fastly

AI, Bots, and the Agentic Future of the Web

How bot insight and intent are transforming the way businesses understand and manage automated traffic

Introduction

With 2026 well underway, bots – AI, agents, automation (wanted and unwanted) – are increasingly disruptive to security, reliability, and entire operating strategies. While traditional AppSec challenges could often be solved with a block or allow, bots and agents require an unprecedented level of nuance that forces organizations to develop new strategies tailored to this type of traffic.

Bot traffic requires more than a broad 'is it a bot' assessment, instead demanding strategic decisioning at the business level to determine which traffic is truly wanted, what risks it poses, and what content and apps that traffic *can* and *should* access. While bot management and other security solutions offer a level of insight into whether a request is from a bot and, if so, what kind, pairing that insight with the bot's *intent* offers critical context needed to inform business strategy.

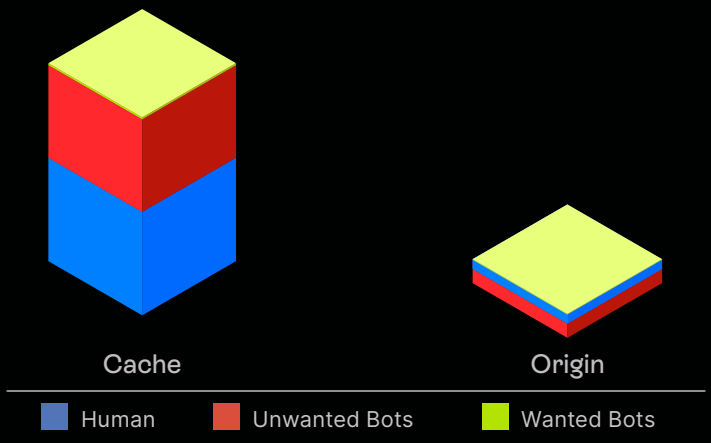
In light of this dynamic, we focused this edition of Fastly's Threat Insights Report on pairing the insights yielded by a bot management solution with bot intent – is the request for an application's most popular (cached) assets or APIs, or not (to origin)? Each has its own considerations, implications, and direct impacts on business operations, as detailed in this report.

The saying "adapt or fall behind" is apt here. As we watch AI change the way we do business online, and the very internet itself, those who succeed in this new "AI era" will be those who adapt, strategically.

FASTLY THREAT INSIGHTS REPORT

Visual Overview of Findings

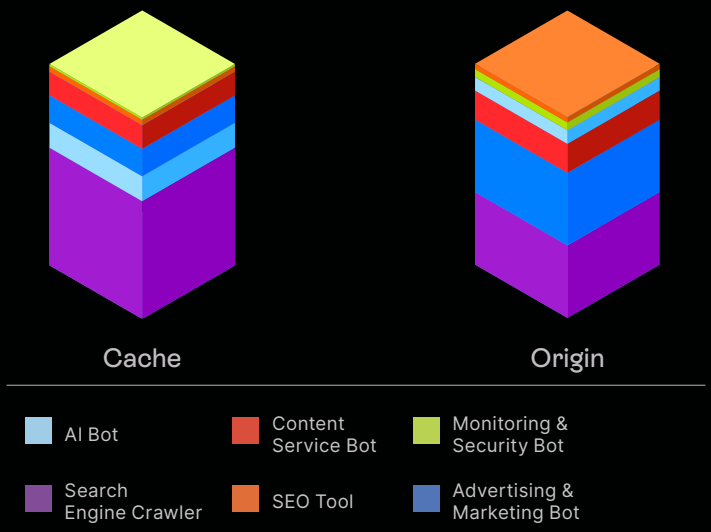
Bots vs. Humans in 2026



The Global Baseline Received

- **51%** human traffic
- **48%** unwanted bot traffic
- **1%** wanted bot traffic

Wanted Bots by Type

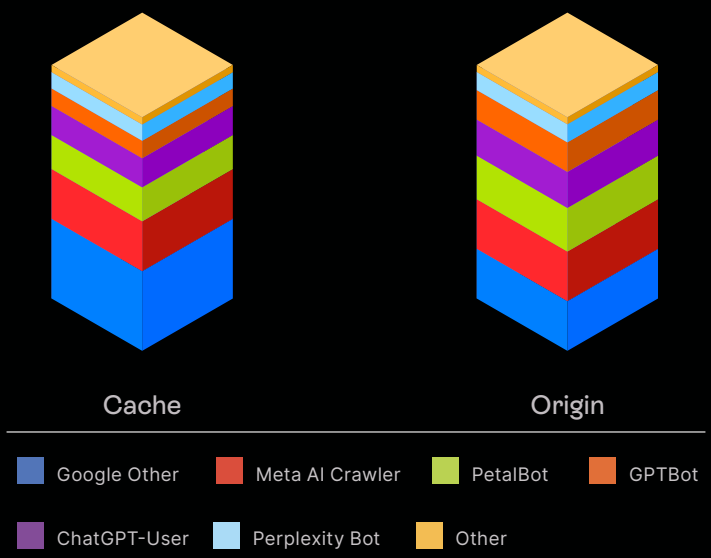


Wanted Bot Traffic

- **8%** AI-driven bots
- **92%** Legacy bots

Of the wanted bot traffic, 8% is AI-driven. The remainder originate from legacy bots scouring the web for search engines, ads, images and so on.

Top AI Bots



Top AI Crawlers

- **28%** Google Other
- **21%** Meta AI Crawler
- **17%** PetalBot

Traffic Trends

In January 2026, **49% of total requests originated from bots, and 51% of requests came from humans.** Pairing this finding with the additional layer of traffic 'intent' provides further insights to inform strategy.

TL;DR

- Total bot and human traffic are nearly equivalent (49 and 51% respectively)
- 47% of cached content requests are from bots → do you want them accessing your most popular content? Any bot? Just specific ones?
- 60% of origin requests are from bots → do you want them inflating egress costs and looking at content that may be irrelevant?

Cached vs origin traffic trends

Modern organizations' most popular content at any given time is often served from a content delivery network's cache to support the increasingly global economy and enhance their end users' experience. When we examine the bot traffic hitting cached content, we find 47% is from bots and 53% is from humans.

While organizations have historically paid little attention to who accesses their cached content, since it costs less than origin egress, as more of the world adopts automation, the high percentage of bots raises major concerns. Bots accessing cached content represent competitive intelligence, data gathering of your most popular content, and even nefarious activities. Deeper visibility into cached content is imperative to operating margins and overall strategy, so your most frequently viewed content and who can leverage it remain within your control.

Origin requests account for a fraction of all requests. In January, we noted 60% of traffic to origin came from bots, and 40% of traffic to origin came from humans. These percentages are to be expected, as bots crawl content indiscriminately, often without any context for what is new, relevant or popular content, and organizations likely have a surplus of older content from operating over many years.

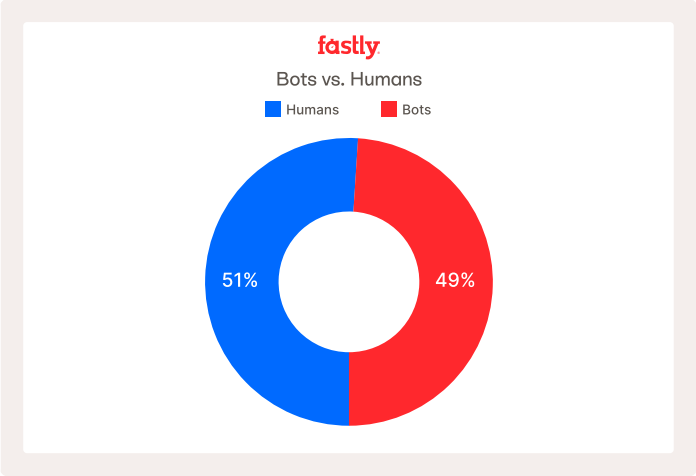


Figure 1. Bots vs. Humans in January 2026. To get a sense for scale, the split is made up of trillions of requests.

