2024

CISCO
TALOS

YEAR IN REVIEW

2024
YEAR IN REVIEW

CISCO TALOS

Table of contents

Introduction

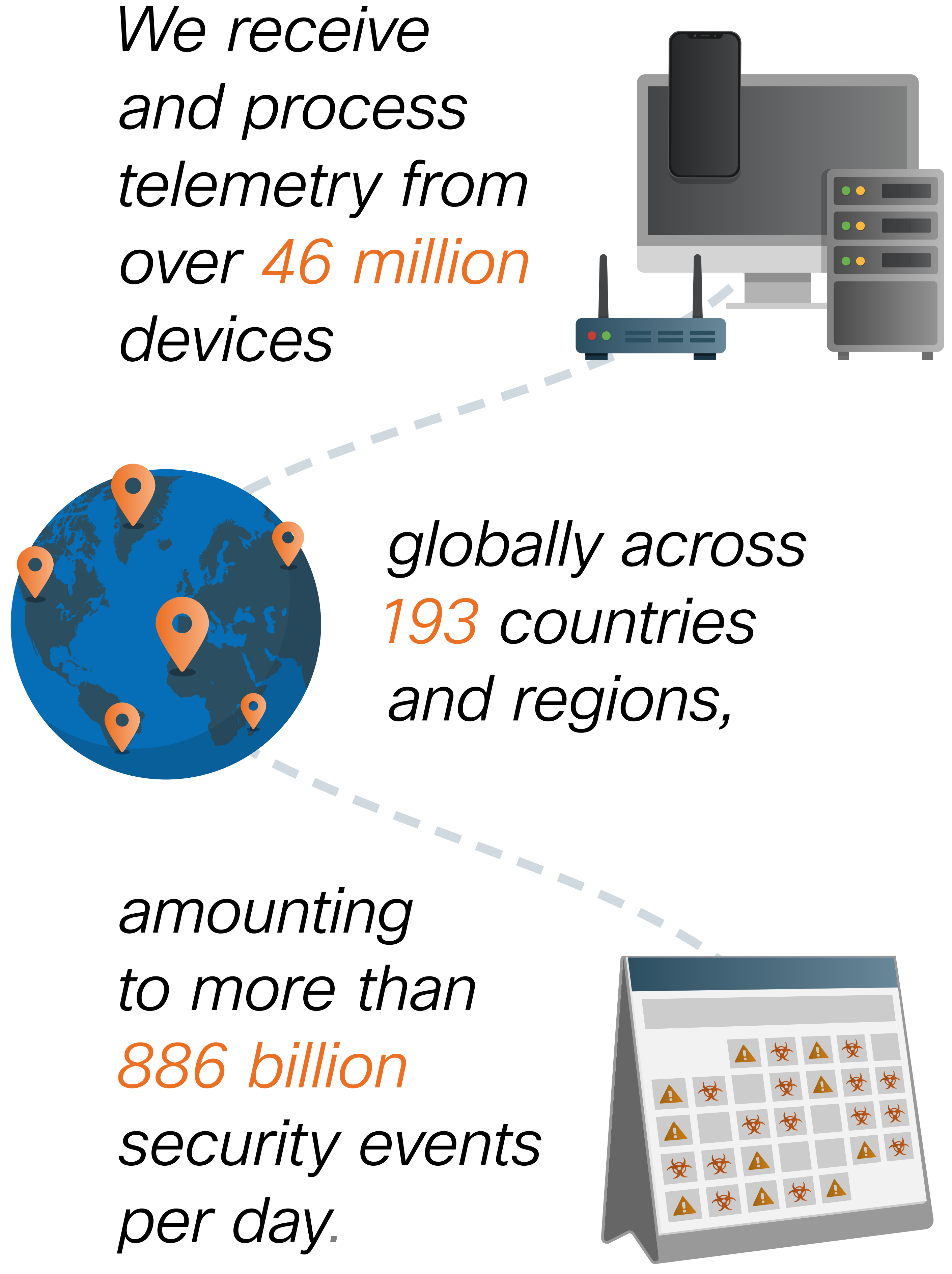
In 2024, threat actors prioritized stealth, simplicity, and efficiency, largely abandoning the use of custom malware and zero-day vulnerabilities in favor of simpler yet highly effective techniques. This environment underscores that it is more critical than ever for organizations to prioritize security fundamentals.

Identity was an enduring theme in 2024, as threat actors looked for ways to compromise users' unique digital footprint and coopt it for their own malicious purposes. This includes identifiers like login credentials, session IDs, API keys, digital certificates, and more. Identity-based attacks were dominant, accounting for 60% of all Cisco Talos Incident Response (Talos IR) cases, and actors relied on these techniques for major phases of their operations – initial compromise, lateral movement, privilege escalation, and more. Difficult to prevent and even harder to detect, identity-based attacks proved to be highly effective in 2024, allowing adversaries to go unnoticed for longer periods of time by using compromised valid accounts, foregoing the use of detectable malware, and sometimes leading to unfettered access to entire networks.

"Easy access" was another dominant theme, with decades-old CVEs topping our list of most-targeted vulnerabilities and security issues like misconfigured systems, MFA weaknesses, and unpatched systems appearing in more than half of Talos IR cases. Ransomware actors in particular honed in on this and attempted to disable poorly configured security solutions in most of the IR incidents we responded to, succeeding nearly 50% of the time. Unsurprisingly, education and healthcare were the top-targeted industry verticals, as actors continue to compromise organizations that tend to have lower cyber budgets and deal with more administrative bureaucracy that makes it difficult for them to maintain quick, agile footing in defense of these threats.

As we look ahead, threats to the very systems that underpin our networks remain persistent, and organizations will need to continue prioritizing security fundamentals and addressing aging infrastructure that poses significant risks. Sophisticated adversaries are capitalizing on vulnerable network hardware, public-facing applications, and cloud applications – all ingress points to network environments that should not be overlooked from a security perspective. Furthermore, as attacks in this space are increasingly relying on identity-based techniques, it's more important than ever that organizations adhere to security fundamentals, as so many of these attacks can be prevented by properly deploying and configuring multi-factor authentication, spotting socially engineered phishing lures, and identifying unusual activity emanating from legitimate accounts on the network.

In the artificial intelligence (AI) space, threat actor use of AI and machine learning (ML) largely fell short of industry projections in 2024, with actors relying on these technologies to enhance their techniques rather than aid in the creation of new ones. Adversaries used generative AI tools, such as large language models (LLMs), to create convincing social engineering campaigns and automate malicious activities.
In 2025, we expect to see their use expand, possibly with actors leveraging the technologies to create capabilities that can compromise AI models, systems, and infrastructure.



We receive and process telemetry from over 46 million devices

globally across 193 countries and regions,

amounting to more than 886 billion security events per day.

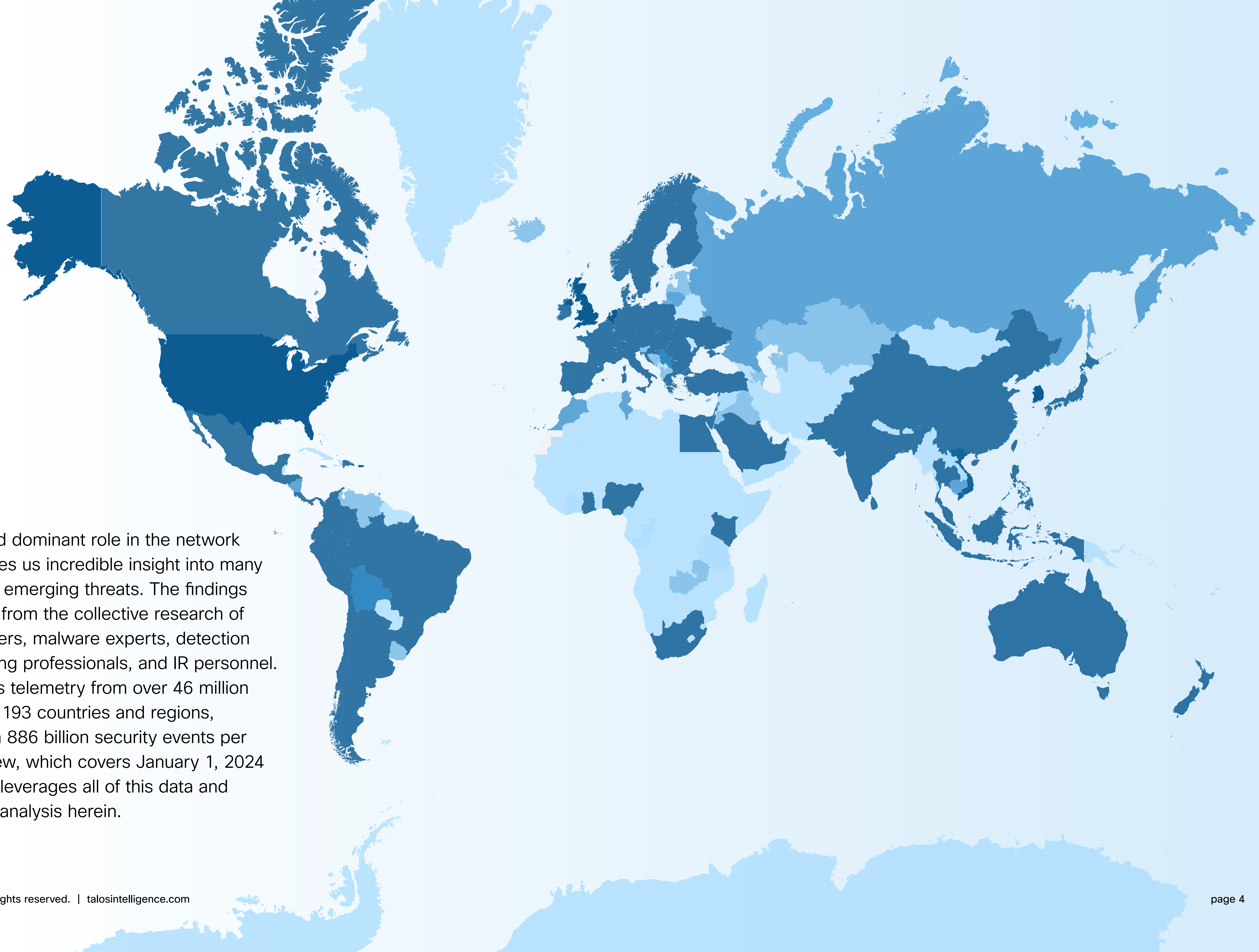


Figure 1

Cisco’s global visibility

Located devices

Less More

Cisco’s global reach and dominant role in the network infrastructure space gives us incredible insight into many of today’s enduring and emerging threats. The findings in this report are pulled from the collective research of hundreds of threat hunters, malware experts, detection specialists, data modeling professionals, and IR personnel. We receive and process telemetry from over 46 million devices globally across 193 countries and regions, amounting to more than 886 billion security events per day. Talos’ Year in Review, which covers January 1, 2024 to December 31, 2024, leverages all of this data and expertise to deliver the analysis herein.

