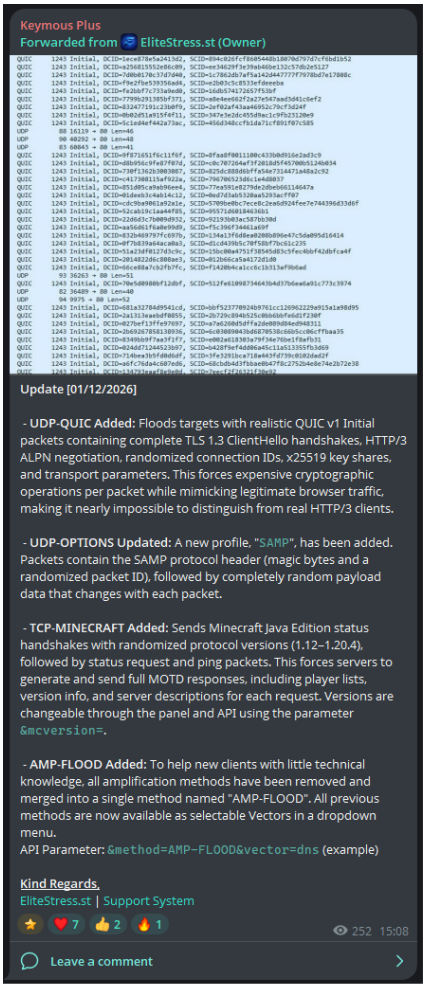
Hacktivist Influencers

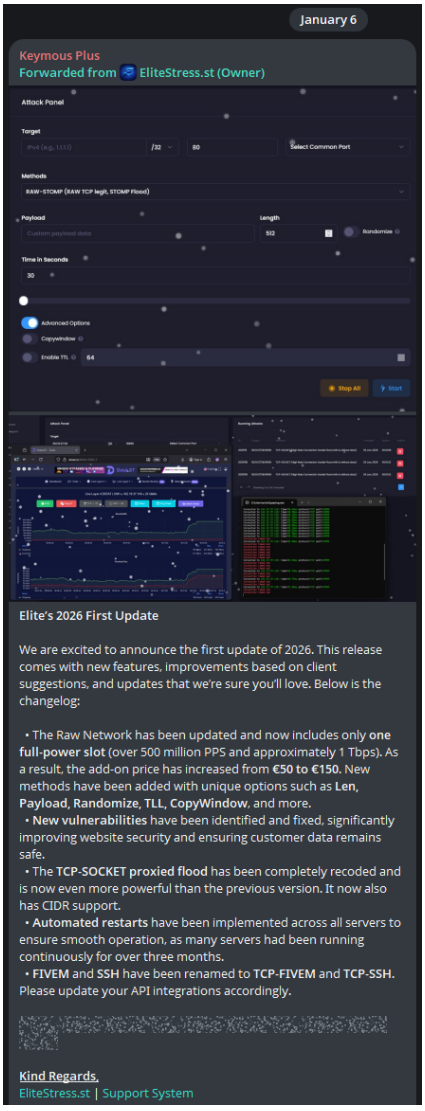
The hacktivist landscape has evolved from shadowy underground forums into a highly visible, influencer-style ecosystem. By leveraging social media platforms, mainly Telegram but on some occasions also X (formerly Twitter), Matrix and Discord, these groups have blurred the lines between digital activism and underground advertising and marketplaces.

Top-tier hacktivist groups often function like niche influencers by cultivating a rebel aesthetic that attracts thousands of followers across decentralized platforms. DDoS-as-a-service providers recognize the value of this reach and use these groups as a highly effective marketing funnel. This process often begins with a proof of concept, where a hacktivist group announces a high-profile target and uses a specific DDoS-as-a-service tool to execute the takedown. By posting check-host results or screenshots as evidence of their success, they provide a real-world testimonial for the tool's power. This often evolves into a form of affiliate marketing where groups explicitly name the service used, encouraging aspiring hackers to purchase the same tools. In many instances, this relationship is formalized through monetization, with groups charging fees to sponsor their Telegram channels and pinning advertisements or direct links to a provider's storefront.