However, deeper visibility into origin content is imperative to determine whether it's worth incurring the additional egress costs for bot requests.

While bot traffic in cached and origin content each has its own concerns and considerations, their sheer volume forces organizations to go beyond simply acknowledging that bots are part of their traffic. You must understand who these bots are, why they're accessing their content, their intent (cache or origin), and whether each is permissible.

We examine this through the lens of wanted vs. unwanted bots (as defined in detail in the appendix) to provide examples of how this granular insight can influence strategic decisions.

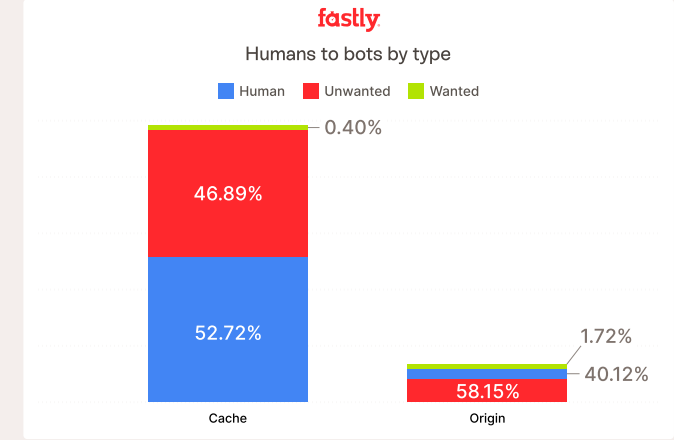


Figure 2. Human to unwanted and wanted bot traffic hitting the cache or going to origin.

The State of Unwanted Bots

Given the high volume of bot traffic, we conducted a deep dive into the types of bots and found that 99% of total bot requests (to both cache and origin) originate from unwanted bots! These are bots that are common types of unverifiable automation, like API clients or those with malicious fingerprints, deceitful user-agents, are malicious scanners, etc. These are bots that, at times, pretend to be real users and can be there for nefarious purposes.

The impact of unwanted bots on business

The inability to distinguish between wanted and unwanted bots leaves organizations guessing or making false assumptions about the nature and makeup of their bot traffic.

This matters at a strategic level; unwanted bots carry out nefarious activities under the guise of legitimate traffic – price scraping and intelligence from competitors, cybercriminals probing for vulnerabilities, or even the automation of malicious activity as commonly seen in account takeover attacks (ATO). Their impacts infringe on how organizations compete in the market, secure their apps and APIs, and generate revenue.

Consider an unwanted bot falsely declaring itself to be ChatGPT; an organization might base its bot blocking and allowance policies on a presumed “ChatGPT problem” that is in fact, a malicious bot problem. This exemplifies both strategic and operational risk that many organizations face without trusted insights, as their strategy is misinformed by false insights, and the malicious bot is likely to be granted unwanted access.

TL;DR

- 99% of total bot requests are unwanted → is that traffic you want?
- Unwanted bots come with consequences – competitors gathering intelligence, attackers using automation, etc.
- Lack of trusted bot traffic insight leads to business decisions made on assumptions rather than realities.

The State of Wanted Bots

Despite the increasing shift toward AI-driven content consumption, in our examination of wanted bot activity, we found that only 1 out of 100 bot requests comes from a verified, identifiable source like AI.

Figure 3 highlights how the vast majority of wanted bot requests, for both origin and cached content, originate from legacy bots scouring the web for search engines, ads, images, and so on. Many organizations have spent years optimizing for these bots already (think classic SEO content and keyword strategies), but must now adjust for the growing impact of AI bots.

AI bots currently represent about 8% of total wanted bot traffic, but this type of bot can have an outsized impact on your business. We are witnessing AI actively reshape entire industries, with digital publishing serving as a prime example. One crawl of a publisher's site means the valuable content can now be served directly from an LLM, and users might never access the source website to gather information. This results in a punishing reality where publishers' very way of doing business could be greatly impacted by AI.

All industries must factor AI into their strategy or risk long-term impact; allowing scraping of inaccurate or outdated content can dilute the value of intellectual property, surface compliance risks, and present a misleading or weakened brand image and reputation. Managing how bots interact with content isn't just a technical concern; it's a governance, security, and brand imperative.

This reinforces the need for organizations to understand what content AI is accessing, and whether that is actually a desired activity and outcome.

AI bots can be broken into crawlers and fetchers (see appendix for definitions) - the following examines trends in each.

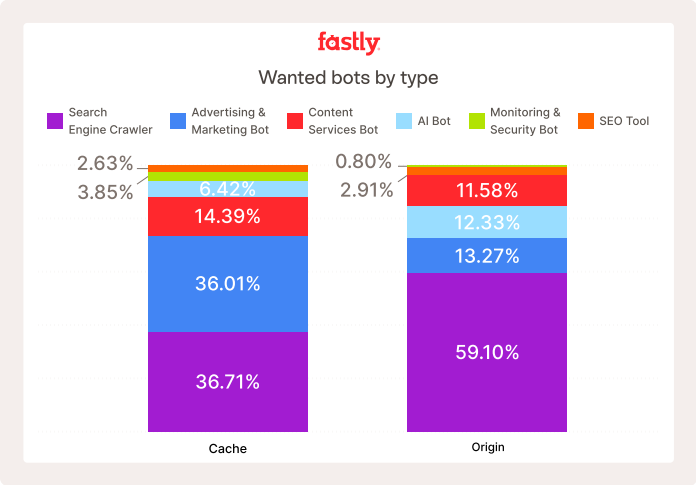


Figure 3. Wanted bot requests to origin and to cached content as a percentage of total.

AI Crawler Traffic

Assessing AI crawler traffic, crawling is primarily attributable to Google Other (36%), Meta's AI Crawler (28%), and Huawei's PetalBot (18%).

Examining this data through the lens of daily crawler activity in January, we see no discernible trend- underscoring that crawlers operate indiscriminately, spidering out to all pages possible after the first "/" in a URL without regard for what's popular or relevant. The similarities between cached and origin traffic per bot also reinforce the theory of indiscriminate crawling.

Crawlers represent a tradeoff; your data being gathered indiscriminately means you have limited control over the content used to represent your business - but blocking these crawlers completely means you might have zero business representation at all.

The challenge in managing crawlers is twofold:

- **Cached** – your most popular content is being ingested for future usage. While users may not have requested it yet, it's likely it will get served in a user's query given its temporal relevance as part of the cache.
- **Origin** – content that is new and yet to be cached or less sought after and may not represent current realities (outdated data), brand values, etc.

As organizations consider the AI they actually want on their apps and APIs, knowing exactly which crawler bots are on your services and what they're accessing enables granular decisions around how policies are created – allowing, rate limiting, or even monetizing others, while some may get outright blocked.

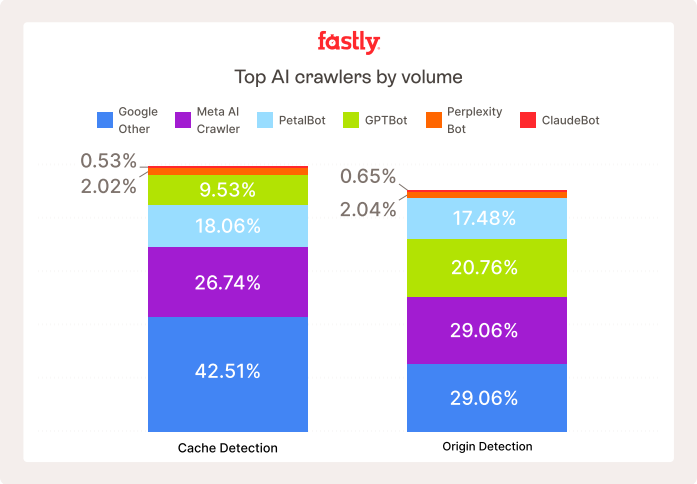


Figure 4. Type of wanted bot requests by total volume and represented as a percent of the total.