Top-targeted vulnerabilities

Actors key in on historic, widely used CVEs in 2024

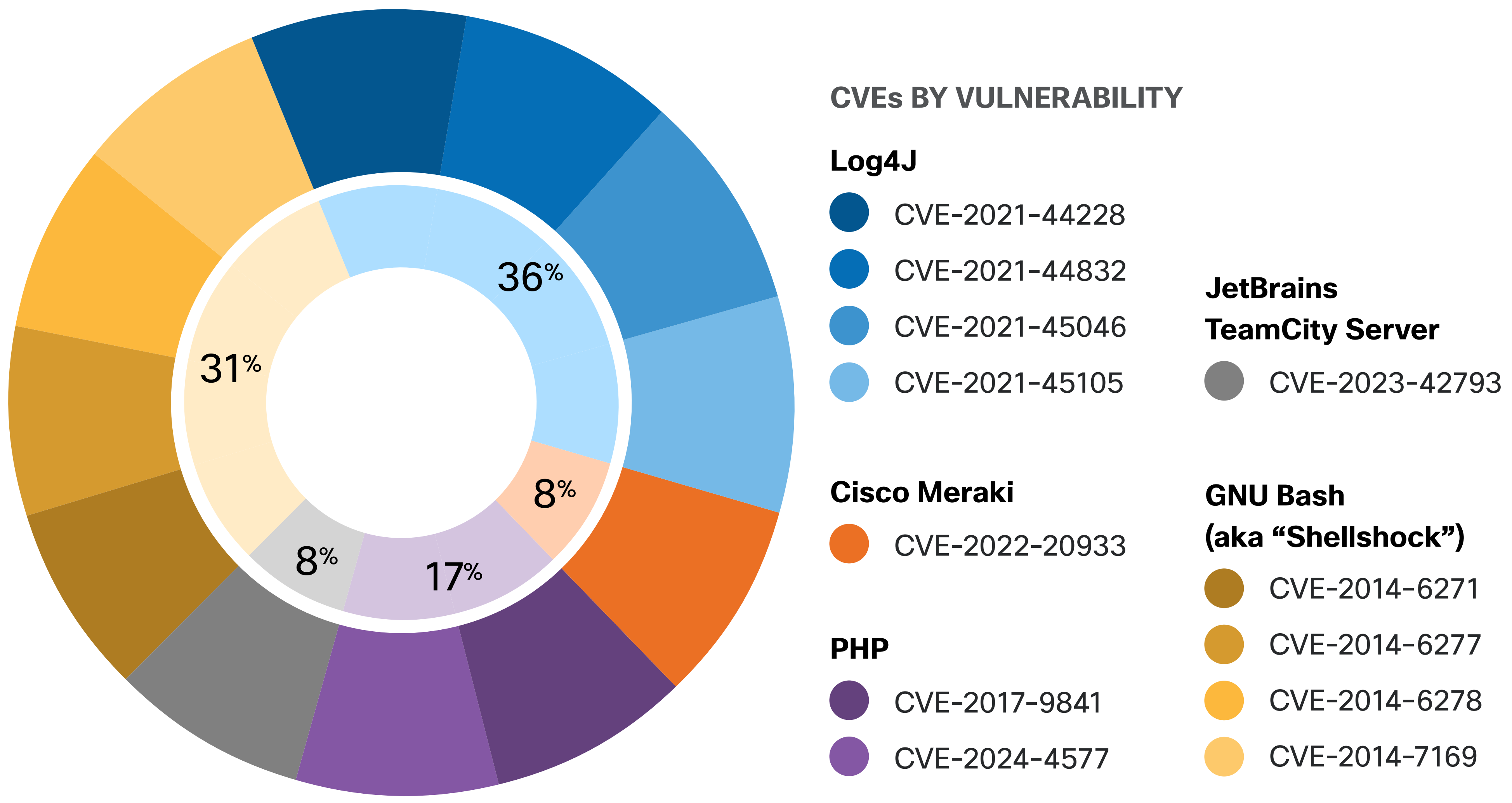
The top-targeted vulnerabilities in 2024 were mostly older CVEs that have been public for several years. Notably, four of the top twelve CVEs that made our list were published a decade ago, and the notorious Log4j vulnerabilities – which were disclosed in early 2021 – are also featured. This is a stark reminder that threat actors frequently target unpatched systems, and failure to apply security updates leaves organizations vulnerable to many attacks that could otherwise be prevented.

Exploitation attempts against the Apache Log4j logging library remain high nearly four years after the vulnerabilities were discovered. Log4J is one of the most widely used open-source programs in the world. While the vulnerabilities, collectively known as "Log4Shell," were patched shortly after discovery, they will likely pose a long-term risk for organizations because Log4j is so deeply embedded in the software supply chain. The U.S. Department of Homeland Security estimates it will take at least a decade to find and fix every vulnerable instance.

Relatedly, all of the vulnerabilities on our list impact widely used software and hardware, creating an incredibly broad attack surface that threat actors can exploit to infiltrate a broad range of sectors and geographies. For example, CVE-2017-9841 and CVE-2024-4577 affect PHP, a common programming language. Estimates show that between 75 and 80 percent of the world's two billion websites rely on PHP, including popular sites like Facebook and Wikipedia and e-commerce platforms like Etsy and Shopify. Vulnerabilities in the underlying code of these websites can allow attackers to gain unauthorized access and even lead to major data breaches. In January 2024, CISA and the FBI published an advisory warning of actors exploiting CVE-2017-9841 to deploy Androxgh0st, a malware known for its ability to establish a botnet that can further identify and compromise vulnerable networks.

Figure 2

Top-targeted vulnerabilities in 2024



Another common scripting language, Bash, was also strongly represented on the list, with four related vulnerabilities appearing among the most frequently targeted CVEs. Bash is the common command-line shell used in many Linux/UNIX systems and older MacOS versions. Collectively, the 10-year-old vulnerabilities are known as "Shellshock" and prompted comparisons to the notorious Heartbleed bug (CVE-2014-0160) from 2014 that sparked a global security crisis.

As mentioned previously, all of the top-targeted vulnerabilities affect software and hardware that are ubiquitous in systems globally, creating a broad attack surface. The types of threat actors that have been observed exploiting these CVEs increases the concern around risks to vulnerable organizations. For example, a variety of advanced threat actors have reportedly leveraged CVE-2023-42793 (impacting JetBrains' TeamCity servers) in their operations, including the Russian state-sponsored threat group APT29 (aka CozyBear), multiple North Korean state-sponsored groups, and the BianLian ransomware gang.

Shellshock's lasting impact

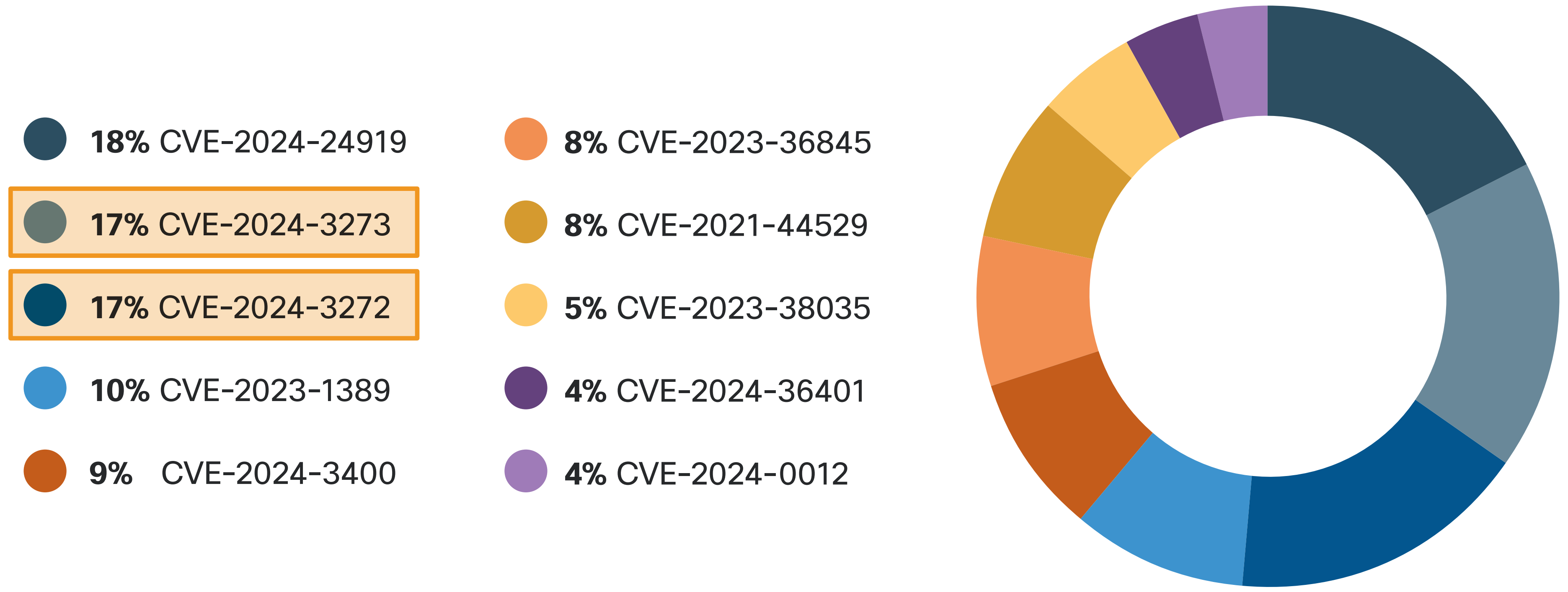
The Shellshock vulnerability, which affects the Bash scripting language in widely used operating systems like Linux and macOS, remains a problem more than a decade after its discovery in 2014. Bash is integrated deeply into applications and system processes globally. Additionally, many web servers, routers and internet-of-things (IoT) devices rely on Bash to execute commands, meaning that vulnerable devices connected to the internet are potential targets. These hardware components are often less frequently updated or harder to patch, especially in industrial or critical infrastructure settings.

Shellshock's direct consequences may not have been as catastrophic as other high-profile breaches and cyber attacks, but it is a persistent problem. For example, in 2019, Talos discovered a global state-sponsored espionage campaign called "Sea Turtle" that manipulated DNS records to gain access to sensitive systems. The adversary relied on several vulnerabilities, including Shellshock, to gain initial access.

While other confirmed public examples of state-sponsored cyber actors targeting Shellshock are limited, it's very likely that other advanced actors have attempted to exploit Shellshock. Many well-known adversaries like the Russian state-sponsored group APT28 and North Korean state-sponsored Lazarus Group exploit critical vulnerabilities in widely used software, making Shellshock a likely tool in their broader espionage and attack campaigns.

Figure 3

Top-targeted network device CVEs



Vulnerabilities affecting EOL devices are among most targeted network device CVEs

We also looked specifically at the top-targeted network device vulnerabilities to see what types of devices attackers are prioritizing in their operations. This list only includes network device vulnerabilities that were added to CISA's Known Exploited Vulnerabilities (KEV) catalog in 2023 or 2024. Of those, three (CVE-2024-24919, a Check Point VPN zero-day; and CVE-2024-3273 and CVE-2024-3272, affecting older D-Link hardware), accounted for more than 50% of network device targeting in that data set (see figure 3).

Many of these vulnerabilities have largely been exploited by known botnets like Mirai, Gafgyt, and others, which can establish control over the compromised devices and command them to carry out distributed denial-of-service (DDoS) attacks and other malicious activity. Because of the access that routers, firewalls, and other network devices afford, their compromise can allow an attacker to easily move laterally, carry out other phases of their attacks, and potentially take over entire networks. At least one of the vulnerabilities (CVE-2023-38035) has been exploited by ransomware operators.

Notably, some of these top-targeted vulnerabilities affect end-of-life (EOL) devices and therefore have no available patches, despite still being actively targeted by threat actors. Examples include CVE-2024-3273 and CVE-2024-3272 (D-Link NAS devices), which were the second and third most targeted network device vulnerabilities on our list, respectively. This underscores the importance of decommissioning and replacing EOL components of an organization's network as soon as possible.