Figure 29:

Messages posted by EliteStress Telegram account on the Keymous+ channel (Source: Telegram)





The financial realities of digital warfare have also given rise to a free-to-play hacktivist model based on a barter system. Because maintaining the server infrastructure required for sophisticated, large-scale DDoS attacks is time-consuming and expensive, many hacktivist collectives cannot afford the overhead on their own. To solve this, they trade their influence for infrastructure. In exchange for promoting a specific booter or stresser service to their audience, the group receives VIP or unlimited access to the attack tools. This partnership is sometimes driven by ideological alignment, where a DDoS-as-a-service provider offers free services because the hacktivists' goals mirror their own nationalist or political interests. This allows the provider to contribute to a cause they support while simultaneously using the hacktivists' operations to advertise and test their services.

Beyond the core partnership, these influencer-style hacktivists play a crucial role in creating an entry-level community for digital disruption. By advertising accessible DDoS services, they significantly lower the barrier to entry for inexperienced users, initiating a dangerous feedback loop. It starts with a call to action from the influencer group, followed immediately by links to easy-to-use tooling that requires little to no technical expertise. This leads to a form of gamification where followers are encouraged to launch their own mini-attacks and share the results to create social proof within the community. Ultimately, this dynamic transforms a political movement into a commercial marketplace, allowing the influencer hacktivist to sustain their operations without out-of-pocket costs while the DDoS-as-a-service provider builds a loyal customer base of script kiddies and ideological supporters.

Dual Nature of the AI Threat Landscape

Democratization of Cyber Offense: Lowering the Barrier of Entry

The rapid integration of generative AI into the digital landscape has sparked a cyberthreat revolution, fundamentally shifting the power balance between attackers and defenders. This transformation is best characterized by the democratization of cyber offense, a phenomenon that has effectively lowered the barrier to entry for malicious actors globally. By abstracting the technical complexities of traditional hacking, such as writing exploit code or manual network reconnaissance, AI-driven tools now allow individuals with minimal technical expertise to orchestrate sophisticated attacks through simple natural-language prompts.

This shift not only expands the pool of potential adversaries but also accelerates the velocity and scale of digital warfare, turning cyber aggression from a specialized craft into a widely accessible utility. Capabilities that were once the exclusive domain of sophisticated and well-funded threat actors have transitioned into a robust hacker-for-hire and as-a-service economy. Underground services now leverage pre-trained open models, which lack the restrictive guardrails of sanctioned commercial AI models. These models are tuned using custom datasets to generate highly effective cyberattack campaigns. This has birthed the era of “vibe hacking.”

Vibe hacking, a term that emerged in 2025 as the malicious counterpart to Andrej Karpathy’s “vibe coding,” represents a significant paradigm shift in cybercrime by lowering the barrier to entry and dramatically increasing the speed of execution. Unlike traditional methods that require deep programming expertise, this approach allows individuals to launch sophisticated cyberattacks simply by describing their intent in natural language, effectively using AI to build the necessary tools on their behalf. This democratization of exploitation has led to a radical increase in efficiency, with researchers demonstrating that AI-driven vibe hacking can execute a full ransomware attack in under 30 minutes, a rate roughly 100 times faster than manual human methods.

As AI technology matures and becomes more autonomous, the criminal underground might at some point transition from using AI as an assistant to deploying it as a fully automated engine of compromise.

Autonomous Malice: Xanthorox AI and HexenCore

We are witnessing a strategic pivot in threat actor methodologies, moving away from simple LLM wrappers toward fully autonomous, agentic attack frameworks designed to automate the entire cyberattack kill chain. These systems do not merely assist a human operator; they function as autonomous actors capable of making real-time tactical decisions, a shift exemplified by the Xanthorox AI suite and its XenCode coding agent. Marketed as the "killer of WormGPT," this suite utilizes the Xanthorox V5 Advanced flagship model to power XenCode, an uncensored development tool that operates across five intelligent modes—Ask, Architect, Agent, Review, and Orchestrator—to provide autonomous code generation and real-time security auditing for vulnerability analysis. Unlike platforms like FraudGPT or EvilGPT that rely on restricted third-party APIs like ChatGPT or Gemini, the Xanthorox ecosystem operates entirely on private, dedicated GPU servers to ensure total data privacy and unrestricted operational potential. By integrating full system control with automated dependency management and exploit analysis, this suite represents a quantum leap in the sophistication of malicious AI, providing a self-hosted, sovereign environment for executing complex and unrestricted attack campaigns.