AI Fetcher Traffic

Fetching is primarily attributable to OpenAI's ChatGPT-User & SearchBot (97%) and DuckAssistBot (2%). Since fetching entails AI bots grabbing content in response to human queries, it is interesting to see that 57% of fetcher requests are predominantly fetching content that cannot or has not yet been cached. While we can't say for certain, one possible reason is that real users are looking for hyper-recent or very specific dynamic content (weather information, game scores, election results, etc.) that hasn't been crawled but serves an immediate, real query.

Given that fetchers hit your apps and APIs in response to direct human interaction, these bots should be factored most heavily into your AI strategy.

Not allowing them directly influences whether your business is surfaced in the LLMs' responses, which, for industries like Publishing, may be permissible as they want to drive that traffic directly to their website instead of serving it in a response. For others, a needed boost in visibility for later conversion may be desirable, so allowing fetchers is part of their business strategy.

Today, most fetching is tied to OpenAI, but as AI continues to take center stage and adoption grows worldwide, we may see fetcher bots grow considerably in volume, forcing additional considerations about whether the category, holistically, is permissible or only some.

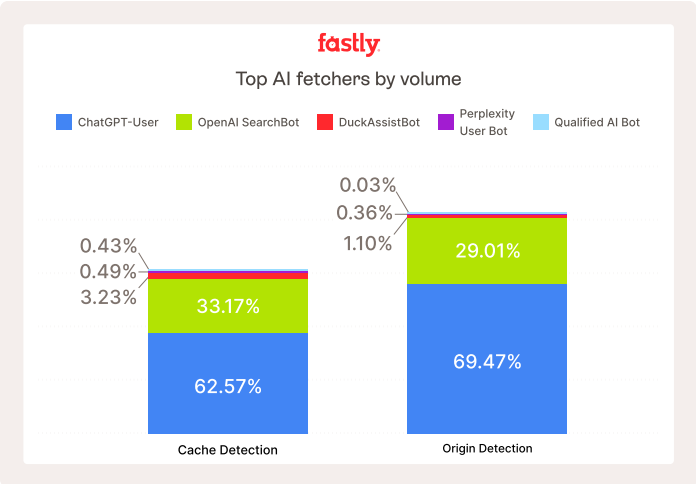


Figure 5. AI Fetcher requests to cache and origin by total volume and represented as percent of total.

TL;DR

- Only 1% of bots are verifiable and while much of the world is putting emphasis on AI, it represents 8% of wanted bot traffic, but with outsized impacts per request
- Crawling is primarily Google, Meta, and Huawei bots – their indiscriminate crawling forces organizations to assess whether their most popular or even outdated content can be scraped to support a potential future query
- Fetching is primarily attributable to two OpenAI bots (98%) – their fetching in response to real users forces organizations to weigh visibility and mindshare vs. content authority during pivotal moments

Bots are interacting with your content on a massive scale - so what can you do about it?

The best thing organizations can do to adapt to the new AI internet is evolve with the growing scale of AI bot traffic. This entails a number of strategies.

Gain visibility

If you take one finding away from this report, it should be that organizations must gather powerful insights into their bot traffic to inform future strategic decisions. It is no longer enough to simply accept that bots are on your services without seeking additional granularity. Nearly half of all requests hitting apps and APIs are bot traffic, making ignoring them highly consequential.

Businesses must strive to gather granular insights to the level of individual bots on their services - only with this depth of visibility can policies be created around which bots receive unique treatment.

Go beyond blocking

Unwanted bots represent 99% of all bot traffic, forcing organizations to combat their scale with capabilities purpose-built to solve them. Fastly's patent-pending Deception technology enables organizations to shift the power back into the defenders' hands by confusing attackers and impacting their economies of scale.

Deception's ATO use-case returns false "invalid" responses to attackers even if the submitted credentials were valid, impeding their ability to retool and forcing them to go elsewhere. This is just the first use-case of this innovative capability, and we're planning to release even more very soon!

Experiment with bot monetization

Wanted bots represent just 1% of all bot traffic, yet they have an outsized impact on many businesses. Monetization is an emerging approach that's gained traction over the last year for generating revenue from sanctioned AI access.

In 2025, AI content transactions facilitated through one of Fastly's monetization partners, TollBit, grew 32x as publishers moved from experimentation to active monetization. Hundreds of publishers are now receiving payouts as AI companies increasingly seek sanctioned access to publisher content for retrieval.

Conclusion

Bots continue to reshape the internet as we know it - organizations must establish business-level strategies that consider their nuanced implications, and adapt accordingly.

The AI era demands that organizations pair deep bot visibility with 'intent' - the purpose of bot traffic, what content it is accessing, and where. This depth of insight comes from integrated platform solutions capable of providing this depth of visibility. Combining insight from solutions like Fastly Bot Management, a Content Delivery Network, and even other AppSec products enables the powerful bot control necessary to address the insights and outcomes detailed in this report.

With capabilities purpose-built to combat malicious bots like Deception and the ability to protect cached content with ContentGuard, organizations gain the visibility, capability, and flexibility needed to unlock their AI and bot strategy. If you have a bot problem and need support, contact us.

Industry and Regional Insights

In the following section you'll find visual summaries for commonly requested industries and regions. If you find yours omitted, reach out and our team is happy to share.

Industries

- Publishing
- Ecommerce
- Travel
- Advertising Technology
- SaaS/PaaS

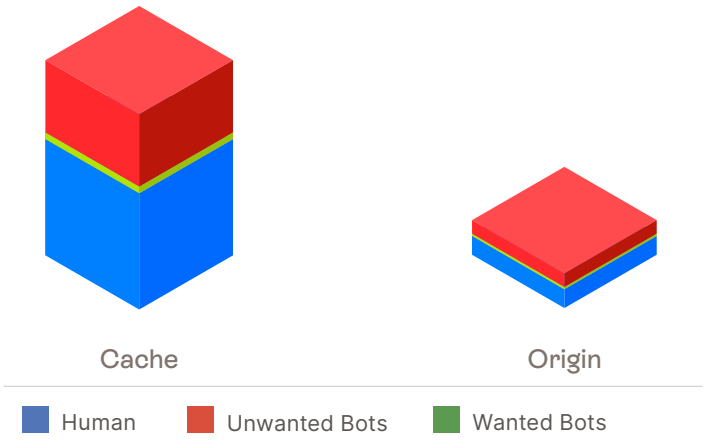
Regions

- JAPAC
- EMEA
- North America
- LATAM

FASTLY THREAT INSIGHTS REPORT

Publishing

Bots vs. Humans in 2026



The publishing industry received

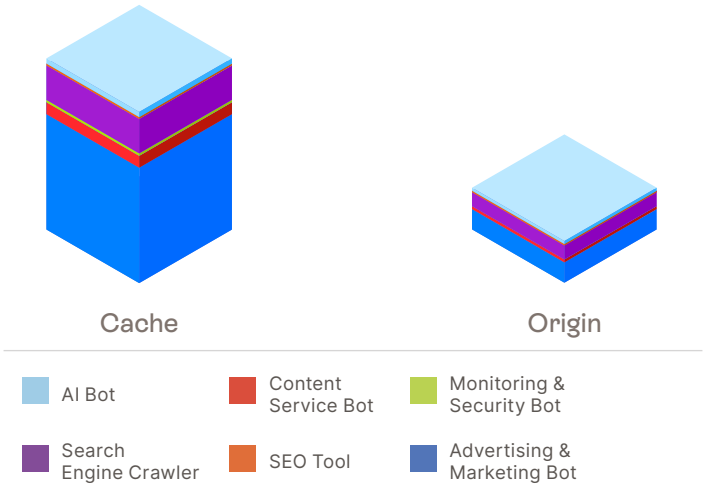
14% more human traffic

22% less unwanted bot traffic

587% more wanted bot traffic

All figures compared to global baseline.