Figure 4

Top 10 most targeted network device vulnerabilities

□ Denotes EOL product

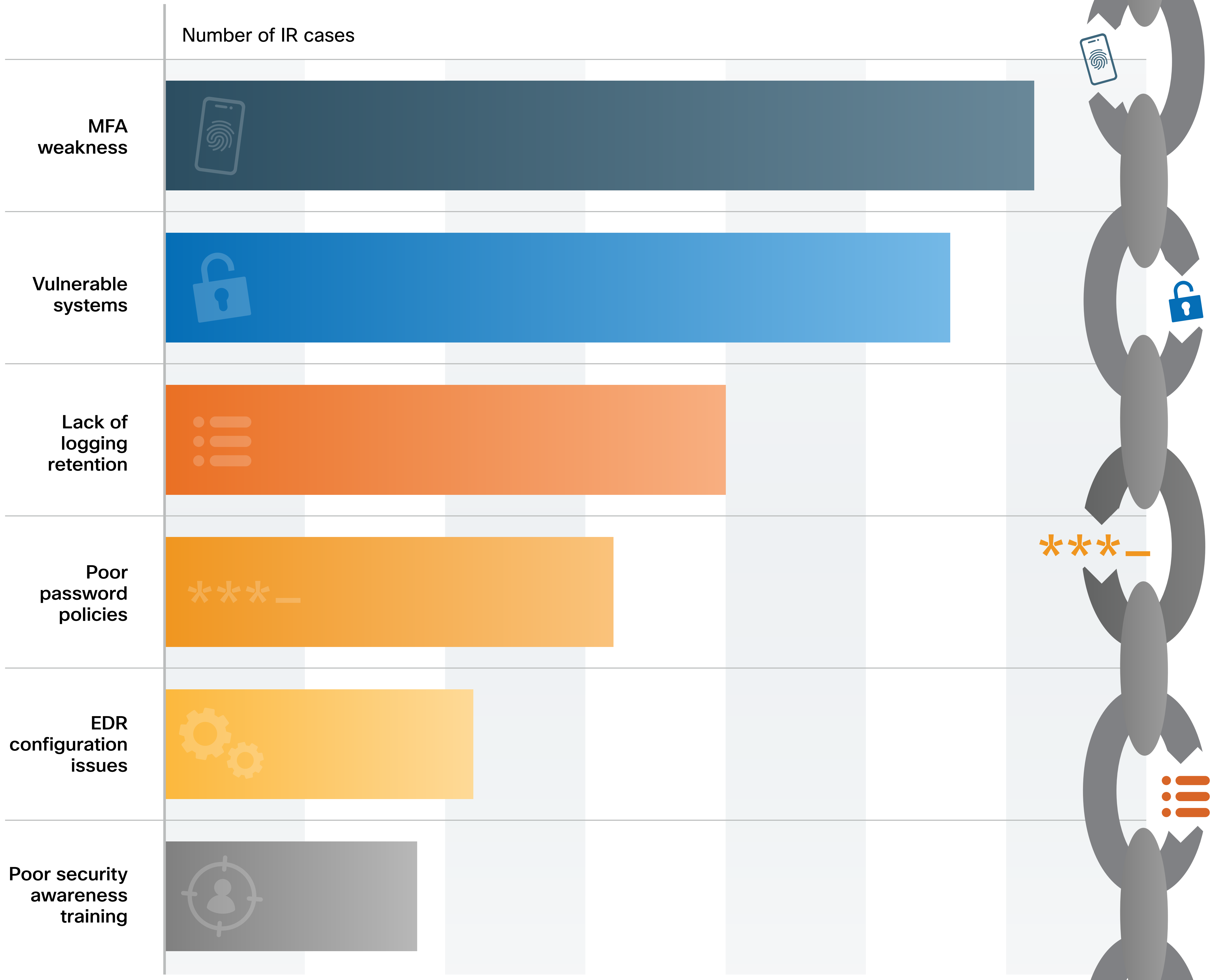
CVE	Manufacturer	Product	Device type	Vulnerability description
CVE-2024-24919	Check Point	Quantum Security Gateways	Firewall/VPN	Attacker can read sensitive data like password hashes.
CVE-2024-3273	D-Link	Multiple NAS Devices	Network attached storage (NAS)	Allows attacker to execute arbitrary base 64-encoded commands on the devices.
CVE-2024-3272	D-Link	Multiple NAS Devices	Network attached storage (NAS)	Allows attacker to execute arbitrary base 64-encoded commands on the devices.
CVE-2023-1389	TP-Link	Archer AX21	Router	Attacker can inject commands, which would be run as root, with a simple POST request.
CVE-2024-3400	Palo Alto Networks	PAN-OS	Firewall	Gives attacker ability to execute commands with root privileges on the firewall.
CVE-2023-36845	Juniper	Junos OS	Network device software	Attacker can inject and execute malicious code.
CVE-2021-44529	Ivanti	Endpoint Manager Cloud Service Appliance	Endpoint device manager	Allows attacker to execute malicious code with limited permissions.
CVE-2023-38035	Ivanti	Ivanti Sentry	Security gateway	Allows attacker to access sensitive API data and configurations, run system commands, or write files onto the system.
CVE-2024-36401	OSGeo	GeoServer	Server	Attacker can conduct remote code execution via specially crafted input.
CVE-2024-0012	Palo Alto Networks	PAN-OS	Firewall	Allows attacker to gain administrator privileges.

It’s no surprise that most of the top-targeted vulnerabilities are many years old, as patch management continues to be a problem for organizations. Unpatched/vulnerable systems was the second most common security weakness observed in 2024, according to Talos IR data (see figure 5).

Actors showed no preference for device size when carrying out their operations, despite reporting that small office and home office (SOHO) devices are often more frequent targets because they presumably might be less secure. In fact, the targeting patterns looked overwhelmingly similar across small (less than 50 users), medium (51-499 users), and large devices (500+ users), indicating that actors are opportunistic and do not prioritize the number of device users.

Figure 5

Top security weaknesses in Talos IR cases



Notably, some of the top-targeted vulnerabilities affect EOL devices and therefore have no available patches, despite still being actively targeted by threat actors.

Talos' top 10 tips for securing network devices:

1 Update devices as aggressively as possible. This includes patching current hardware and software against known vulnerabilities and replacing EOL hardware and software.

2 Implement robust authentication methods. Use multifactor authentication, select complex passwords and community strings, and avoid default credentials.

3 Adhere to security best practices, including conducting regular updates, managing access controls, implementing user education, and enforcing network segmentation.

4 Encrypt all monitoring and configuration traffic, including SNMPv3, HTTPS, SSH, NETCONF, and RESTCONF.

5 Stay informed and up-to-date on security advisories from the U.S. government and industry. Consider suggested configuration changes to mitigate issues captured by these reports.

6 Lock down and actively monitor credential systems, such as TACACS+ and any jump hosts.

7 Store configurations centrally and push to devices. Do not allow devices to be the trusted source for their configurations.

8 Use authentication, authorization, and accounting (AAA) to deny configuration changes for key device protections, such as local accounts, TACACS+, and RADIUS.

9 Monitor your environment for unusual changes in behavior or configuration. Be on the lookout for exposure of administrative or unusual interfaces (such as SNMP, SSH, and HTTP(S)), and monitor syslog and AAA for unusual activities.

10 Profile your devices' baseline to identify any changes. Fingerprint network devices via NetFlow and port scanning for a shift in surface view, including new ports opening/closing and traffic to/from. When possible, develop NetFlow visibility to identify unusual volumetric changes.

Additional resources:

- Our blog on Salt Typhoon activity provides an in-depth look into tactics, techniques, and procedures (TTPs) leveraged by this sophisticated threat actor to target network infrastructure, as well as detection and prevention guidance.
- Talos also coordinated with CISA and other partners on this guide for hardening communications infrastructure.
- Finally, our 2023 blog detailing sophisticated attacks on network infrastructure by state-sponsored actors contains attack chain information and actionable security recommendations.

Email threats

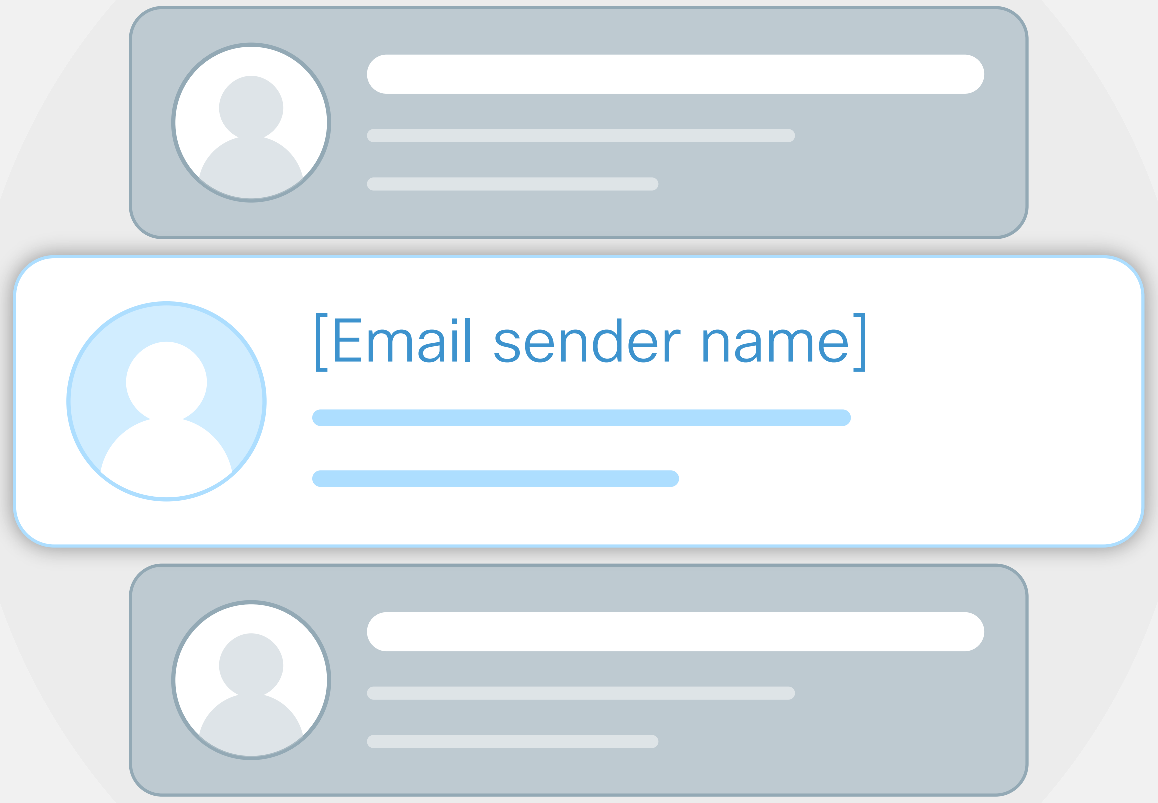
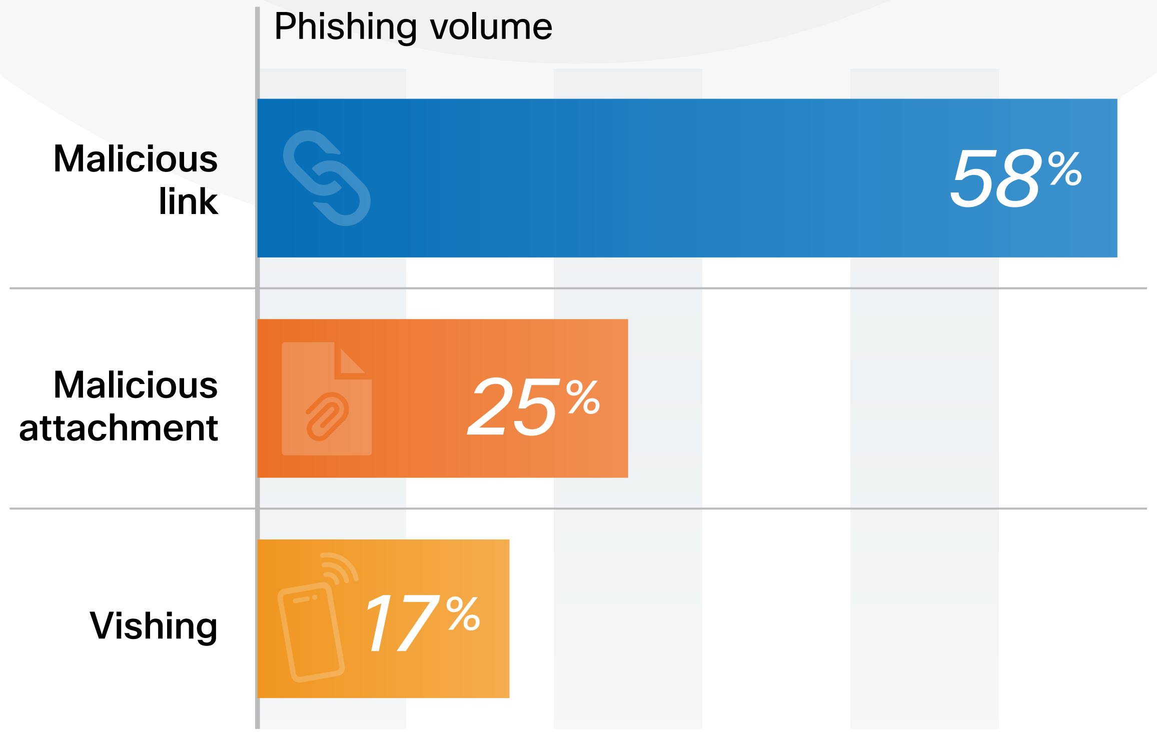