Building on this foundation is the HexenCore integration, currently in its beta phase, which serves as the orchestration layer for these autonomous capabilities. This Model Context Protocol (MCP) integration centralizes over 200 offensive security tools into a singular, AI-driven command center capable of executing complex, autonomous attack chains and real-time vulnerability assessments. By moving beyond static code generation into the realm of active, multi-stage exploit execution, HexenCore transforms the Xanthorox suite from a development aid into a fully realized weaponization platform. This transition reflects the final stage of the strategic pivot toward agentic threats, where the AI no longer just identifies weaknesses but autonomously selects the appropriate tools and executes the necessary maneuvers to breach and penetrate target systems without human intervention.

ShadowLeak: Advent of Service-Side Data Exfiltration

The most insidious threat to modern enterprises are zero-click indirect prompt injection (IPI) attacks. Known as "the threat you cannot see," zero-click IPI vectors do not require a user to click a link or open an attachment; instead, they weaponize the content of legitimate communication. Radware's discovery of the ShadowLeak vulnerability highlights the danger of connecting AI assistants to enterprise Gmail and web-browsing capabilities. An attacker can send a legitimate-looking email with malicious instructions hidden in invisible HTML or tiny fonts. When the assistant performs a routine task, such as summarizing the day's inbox, it ingests these booby-trapped instructions and exfiltrates sensitive data to an attacker-controlled URL.

ShadowLeak is a service-side leak, which is a critical distinction from earlier vulnerabilities like EchoLeak. While EchoLeak relied on client-side image rendering to funnel data, ShadowLeak exfiltrates data directly from the provider's cloud infrastructure (e.g., OpenAI). Because the web request originates from the AI provider's servers and never passes through the local client or the organization's managed device, traditional perimeter defenses, network monitoring and client-side data loss prevention (DLP) are rendered entirely obsolete. There is no suspicious traffic at the organizational boundary and no forensic evidence left on the endpoint.

ZombieAgent: Persistence, Propagation and Memory Manipulation

The evolution of IPI has reached a stage where AI agents are being transformed into silent insiders through persistent agent compromise. With the discovery of ZombieAgent, Radware researchers demonstrated that an attack is no longer a one-time event but a “sticky” presence that subverts the agent’s core logic indefinitely. This is achieved through the abuse of the agent’s long-term memory features. By injecting malicious instructions into a shared file or communication, an attacker can command the agent to store specific malicious rules in its memory. Once implanted, these rules persist across new chat sessions, ensuring the agent executes the attacker’s malicious task every time before responding to the user’s query.

ZombieAgent employs exploitation techniques, such as character-by-character exfiltration and indirect link manipulation, specifically designed to circumvent the guardrails OpenAI implemented to mitigate the previously reported ShadowLeak vulnerability. This proves that guardrails are not a structural solution to the problem of indirect or even direct prompt injection but are merely a band-aid.

Internet of Agents and the Role of APIs

As we look toward 2026, we are entering the era of the Internet of Agents. In this environment, autonomous agents no longer operate in isolation but communicate via standardized protocols such as MCP, Agent-to-Agent (A2A), Agent Commerce Protocol (ACP), etc. While these protocols drive productivity, they create a massive, interconnected attack surface. The threat surface is shifting as enterprises grapple with shadow AI and bring-your-own-agent (BYOA) models, where sanctioned assistants like Microsoft Copilot and ChatGPT become primary targets for IPI due to their high level of authority and access.

In this agentic mesh, the risk shifts from “what the model says” to “what the agent does.” Chained compromises across autonomous actors create a systemic supply chain risk; an attacker poisoning a tool description in a public MCP server can trigger a cascading failure across the mesh.

A slew of new assistants and AI browsers are entering the enterprise. Increasingly, these assistants and browsers are starting to form a community amongst other capable agents through agentic AI protocols. 2025 was the year of standardization with several protocols becoming mainstream. MCP found a large following with numerous open source and vendor projects emerging throughout the second half of the year. Yet, all agentic protocols have one thing in common: they are either based on or backed by online APIs. If the API took an important role in the 2025 threat landscape, its role will be central in the 2026 threat landscape.