Wanted Bots by Type



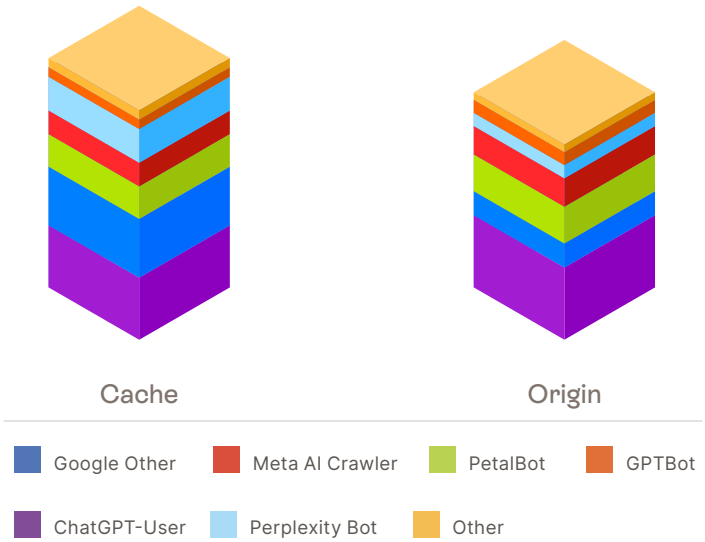
Wanted Bot Traffic

4% AI-driven bots

48% less than global baseline

4% of the publishing industry's wanted bot traffic is attributable to AI. This is 48% less than the global baseline.

Top AI Bots



Top AI Crawlers

32% ChatGPT User

20% Google Other

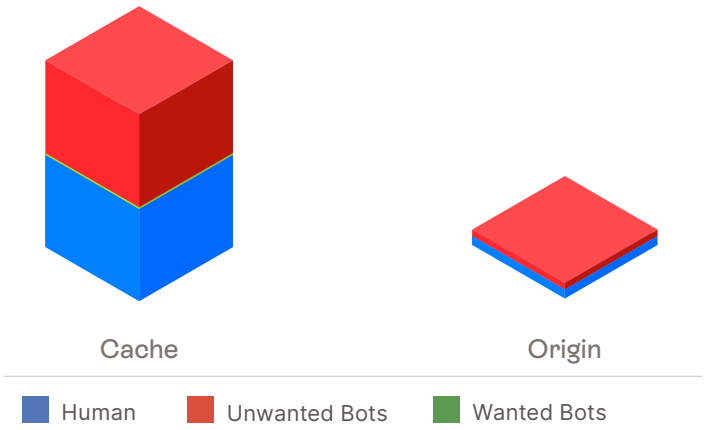
16% PetalBot

*The global baseline = the % across all observed regions and industries in Fastly's total traffic.

FASTLY THREAT INSIGHTS REPORT

E-commerce

Bots vs. Humans in 2026



The e-commerce industry received

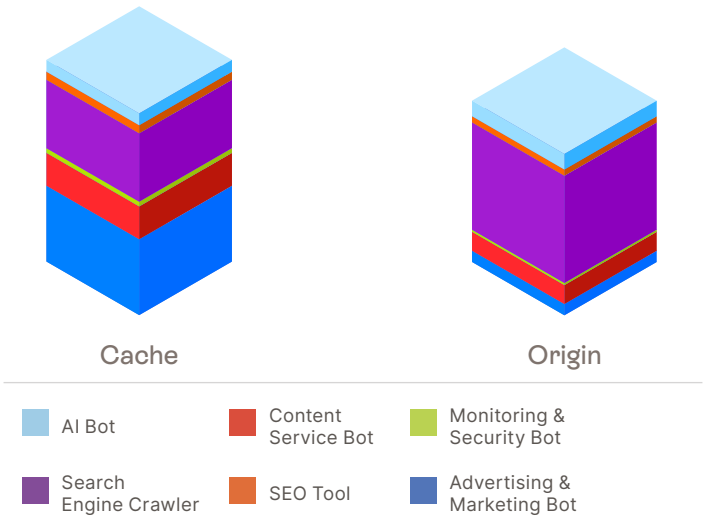
5% more human traffic

5% less unwanted bot traffic

43% more wanted bot traffic

All figures compared to global baseline.

Wanted Bots by Type



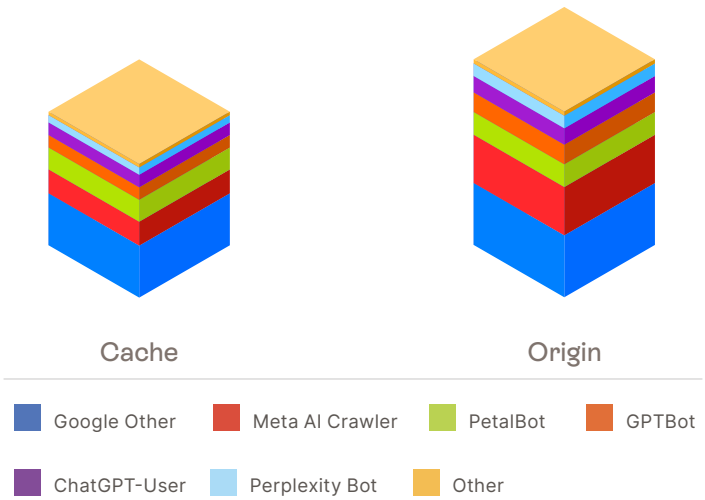
Wanted Bot Traffic

8% AI-driven bots

7% less than global baseline

8% of the e-commerce industry's wanted bot traffic is attributable to AI. This is 7% less than the global baseline.

Top AI Bots



Top AI Crawlers

36% Google Other

22% Meta AI Crawler

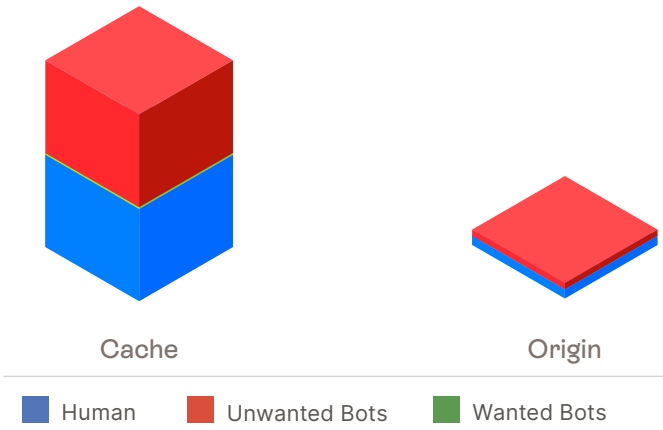
14% PetalBot

*The global baseline = the % across all observed regions and industries in Fastly's total traffic.