Figure 6

Types of phishing in Talos IR cases



Attackers spoof well-known brands in phishing lures

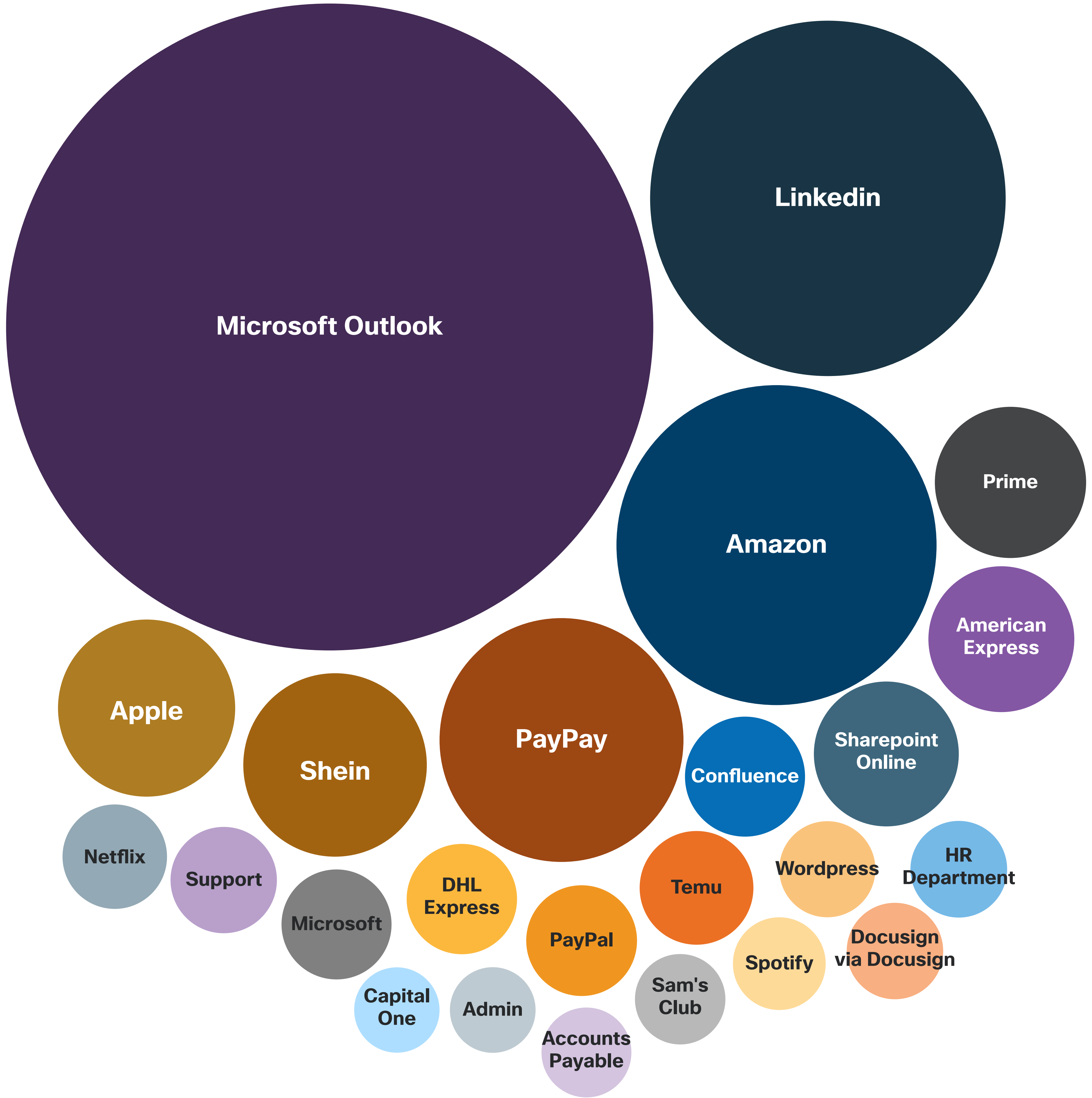
Phishing continues to be a main method of compromise, as threat actors can easily and anonymously send out high volumes of emails to reach their victims. We saw adversaries gain initial access via phishing in nearly a quarter of Talos IR incidents. In those cases, embedded malicious links appear to be more successful than other modes of phishing, like email attachments or voice phishing (vishing). Social engineering is a hallmark in this space, and artificial intelligence (AI) tools have made it even easier for actors to create believable lures crafted specifically for their targets.

We analyzed our email telemetry to identify prevailing social engineering preferences or trends. Figure 7 shows the most common brands appearing in sender display names from emails that were blocked. The findings give us a sense of the types of companies that threat actors might be trying to spoof to trick victims into engaging with their malicious phishing lures.

Microsoft Outlook was the most commonly spoofed brand, appearing as the sender name in 25% of blocked emails. Tech behemoths Amazon and Apple also had high prevalence. Other names that top the list include popular online payment services, international retail companies, and common enterprise collaborative applications. It's no surprise that these companies are among threat actors' favorites to imitate given their global ubiquity: High brand recognition likely coincides with increased trust and higher click rates.

Figure 7

Top spoofed brands in sender names



Actors use simple subject lines in phishing lures but still leverage major events

In addition to making emails appear like they're coming from a legitimate, trusted sender, threat actors also rely on the email's subject line to creatively and convincingly trick victims into opening the email. Figure 8 shows the most common terms appearing in subjects of blocked emails. These terms were overwhelmingly ordinary, common words one would expect to see in their daily inbox. Threat actors largely abandoned the use of urgent or time-sensitive subjects in their lures, instead opting for terms that are far less sensational and perhaps more likely to be mistaken as benign messages.

While ordinary words and phrases were the most prevalent terms appearing in email subject lines for the year writ large, we also saw evidence that threat actors remain attuned to major national events, and we see them quickly incorporate those themes into phishing lures and spam email to get higher click rates. In the example below, we saw the terms "Biden," "Harris," and "Trump" appearing more or less frequently across varying subject lines as threat actors tuned their message based on current events. Some of these changes aligned with major events in the 2024 presidential race. For example, when former Vice President Harris announced her campaign, we saw an increase in lures leveraging "Harris," while lures with "Biden" started to drop off. In the immediate weeks following the election, email lures featuring "Harris" dropped off dramatically, while those leveraging "Trump" remained consistently high.

Figure 8

Common terms in subject lines of malicious emails

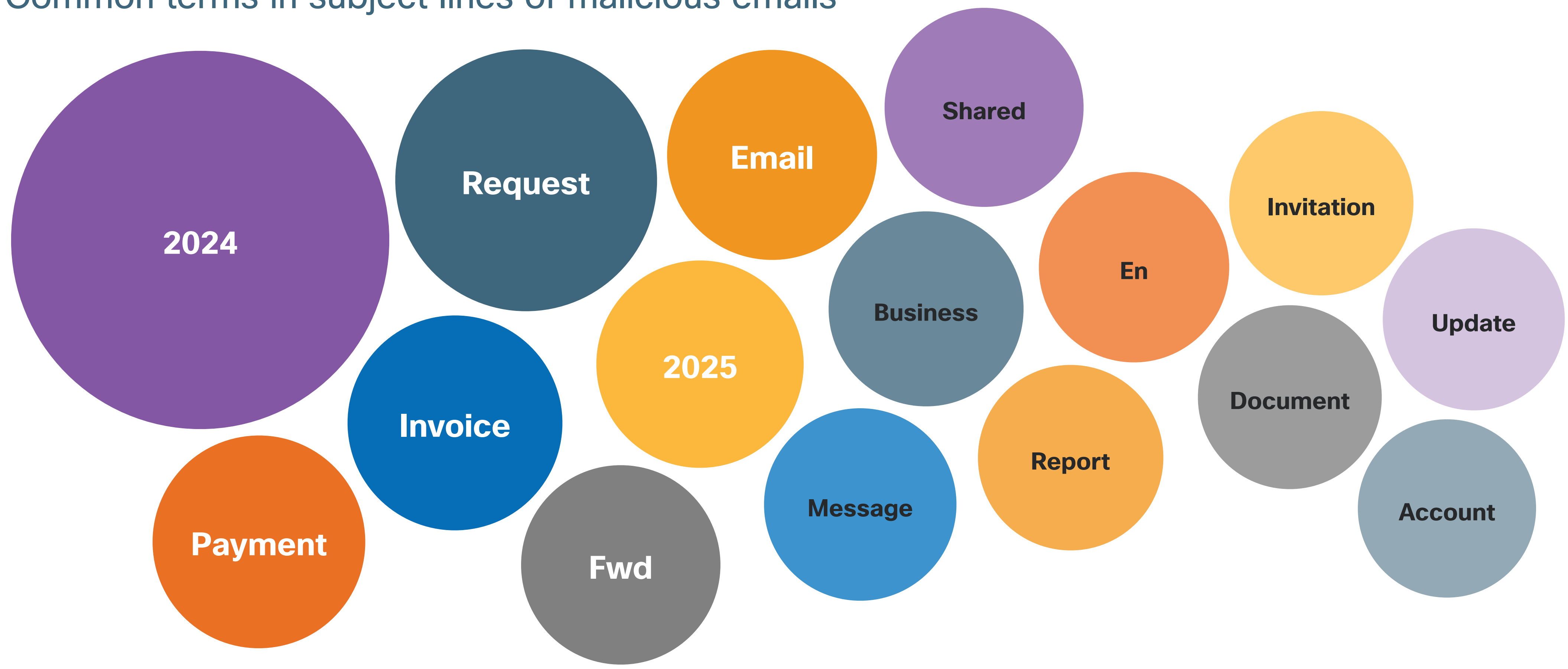
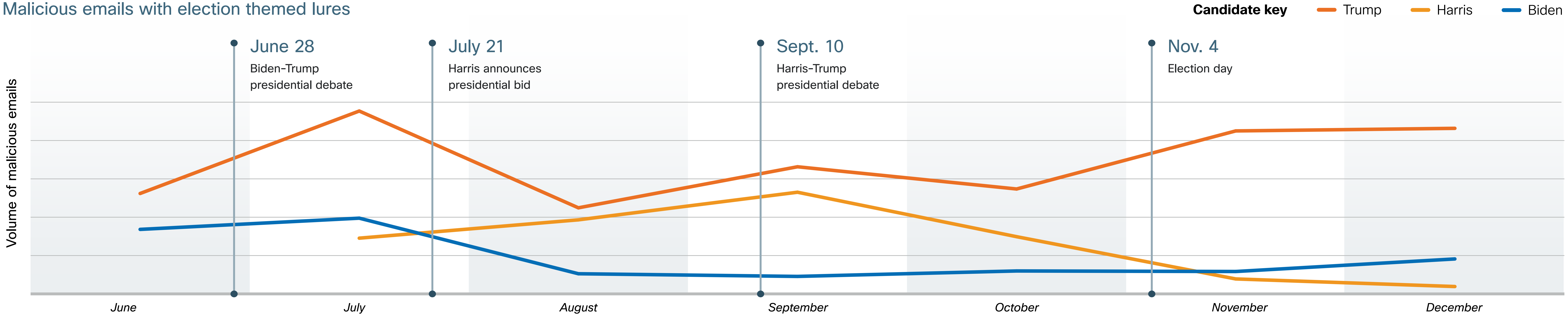


Figure 9

Malicious emails with election themed lures



2024
YEAR IN REVIEW

Most used tools

CISCO TALOS

Actors largely preferred LoLBins, enabling them to blend in with normal traffic

When we observed tool usage in Talos IR engagements, actors prioritized living-off-the-land binaries (LoLBins) — or tools and utilities found natively on an endpoint — more than commercial or open source tools (figure 11). While we saw LoLBins more frequently, the number of different LoLBins paled in comparison to the variety of commercial and open source tools (figure 10). This is because there are only a limited number of LoLBins native to targeted endpoints that will be of use to a threat actor. By contrast, the number of commercial and open source tools is ever growing, with actors continuously developing or creating these for their own use to deploy onto compromised systems.

Common LoLBins, such as PSExec, PowerShell and remote desktop protocol (RDP) are just three of the top five tools that were used to facilitate large components of an adversary's attack chain. For example, Talos IR responded to a ransomware incident in which PsExec, originally designed for network management, was used to execute malicious batch (BAT) files as well as the final BlackBasta ransomware binary, underscoring the impact of the misuse of common LoLBins. The table on the next page shows how some of these top observed tools, like PsExec, and commercial tool frameworks like Mimikatz, are intended to be used, and how threat actors leverage them in their operations.