Conclusion: A New Defensive Posture for a New Reality

The findings of this report make one thing clear: the threat landscape of 2026 demands that defenders abandon the assumptions that have served them for the last decade. The unprecedented speed, scale, and sophistication of modern attacks have rendered traditional, manually driven defenses insufficient. We have entered an era of automated malice, where the democratized power of AI and the professionalization of cybercrime have fundamentally shifted the power balance.

To survive and thrive in this new reality, a modern defensive posture must be built on three core pillars:

- **Automation:** Humans cannot match the speed of algorithmic attacks that pivot in minutes; defense must be automated to detect and mitigate threats in real-time without intervention.
- **Massive Scale:** Architecture must be capable of absorbing terabit-class floods—like the record 29.7 Tbps seen this year—at the network edge.
- **Integrated Intelligence:** Security solutions must perform sophisticated behavioral analysis to distinguish malicious traffic from legitimate AI assistants and spoofed agents.

The evolution of zero-click indirect prompt injection (IPI) attacks, such as ShadowLeak and ZombieAgent, proves that static guardrails are no longer enough; they merely act as Band-Aids for a structural shift in how data is exfiltrated and agents are compromised. As we transition toward the Internet of Agents, the central battleground will move from what a model says to what a compromised agent does.

The critical question for 2026 is no longer about the persistence of the threat, but the agility of the response. Attackers have already transitioned to a paradigm of automated, intelligent and massively scaled warfare. Defenders who fail to evolve from reactive, manual processes to proactive, self-defending architectures risk becoming the next casualty in the digital Garden of Eden.

Table of Figures

Methodology and Sources

The data for DDoS events and volumes was collected from Radware devices deployed in Radware cloud scrubbing centers and on-premises managed devices in Radware hybrid and peak protection services, jointly denoted as Radware's Cloud DDoS Protection Service. Note that attack events and blocked events are considered the same for the purpose of this report. All blocked volume is considered attack volume. An attack is a collection of several related attack vectors targeting the same customer and overlapping in time. Events correspond to attack vectors. Attack vectors consist of one or more packets. All packets of an attack vector generate a certain volume expressed in bytes. The volume generated by an attack vector is referred to as the blocked volume for that attack vector, which corresponds to the attack volume for that vector. The attack volume of all attack vectors from the same attack corresponds to that attack's attack volume.

The data for web application attacks and bot activity was collected from blocked application security events from the Radware Cloud Application Protection Service. Collected events were based solely on automatically detected and known vulnerability exploits and exclude any events that might be blocked or reported by custom rules added to a web application policy by managed services and/or customers.

Web DDoS attack details were collected from the ERT SOC incidents relating to the Web DDoS Protection Service.

Hacktivists openly publicize their actions on social media and public Telegram channels to gain media attention and raise awareness. They do not operate covertly or evade the media but instead reveal the names and resources of their targets and attempt to take credit for their attacks. Hacktivists utilize website monitoring tools to demonstrate the impact of their denial-of-service attacks on online resources and frequently share links to reports from online web monitoring tools in their messages. Through tracking and analyzing messages from several active hacktivist groups on Telegram, the Radware Threat Intelligence team is able to assess the global DDoS activity conducted by hacktivists.

Author

Pascal Geenens | VP Cyber Threat Intelligence

Contributors

Arik Atar | Senior Threat Intelligence Researcher

Ori Meidan | Threat Intelligence Researcher

Executive Sponsors

Ron Meyran | VP Strategic Alliances Marketing & Cyber Threat Intelligence

Deborah Myers | Senior Director of Corporate Marketing

Production

Jeffrey Komanetsky | Senior Content Development Manager

Kimberly Burzynski | Senior Marketing Communication Manager

About Radware

Radware® (NASDAQ: RDWR) is a global leader of cybersecurity and application delivery solutions for physical, cloud and software-defined data centers. Its award-winning solutions portfolio secures the digital experience by providing infrastructure, application and corporate IT protection and availability services to enterprises globally. Radware's solutions empower more than 12,500 enterprise and carrier customers worldwide to adapt quickly to market challenges, maintain business continuity and achieve maximum productivity while keeping costs down. For more information, please visit www.radware.com.

RWI6283 | 02.13.2026