FASTLY THREAT INSIGHTS REPORT

Travel

Bots vs. Humans in 2026



The travel industry received

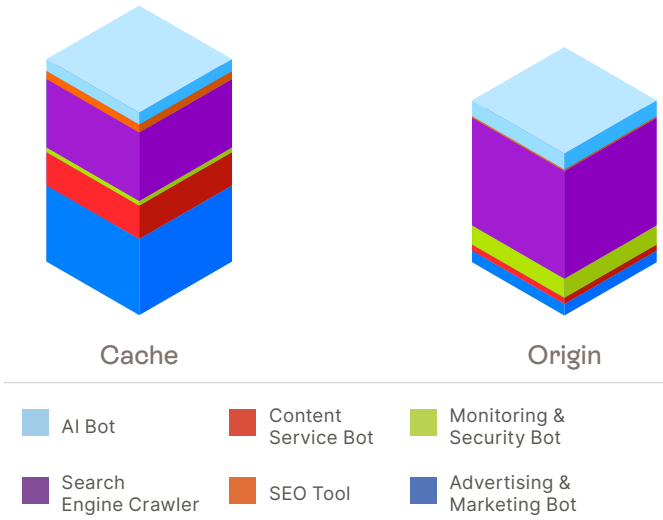
46% more human traffic

54% less unwanted bot traffic

420% more wanted bot traffic

All figures compared to global baseline.

Wanted Bots by Type



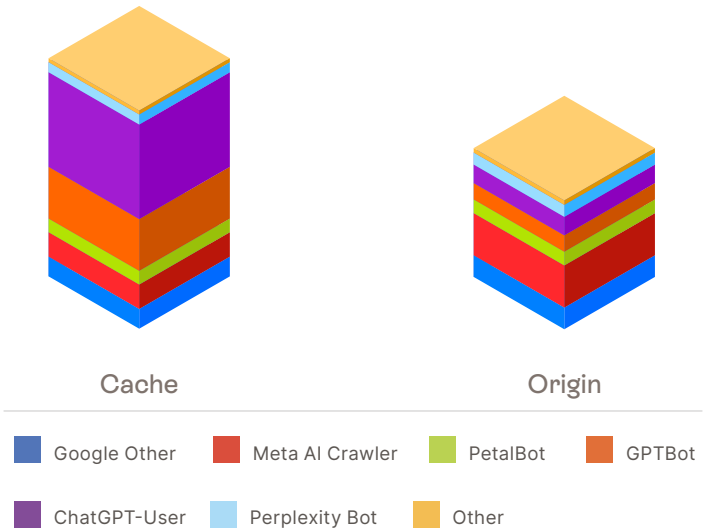
Wanted Bot Traffic

14% AI-driven bots

70% more than global baseline

14% of the travel industry's wanted bot traffic is attributable to AI. This is 70% more than the global baseline.

Top AI Bots



Top AI Crawlers

33% GPT Bot

20% ChatGPT User

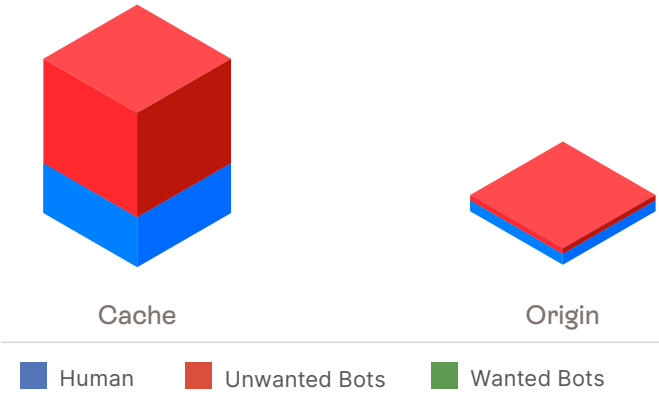
19% Meta AI Crawler

*The global baseline = the % across all observed regions and industries in Fastly's total traffic.

FASTLY THREAT INSIGHTS REPORT

Advertisement Technology

Bots vs. Humans in 2026



The AdTech industry received

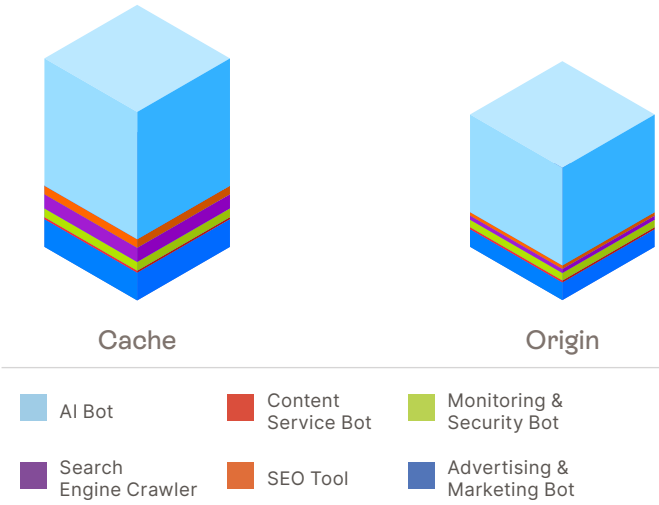
31% less human traffic

33% more unwanted bot traffic

60% less wanted bot traffic

All figures compared to global baseline.

Wanted Bots by Type



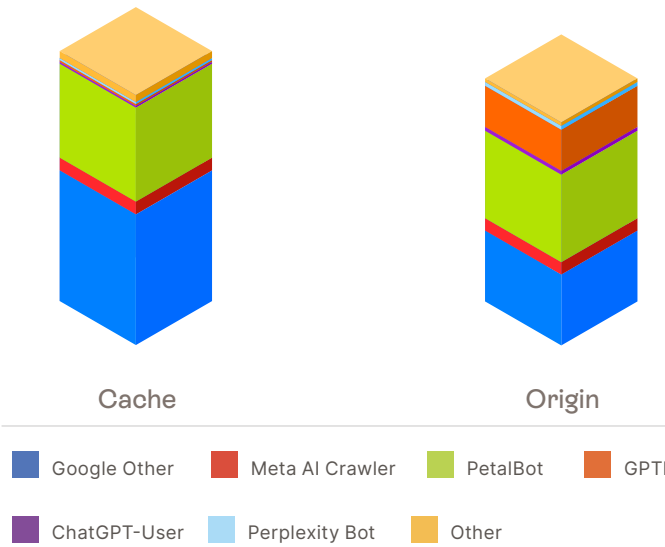
Wanted Bot Traffic

5% AI-driven bots

42% less than global baseline

5% of the AdTech industry's wanted bot traffic is attributable to AI. This is 42% less than the global baseline.

Top AI Bots



Top AI Crawlers

43% Google Other

39% Petalbot

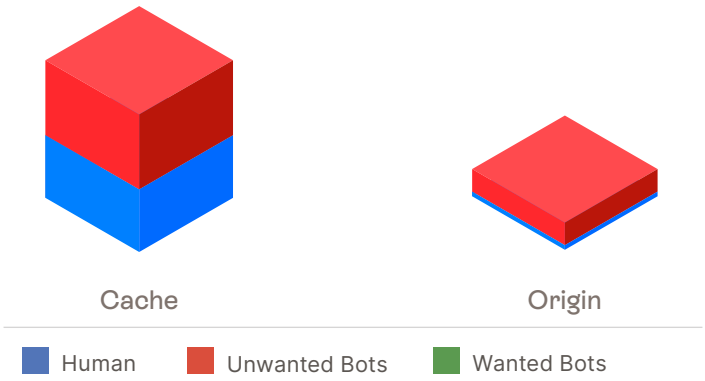
9% GPTBot

*The global baseline = the % across all observed regions and industries in Fastly's total traffic.