Figure 10

Classification of tools observed in Talos IR incidents

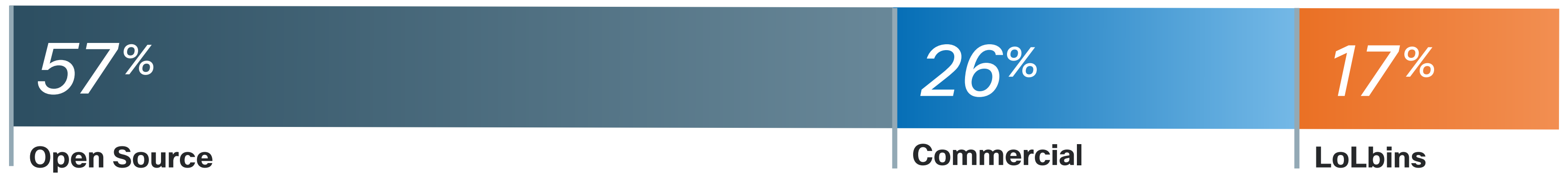
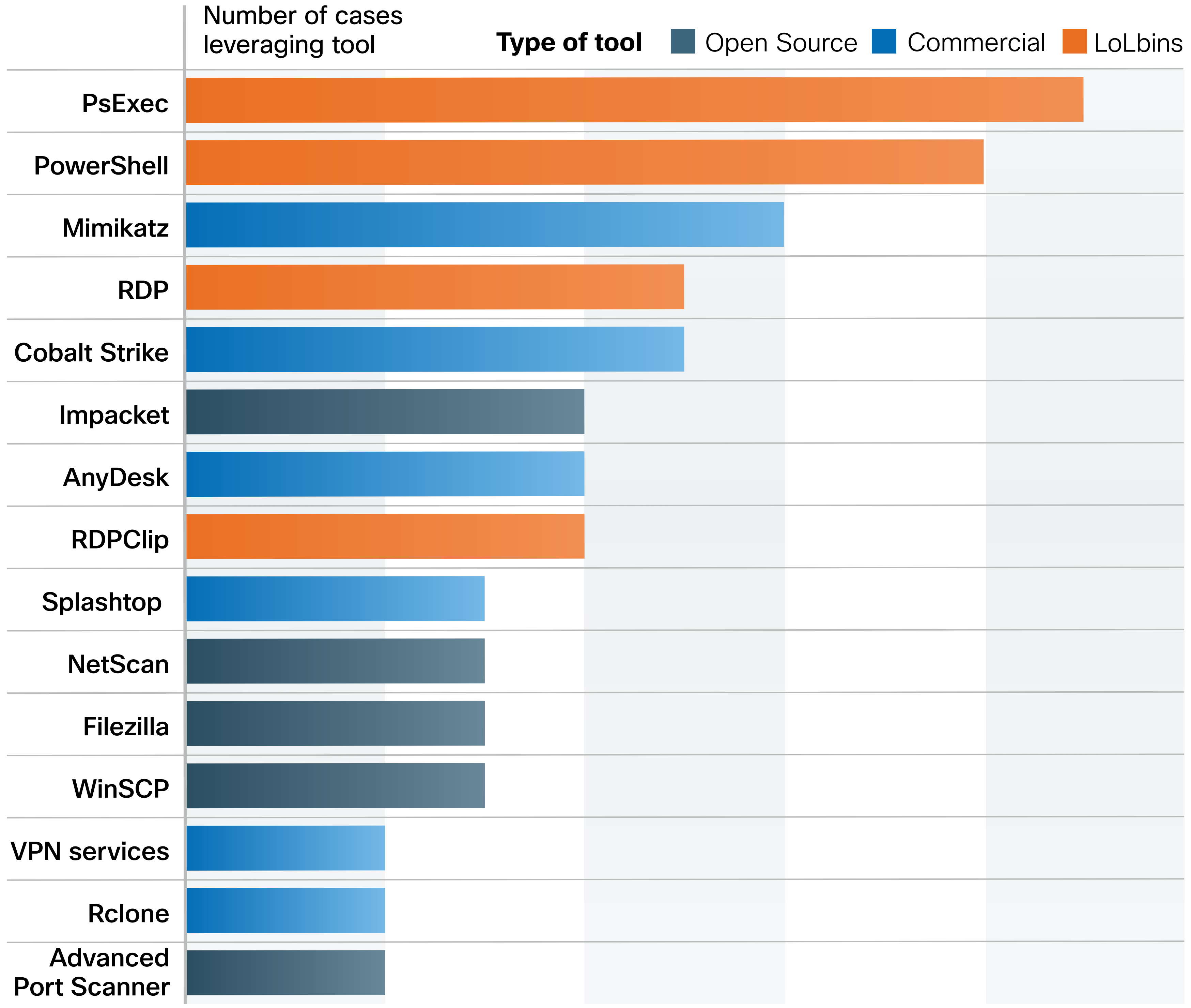


Figure 11

LoLBins used across the attack chain in Talos IR incidents



LoLBins enable actors to blend in with regular network traffic and avoid triggering antivirus or endpoint detection solutions. Furthermore, the nature of these binaries as default Windows utilities that come preloaded on Windows operating systems can make it difficult to define normal usage patterns, complicating efforts by defenders to identify abuse. LoLBins also likely help improve the efficiency of malicious operations, as an attacker does not need to take the time to install additional software or test the efficacy of external tools and exploits.

One example of a newly observed open-source tool this year in Talos IR engagements was DonPAPI, which automates credential dumping remotely on multiple Windows computers. This tool locates and retrieves Windows Data Protection API (DPAPI) protected credentials, also known as DPAPI dumping. From an identity perspective, open-source tools like DonPAPI pose a significant risk to organizations based on their wide availability on code repositories like GitHub and the ease of installation. DonPAPI searches for certain files, including Wi-Fi keys, RDP passwords, and credentials saved in web browsers, to help authenticate and move laterally to identify other assets in the environment. Ransomware groups have reportedly used DonPAPI for a few years now, highlighting the emphasis adversaries put on obtaining credentials using these types of tools.

Since organizations regularly use many of these tools to support daily operations, it can be difficult to discern when their use or presence on an endpoint might be nefarious. The table on the following page shows the most commonly seen tools in Talos IR cases from each category (e.g., LoLBin (PsExec), open-source (Impacket), and commercial (Mimikatz)) are intended to be used, and how actors are coopting them for their own malicious purposes.

> *From an identity perspective, open-source tools like DonPAPI pose a significant risk to organizations, based on the wide availability on code repositories like GitHub and the ease of installation.*

Use and abuse of most common tools in Talos IR cases

	PsExec	Impacket	Mimikatz
✓ Intended use	Part of Microsoft's Sysinternals suite of tools; allows users to run commands on local and remote systems.	Open-source Python library for performing network audits.	A credential-dumping utility commonly used by penetration testers and red teams to extract plain text passwords.
⚠ Malicous capabilities	Has the ability to execute processes on other systems remotely, remotely create accounts on target systems, download or upload a file over a network share.	Impacket modules like SecretsDump allow actors to steal account and password information from Active Directory databases.	Contains functionality to acquire information about credentials, including from LSASS memory, registry hives, DPAPI, among others.
Threat actor abuse	Many ransomware operations use PsExec to run their payload on all systems in the domain.	APTs and other actors frequently use Impacket to gain a foothold in the victim environment and move laterally.	Cybercriminals to APT groups use Mimikatz to steal account logins and credentials to aid in moving laterally in the victim environment.

2024

YEAR IN REVIEW

Ransomware

Higher education hit hardest in 2024

Ransomware actors targeted education entities more than any other sector in 2024. This is in line with trends from previous years, where education was also the most targeted in 2022, and the second most targeted in 2023. Ransomware attacks were also high against public administration, manufacturing, and healthcare entities, suggesting ransomware actors focused their operations against large organizations that traditionally have a low downtime tolerance and/or limited security budgets (see figure 12).

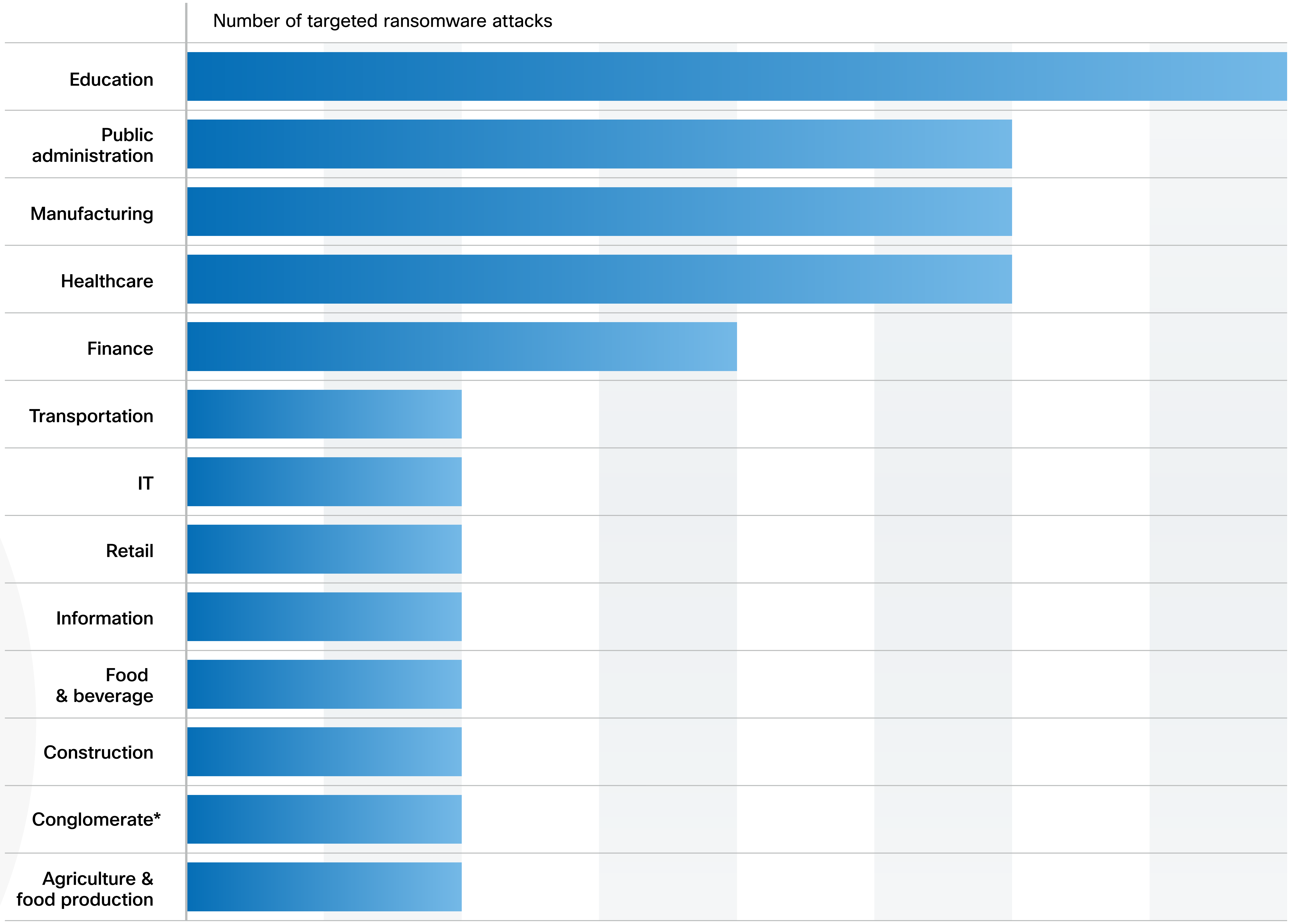
Interestingly, almost all the ransomware attacks against the education sector targeted higher education entities. Universities typically have greater cybersecurity budgets than primary and secondary schools, presumably leaving them better defended, but the data they house such as proprietary and/or government-funded research is likely of greater value for a ransomware attack. Universities also rely more heavily on their IT infrastructure for things like online classes and student research, incentivizing these institutions to minimize disruptions to their operability.

9,860

Employees on average at targeted organizations

Figure 12

Targeted sectors



Conglomerate organizations and their subsidiaries are not included in any other verticals.

What makes education such an attractive sector?

Insufficient funding

Primary, secondary, and high schools in particular often lack funding for proper information security capabilities, enabling adversaries to cause maximum damage with minimal ability for the victim to recover quickly.

Irregular monitoring

Effective cybersecurity often requires the ability to monitor and respond to threats 24/7, which schools may not have the bandwidth for outside of school hours, on weekends, and during breaks.

Minimal network segmentation

Many university networks have minimal segmentation between student networks, research networks, and administrative networks, providing adversaries a large attack surface and opportunity for lateral movement.

Attractive data stores

Schools, particularly universities, store a range of data that is of interest to ransomware actors, including financial account information, personally Identifiable information (PII), and classified research data.

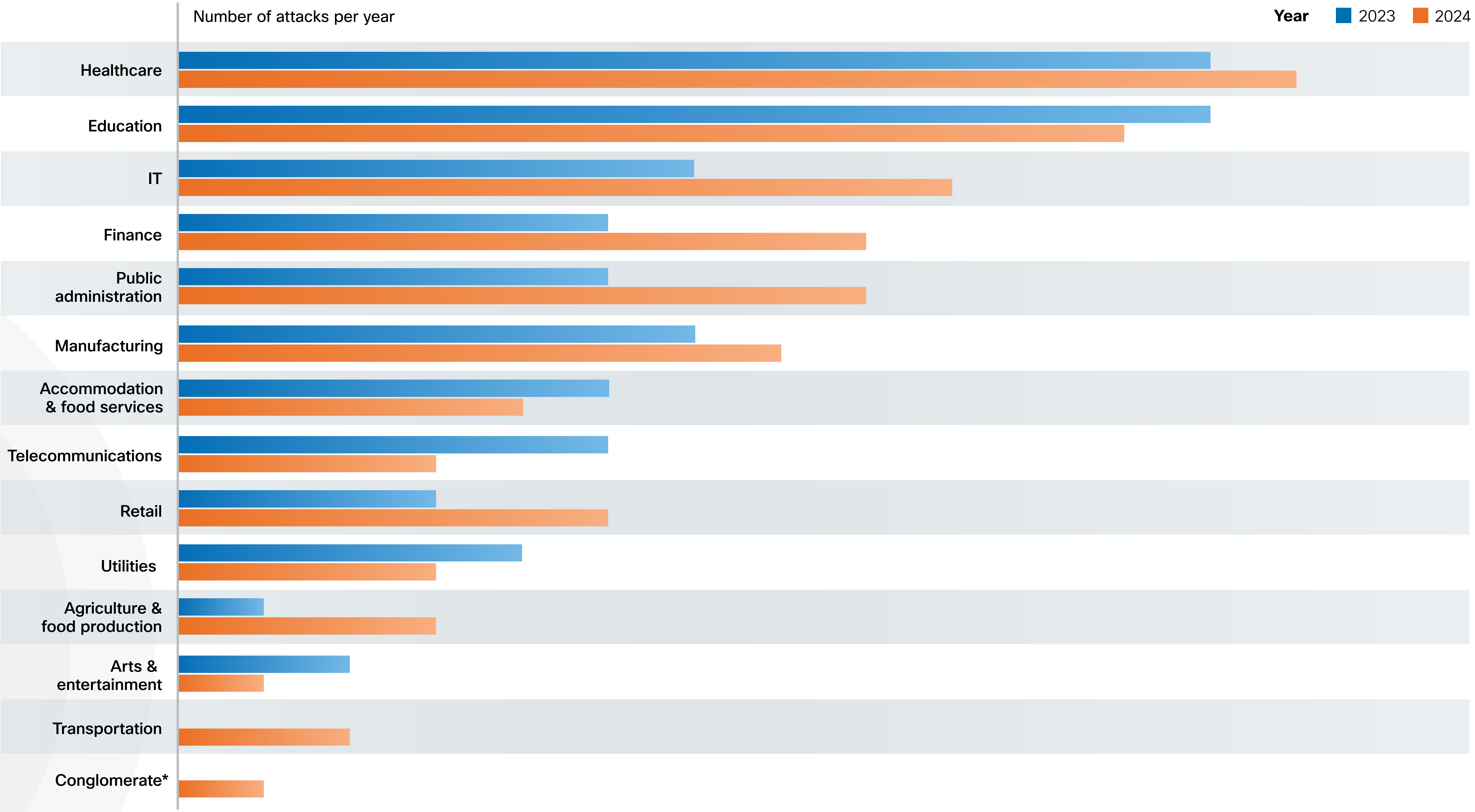
Poor device hygiene

Students often lack cybersecurity training and/or are less inclined to be concerned with the security of a school's network, leading to a low level of personal device hygiene. This poor hygiene creates easier lots of points for actors looking to compromise easy targets and gain access to a school's network.

The industry targeting trends for both ransomware and our broader data set are somewhat similar, with healthcare and public administration (i.e., local government) rounding out the top five in both instances. Interestingly, the targeting across sectors looked largely similar in 2023 and 2024, meaning that threat actor preference has remained fairly consistent (see figure 13).

Figure 13

Top affected sectors from both ransomware and our broader data set remain consistent across 2023 and 2024 IR data



*Conglomerate organizations and their subsidiaries are not included in any other verticals.

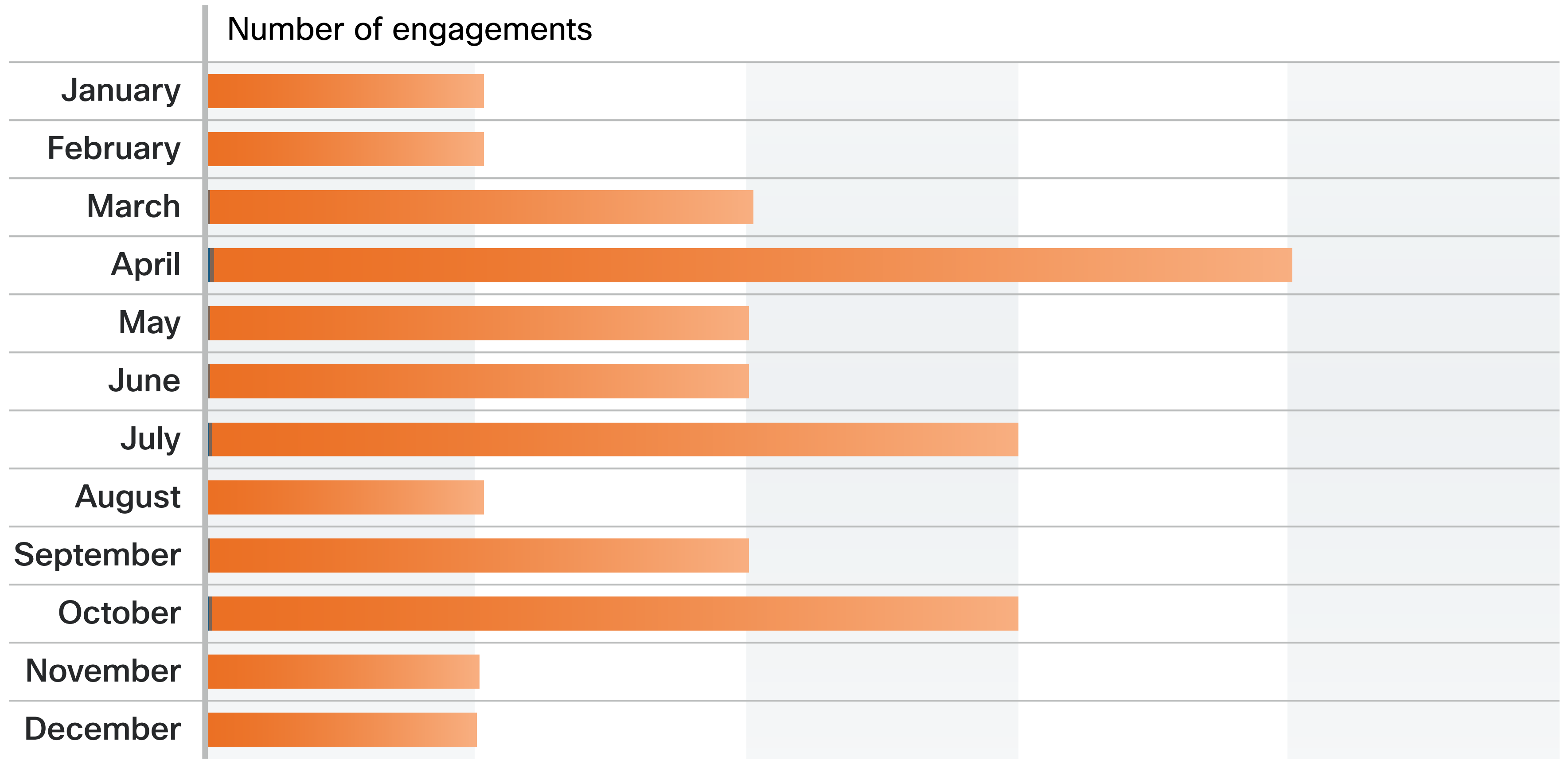
Ransomware attacks more frequent in spring and summer

Ransomware actors appeared to be more active on average during the spring and summer, based on Talos IR findings (see figure 14). These months overlap with times when schools are closed for break or employees or students might be more likely to be on vacation – possibly contributing to slower response times to cyber incidents or a more relaxed cyber security posture in general. Education entities typically operate at a reduced capacity during the summer months, and their calendars are often available online for the public – and possible attackers – to reference.

We’ve seen instances where adversaries take advantage of personnel being “out of office,” including one Talos IR case where a LockBit ransomware operator gained control of an IT account belonging to an employee on vacation. The threat actor easily gained access and created another account with admin rights to the entire domain to facilitate lateral movement.

Figure 14

Ransomware attacks by month

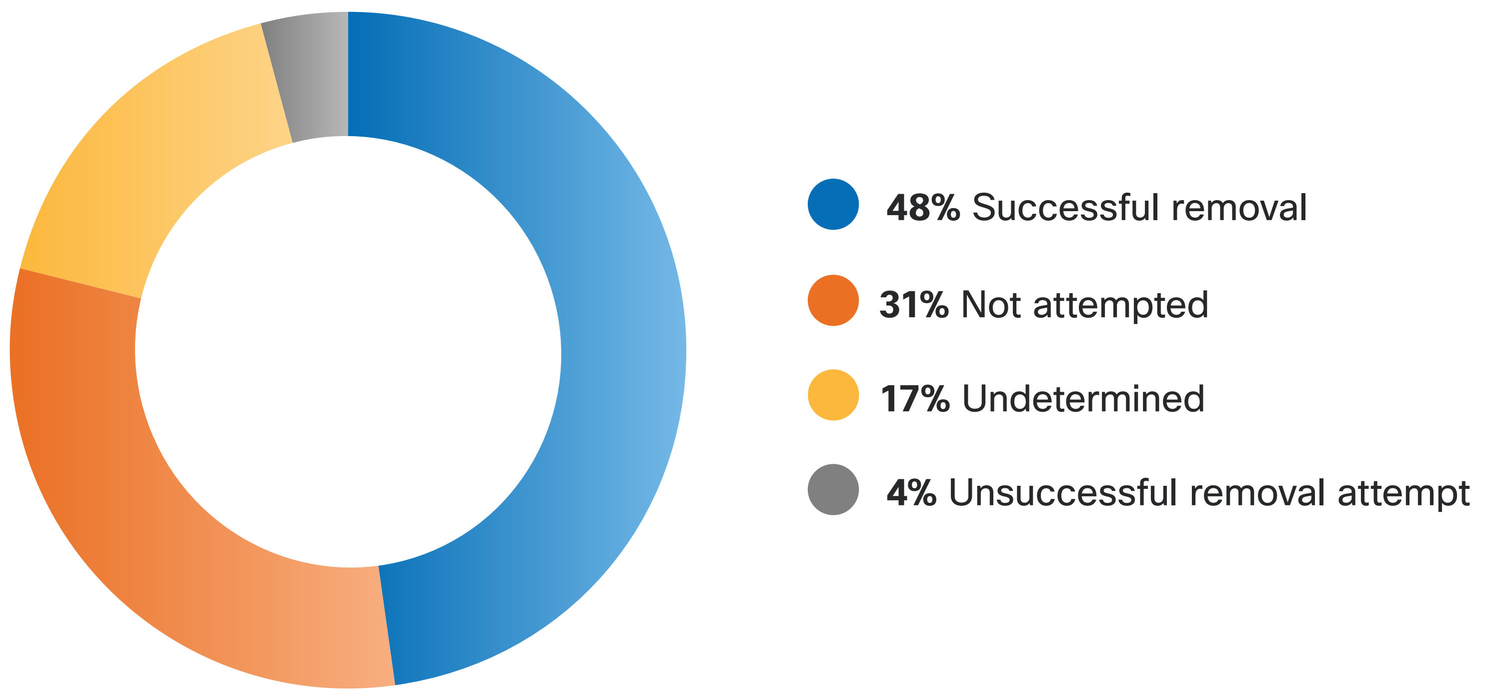


Actors prioritize disabling security solutions frequently and early on in their operations

Ransomware operators endeavored to disable targets’ security solutions in most of the Talos IR cases we observed, almost always succeeding (see figure 15). This was often one of the first actions actors took upon logging into a compromised network, taking advantage of endpoint solutions that did not require an agent or connector password and/or that were not configured properly.