FASTLY THREAT INSIGHTS REPORT

SaaS/PaaS (High Tech)

Bots vs. Humans in 2026



The High Tech industry received

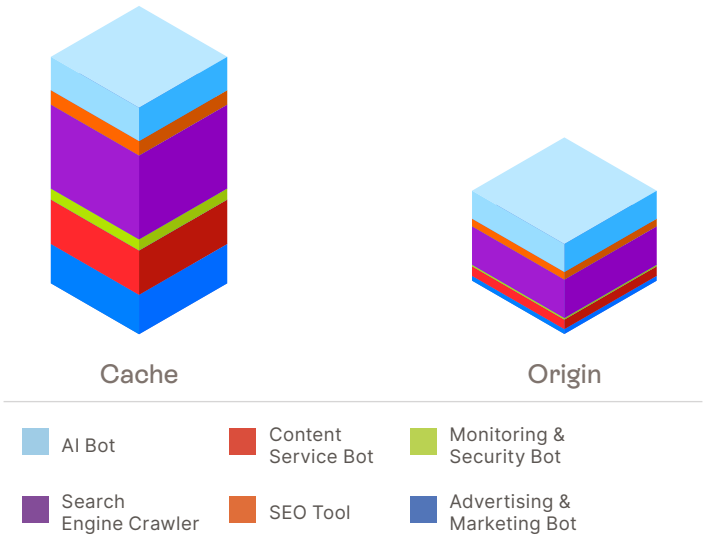
21% less human traffic

22% more unwanted bot traffic

4% less wanted bot traffic

All figures compared to global baseline.

Wanted Bots by Type



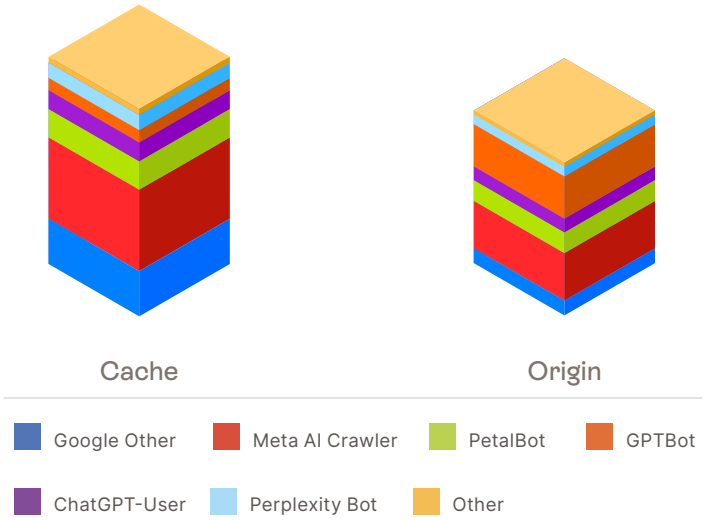
Wanted Bot Traffic

20% AI-driven bots

136% more than global baseline

20% of the High Tech industry's wanted bot traffic is attributable to AI. This is 136% more than the global baseline.

Top AI Bots



Top AI Crawlers

35% Meta AI Crawler

17% Google Other

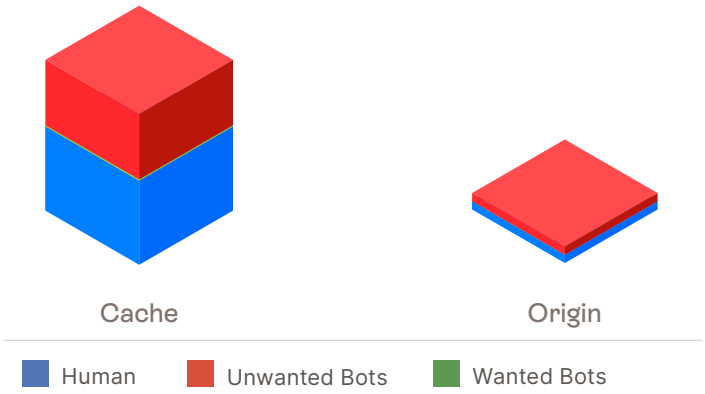
16% GPTBot

*The global baseline = the % across all observed regions and industries in Fastly's total traffic.

FASTLY THREAT INSIGHTS REPORT

North America (NA)

Bots vs. Humans in 2026



The NA region received

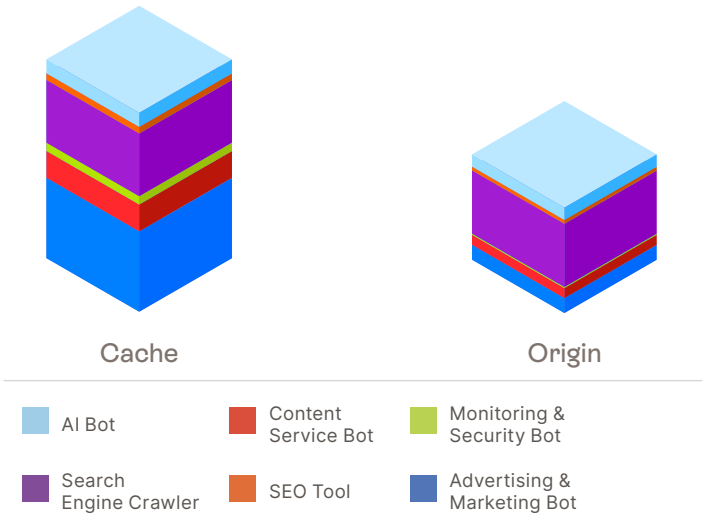
7% more human traffic

8% less unwanted bot traffic

11% more wanted bot traffic

All figures compared to global baseline.

Wanted Bots by Type



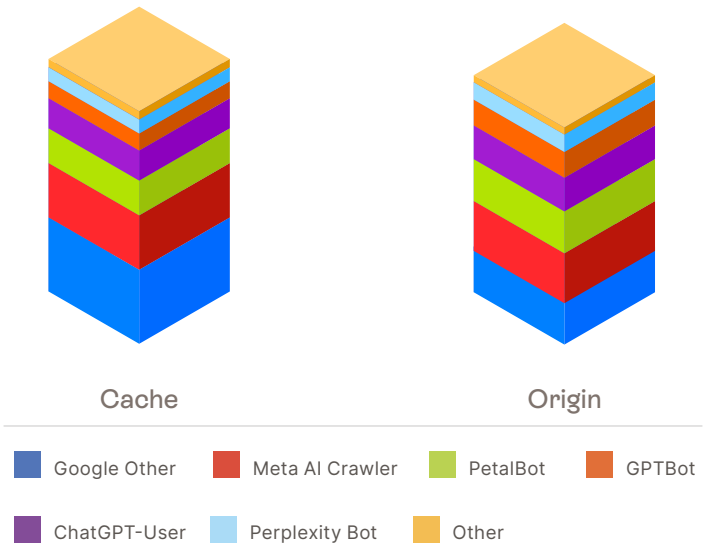
Wanted Bot Traffic

8% AI-driven bots

4% less than global baseline

8% of the NA region's wanted bot traffic is attributable to AI. This is 4% less than the global baseline.

Top AI Bots



Top AI Crawlers

26% Google Other

23% Meta AI Crawler

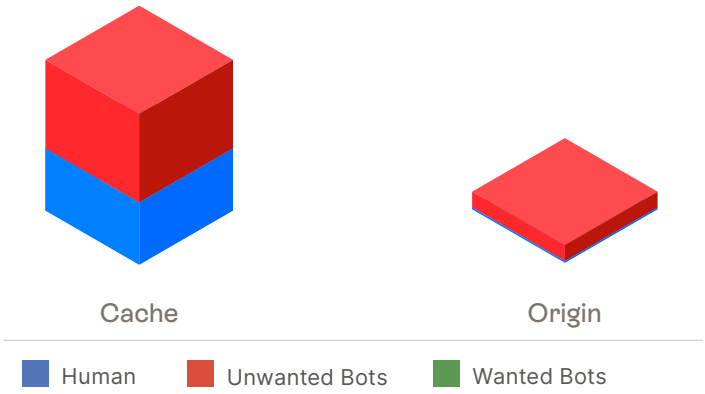
16% ChatGPT User

*The global baseline = the % across all observed regions and industries in Fastly's total traffic.