Figure 15

Disablement of security solutions



Actors were quick to uninstall endpoint security products, which detect and quarantine the deployment of threats like ransomware on the system. They also modified certain solutions, like creating new firewall rules that can allow the adversary remote access, and removed evidence of their activity by deleting shadow copies and clearing event logs related to System, Application, and Security, a commonly observed ransomware TTP. These actions not only severely inhibit detection capabilities, but they also make system recovery much more difficult.

Separately, we also saw ransomware actors abuse poorly configured security solutions. Many out-of-the-box security products come with baseline/default policies enabled, but organizations often fail to configure these products specifically for their own network’s needs. Therefore, we saw many cases where ransomware operations were successful in environments where security policies were set to “audit-only” mode, meaning that the product only alerted an administrator to malicious activity but did not automatically block it.

We repeatedly noticed alerts generated for an initial compromise, followed by alerts on suspicious behaviors for privilege escalation and lateral movement, and finally for execution of a malicious payload, all without a single event being blocked or actioned. If solutions are deployed passively, a security team may have only a short time window to see an alert, validate if it’s a true positive, and mitigate the activity with a response.

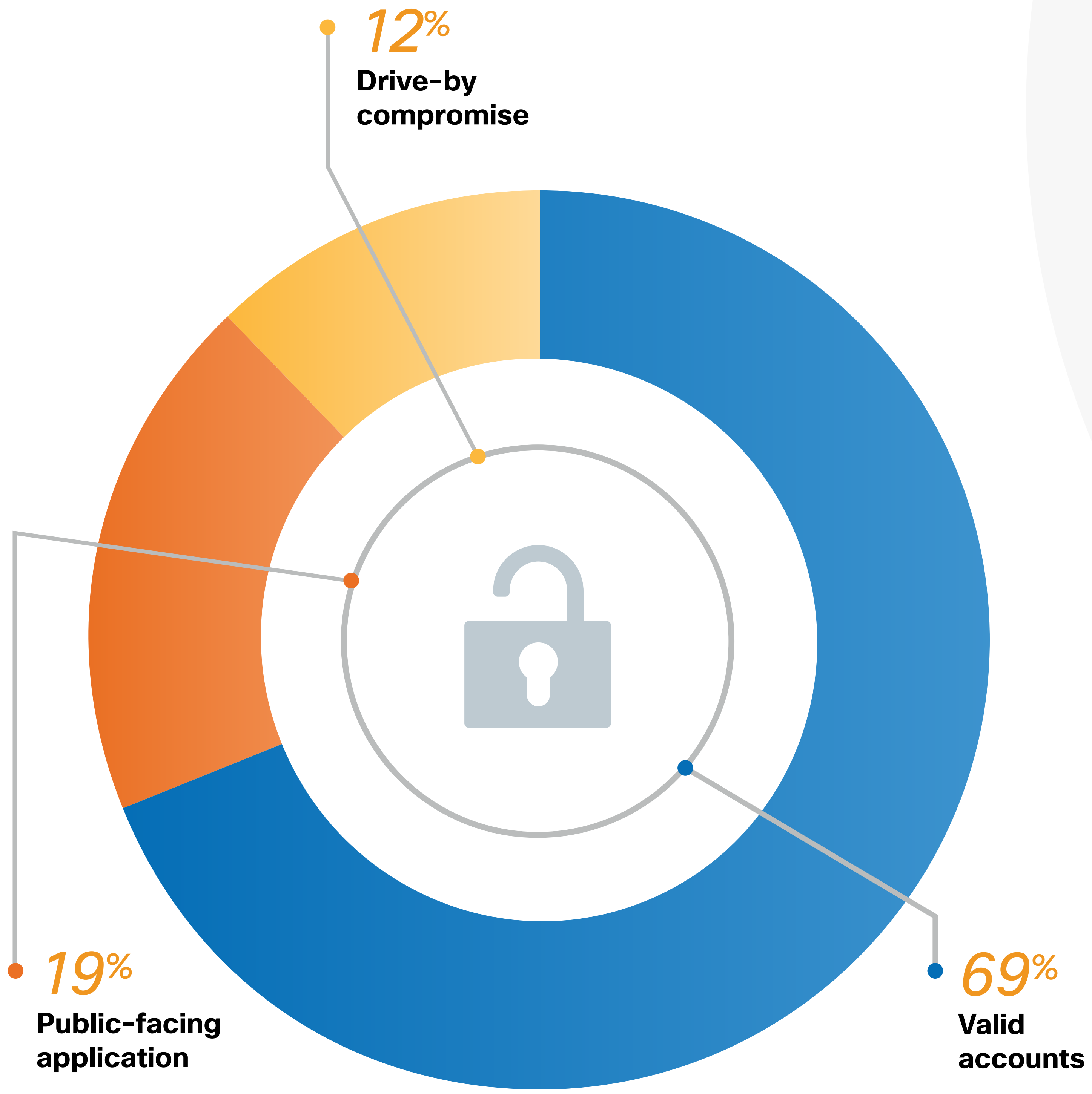
> *We often see organizations have deployed endpoint protection in a passive manner, meaning the product is producing alerts to the user but not blocking malicious activity.*

Initial access largely achieved via valid accounts

Ransomware actors overwhelmingly leveraged valid accounts for initial access in 2024, with this tactic appearing in almost 70% of related cases (see figure 16). As we outline in a later section of this report, actors are increasingly using identity-based attacks across the threat landscape, and with great success. In many cases, it's much easier and safer for adversaries to simply log in to legitimate user accounts using stolen credentials than to use more complex means like exploiting vulnerabilities or deploying malware. This tactic is facilitated in large part by the sale of compromised credentials on dark web forums, enabling ransomware actors to essentially buy their key into a targeted organization.

Ransomware actors exploited public-facing applications nearly 20% of the time. Public-facing applications can be accessed by anyone on the internet, not just internal users within a company, making this an incredibly vast attack vector. These include applications that support online shopping platforms, customer login portals, social media sites, online banking systems, email servers, customer service portals, and more. Attackers often exploit known vulnerabilities or misconfigurations to gain access. These types of attacks typically require more technical skill, with actors relying on techniques such as SQL injection, cross-site scripting (XSS), or remote code execution, which likely explains why we see this less often than the simpler method of compromising valid accounts.

Figure 16
Initial access



The prevalent use of valid accounts for initial access shines a light on the role of initial access brokers (IABs) in the ransomware ecosystem. Compromised credentials remain a valuable commodity, keeping IABs in business and streamlining adversaries' operations.

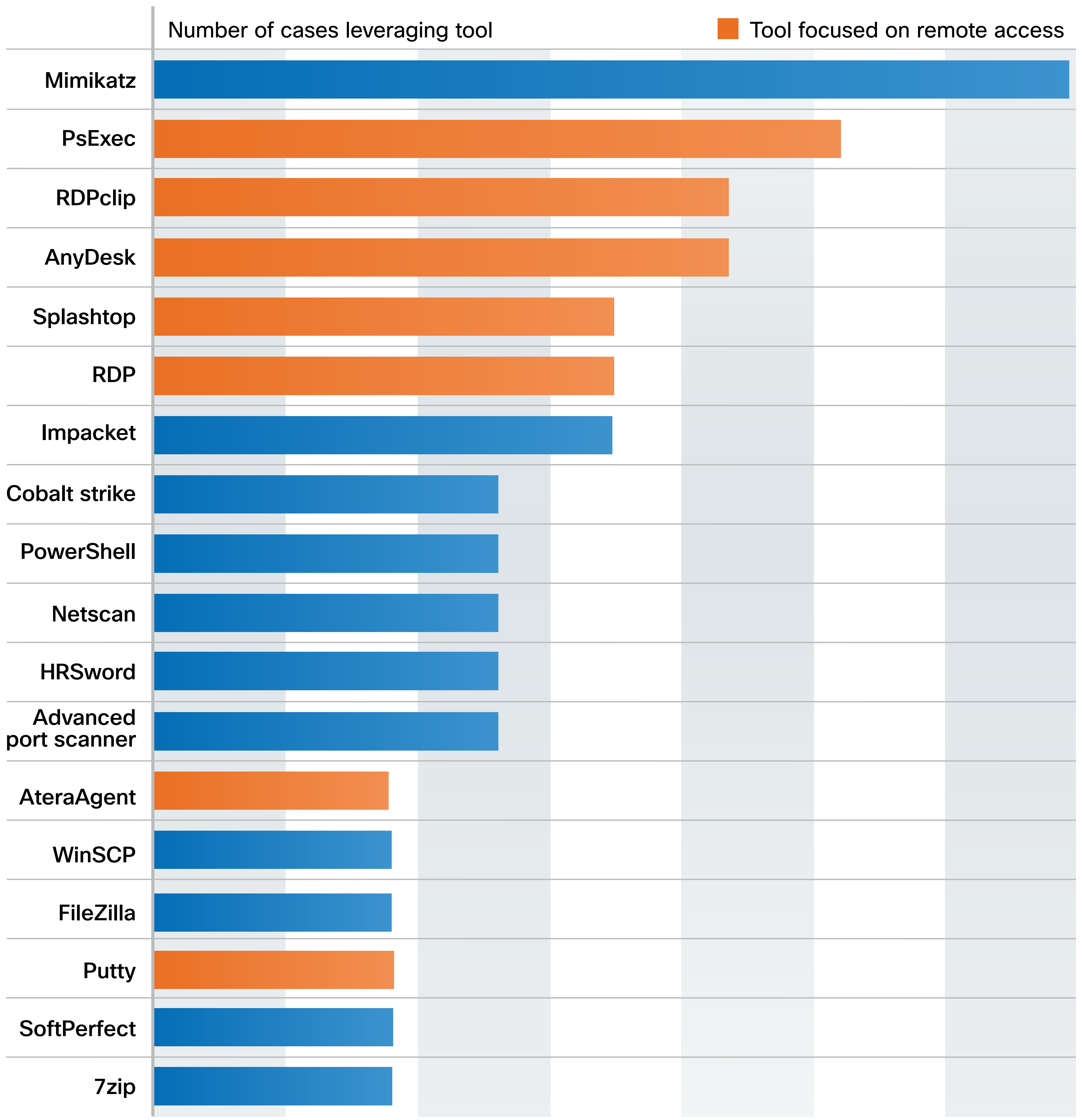
Actors rely heavily on remote access tools, commercial products, and LoLBins

Based on our review of the tools ransomware actors most frequently used in 2024, we saw a focus on remote access (see figure 17). Specifically, actors leveraged commercial products and LoLBins for command and control in their campaigns. Many organizations rely on legitimate remote access applications such as AnyDesk and Splashtop for daily operations, such as remote work or IT help, making detecting or blocking malicious use of these tools more challenging.

While effectively blocking all unauthorized remote management tools may be a challenge, security can still be greatly improved through policy and technical controls. Adopting just one or two approved remote access solutions and banning all others is a good practice, as security teams can also ensure the chosen solutions are thoroughly tested and deployed as securely as possible. Additional controls, such as auditing and blocking DNS queries associated with these tools, blocking hashes associated with remote access software installers, and employing an application allowlisting program can also be leveraged to mitigate this threat.