FASTLY THREAT INSIGHTS REPORT

JAPAC (Japan and Asia Pacific)

Bots vs. Humans in 2026



The JAPAC region received

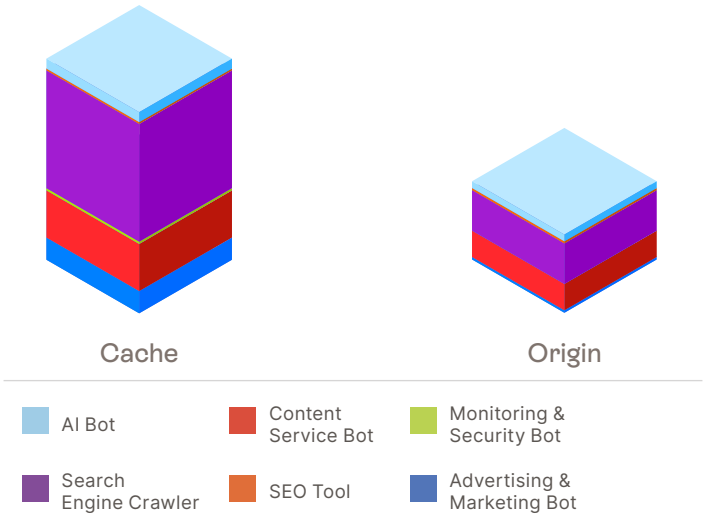
27% less human traffic

29% more unwanted bot traffic

59% less wanted bot traffic

All figures compared to global baseline.

Wanted Bots by Type



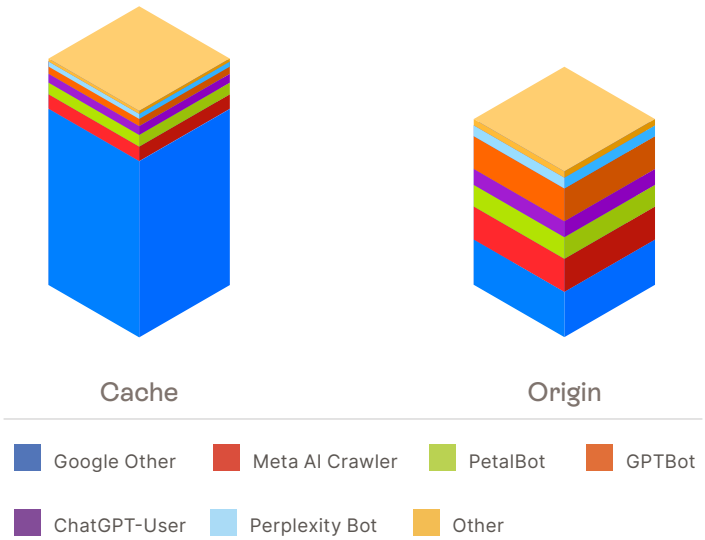
Wanted Bot Traffic

5% AI-driven bots

35% less than global baseline

5% of the JAPAC region's wanted bot traffic is attributable to AI. This is 35% less than the global baseline.

Top AI Bots



Top AI Crawlers

56% Google Other

12% Meta AI Crawler

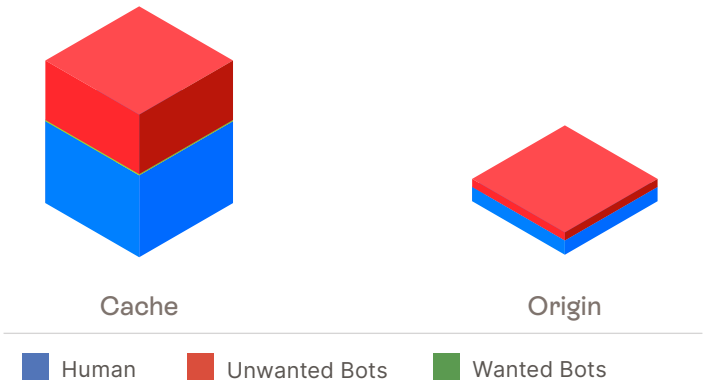
11% GPTBot

*The global baseline = the % across all observed regions and industries in Fastly's total traffic.

FASTLY THREAT INSIGHTS REPORT

EMEA (Europe, the Middle East, Africa)

Bots vs. Humans in 2026



The EMEA region received

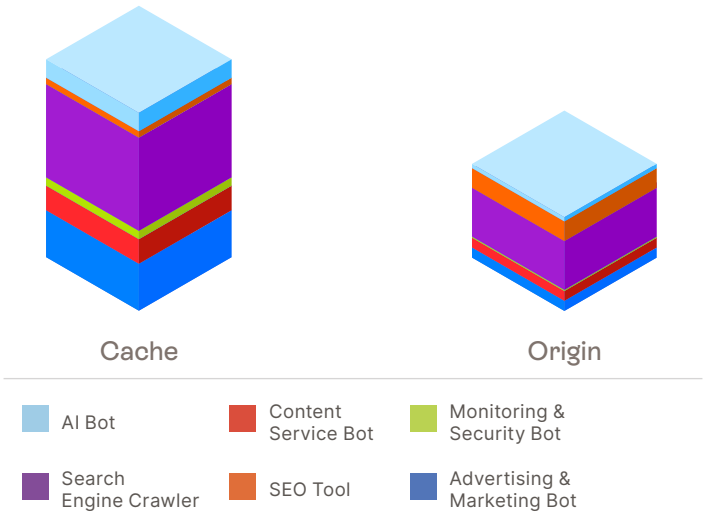
13% more human traffic

15% less unwanted bot traffic

52% more wanted bot traffic

All figures compared to global baseline.

Wanted Bots by Type



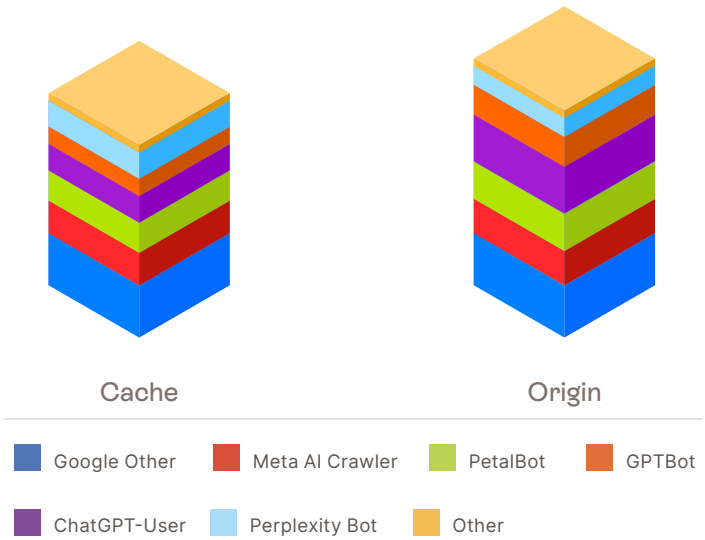
Wanted Bot Traffic

13% AI-driven bots

58% more than global baseline

13% of the EMEA region's wanted bot traffic is attributable to AI. This is 58% more than the global baseline.

Top AI Bots



Top AI Crawlers

25% Google Other

18% ChatGPT User

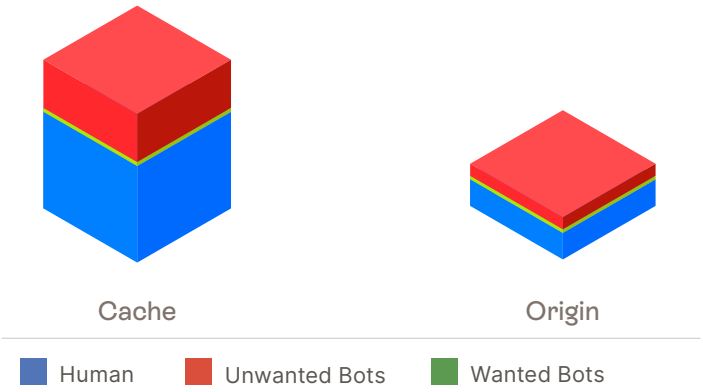
16% PetalBot

*The global baseline = the % across all observed regions and industries in Fastly's total traffic.

FASTLY THREAT INSIGHTS REPORT

LATAM (Latin America)

Bots vs. Humans in 2026



The LATAM region received

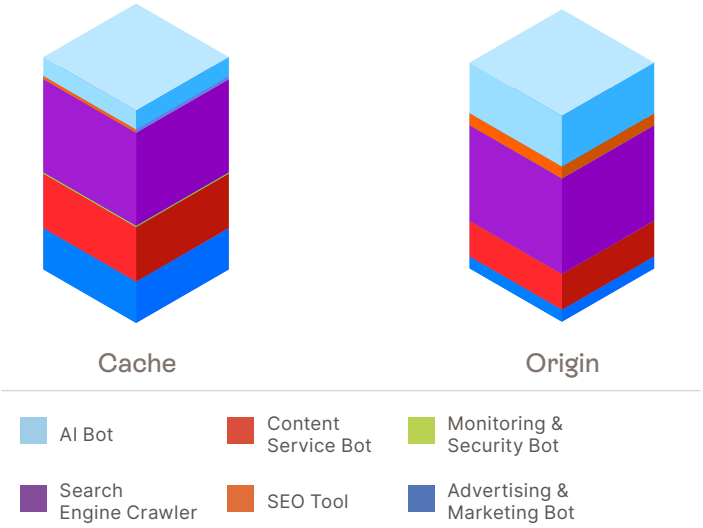
25% more human traffic

32% less unwanted bot traffic

470% more wanted bot traffic

All figures compared to global baseline.

Wanted Bots by Type



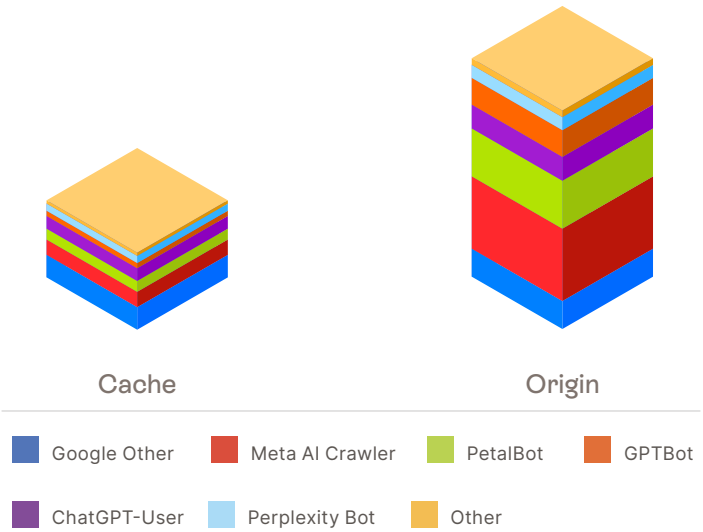
Wanted Bot Traffic

17% AI-driven bots

101% more than global baseline

17% of the LATAM region's wanted bot traffic is attributable to AI. This is 101% more than the global baseline.

Top AI Bots



Top AI Crawlers

30% Meta AI Crawler

20% PetalBot

17% Google Other

*The global baseline = the % across all observed regions and industries in Fastly's total traffic.

About this report

The Fastly Threat Insights Report highlights security trends, attack vectors, and threat activity across the application security landscape. Drawing from trillions of application and API requests across our global customer base, the report offers a real-time view into what's materially impacting organizations in the context of larger trends.

This quarter's insights are derived from traffic analyzed across Fastly's Next-Gen WAF and Bot Management products between January 1, 2026, and January 31, 2026. These solutions collectively protect over 130,000 applications and APIs* and inspect around 6.5 trillion requests per month**. Fastly's broad visibility, spanning edge and cloud-native architectures, combined with our presence across a wide range of industries, including leading e-commerce, streaming, media and entertainment, financial services, and technology organizations, provides us with a unique and comprehensive view of the global web application threat landscape.

Methodology

Some bots provide clear verification mechanisms, such as publishing IP ranges or supporting reverse DNS checks, that make it straightforward to confirm the authenticity of their traffic. In these cases, impostor traffic is easily distinguished and has been excluded from the dataset used in our analysis. However, not all bots offer this level of transparency. For bots lacking public verification data, we rely on alternative heuristics such as the originating network, behavioral patterns, and other identifiable traffic signatures to assess authenticity. While these methods are generally effective, they do not offer absolute certainty. We have worked to accurately classify the data presented. We encourage bot operators to adopt transparent verification practices to improve attribution. We have also chosen not to reclassify general-purpose search engine crawlers as AI bots. This is a deliberate decision to avoid misleading conclusions, particularly since blocking such crawlers could unintentionally impact a website's visibility in the associated search engines.

**130,000 App deployments protected as of 04-22-2025*
***~6.5 trillion Average NGWAF Requests Inspected Monthly for the trailing 6 months as of 04-22-2025*

To protect the privacy of our customers and avoid unintentionally singling out individual websites, certain details have been intentionally excluded or aggregated in the visualizations and analysis presented in this report. Fastly serves a significant portion of the web, giving it a global perspective on bot activity. This report focuses on overarching patterns in bot behavior, rather than on any individual site or service.

Definitions

Name	Definition
Bot Traffic	Any non-human internet traffic can be beneficial (e.g., search engine crawlers) or malicious (e.g., carding).
Unwanted Bots	Unwanted bots account for a significant portion of internet traffic, generated by automation tools that provide no business value to websites. Many of these bots are malicious, posing risks such as fraud, data scraping, account takeovers, and infrastructure strain.
Wanted Bots	Wanted bots are legitimate automation tools that send requests to websites, typically in ways that benefit the site. Fastly maintains a curated list of these bots, organized by their specific purposes. These bots play an essential role in many online functions, including search engine indexing, site performance monitoring, and security.
Headless Bots	A headless bot is an automated program that runs without a graphical interface, performing tasks in the background. It can interact with websites, APIs, or other systems, simulating a human interaction often for data scraping, automated product purchase, account creation, account takeover, etc.
Common Headless Automation	Common Headless Automation refers to a classification of headless bots. They are a collection of widely used browser extensions, programs and tools that interact with websites and APIs.
AI Crawler	AI crawlers are automated software programs that systematically visit websites and online resources to collect data used by artificial intelligence systems. They operate without direct human control, following programmed rules to discover, read, and process content at scale. Unlike manual data collection, AI crawlers can scan millions of pages efficiently, making them a core component of modern AI development and deployment.
AI Fetcher	AI fetchers are automated systems that retrieve specific pieces of content for the use of artificial intelligence applications. Unlike AI crawlers, which systematically scan large portions of the web, AI fetchers typically access individual URLs or small sets of resources in response to a direct request.
CDN Cache	In the context of CDNs, CDN caching reduces latency by allowing you to deliver your content from a location closer to the user, significantly improving load times. A CDN cache works by storing copies of content on edge servers that are distributed around the globe. Whenever a user requests content, the CDN delivers it from the nearest edge server rather than waiting for your origin server to respond. This decreases the workload on the origin server and supports higher traffic volume by distributing requests across multiple servers.
CDN Origin	A CDN origin server is the main web server in a CDN that hosts all of the original content from a website - the images, videos, HTML and so on. It is the essential 'source of truth' for the CDN as it has the complete picture of the website.