Figure 17

Top tools seen in Talos IR ransomware engagements



Ransomware operators impersonate IT personnel to gain remote access

Starting in November 2024, according to Talos IR observations, actors distributing BlackBasta and Cactus ransomware launched a campaign that leveraged social engineering to attain remote access to targets' computers.

The actors first sent a flood of email spam to a victim mailbox, then proceeded to call the victim a few days later, usually via Microsoft Teams, posing as IT support and offering help for the email flood issue. Targets were directed to initiate a Microsoft Quick Assist remote access session and to install the software if they didn't already have it on their system. Once the QuickAssist session was established, the adversary loaded tooling to collect information about the target system, establish persistence, elevate privileges, and ultimately deploy ransomware. BlackBasta ransomware was observed in earlier attacks, with the actors pivoting to Cactus ransomware later in the campaign.

This campaign underscores how organizations' reliance on remote access tools for legitimate purposes can be manipulated by adversaries. It also serves as a reminder for organizations to educate users on recognizing approved ways in which their IT personnel will engage with them.

Threat actor spotlight: RansomHub

RansomHub is a financially motivated RaaS group that has been increasingly active since at least February 2024. The ransomware is likely an updated version of Knight ransomware, which was for sale on underground forums in February 2024. RansomHub affiliates commonly leverage double extortion, encrypting a victim's data while also stealing information and threatening to publish it on their data leak site unless a ransom is paid.

RansomHub currently plays a significant role in the ransomware threat landscape. They have attracted affiliates associated with well-known ransomware groups LockBit and ALPHV, as well as Scattered Spider, a financially motivated cybercrime gang that previously used ALPHV ransomware for their operations. RansomHub typically targets large organizations, likely in pursuit of hefty payouts; the average employee count of organizations targeted in RansomHub incidents we responded to this year was over 18,000 employees.

In line with the trend detailed above, we observed RansomHub operators successfully uninstalling endpoint protection on compromised hosts, including critical servers, in the majority of RansomHub engagements this year, enabling them to quietly deploy their ransomware.

LockBit remained top player while newcomer RansomHub quickly ascended to the #2 spot

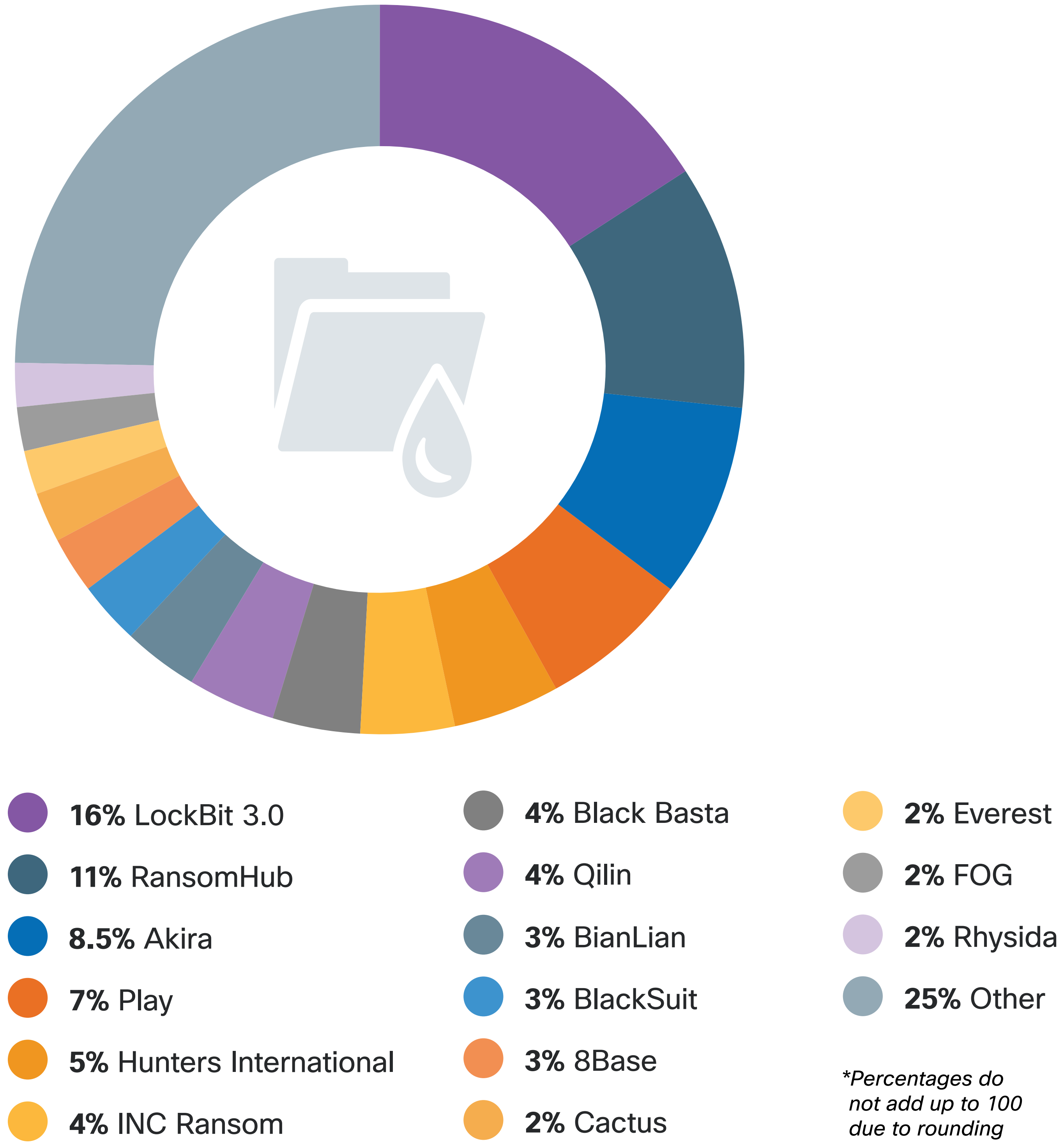
For the third year in a row, LockBit was the most active ransomware-as-a-service (RaaS) group, based on our monitoring of posts made to ransomware actors' leak sites. LockBit had the highest volume of posts (i.e., alleged victim compromises) among the 60+ groups we track, effectively claiming 16% of the market share in this crowded space (see figure 18). LockBit appearing as the frontrunner for the third year in a row is incredibly notable – in a dynamic space defined by constant change and the rise and fall of new ransomware groups, this type of longevity is unexpected. Moreover, LockBit was the target of a major law enforcement takedown operation in early 2024, but was able to rebound and quickly reconstitute, returning to normal activity levels soon after. Of note, LockBit's builder was leaked in September 2022, likely contributing to the ransomware's dominance as it expanded the pool of operators leveraging this encryptor.

Notably, newcomer RansomHub – a suspected successor of the Knight ransomware group that was first seen in February 2024 – followed close behind, accounting for 11% of posts.

In addition to RansomHub, Akira, Hunter's International, INC Ransom, Qilin, and BlackSuit ranked in the top ten for most active RaaS groups this year but not last year, demonstrating how dominance shifts quickly in this threat landscape. There are many plausible explanations for certain groups gaining momentum while others become more stagnant, such as rebranding of existing groups, source code leaks, dispute amongst operators, and law enforcement intervention.

Figure 18

2024 volume of posts made to data leak sites by ransomware groups



*Percentages do not add up to 100 due to rounding

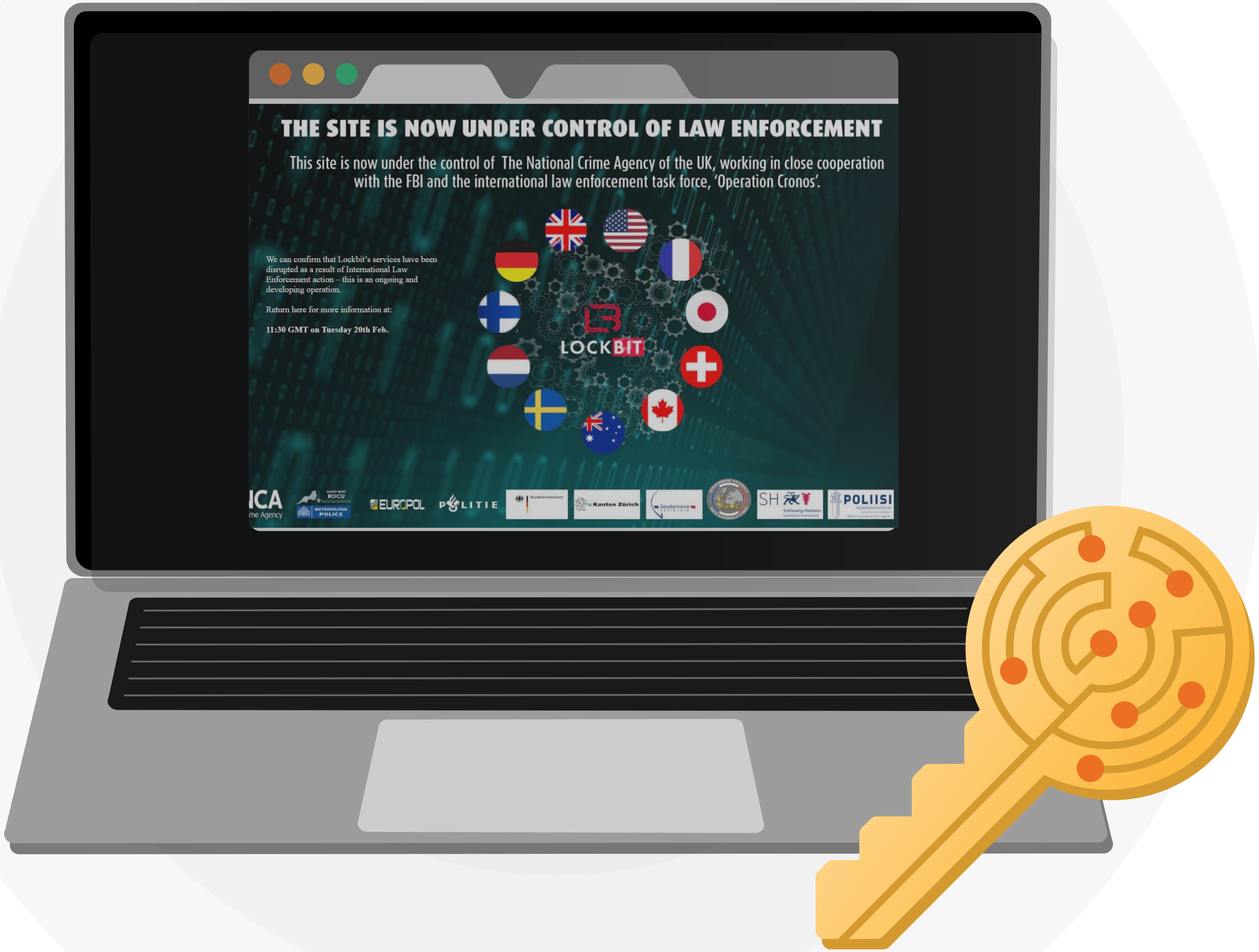
> *Many of the RaaS groups that ranked as most active this year did not make the top ten last year, demonstrating how dynamic this space is.*

Release of decryptor is the game-changer in disrupting ransomware gangs

2024 saw a number of disruptive operations led by law enforcement, with varying impacts on the targeted ransomware groups. One thing was clear, though – ransomware actors are far less likely to fully rebound from a takedown if associated decryption tools are made publicly available. ALPHV's dominance plummeted after an FBI disruption at the end of 2023. This group was ranked second in our 2023 report and dropped to 22nd this year. As part of this disruption, the FBI seized several websites operated by the group and offered a decryption tool to affected victims, enabling them to restore their systems. Though the group stood up new servers after the takedown, the decryption operations significantly impacted the group's revenue, and in March, administrators made the decision to shut down operations and declared their intent to sell their source code.

By contrast, the LockBit ransomware group was also targeted in a major takedown, but this operation did not include the release of a decryptor. Dubbed Operation Cronos, authorities in Ukraine, Poland, and the United States executed simultaneous actions against LockBit in February, taking control of key darknet infrastructure and arresting several affiliates. Though LockBit activity dropped– their posts on data leak sites went from 926 last year to 783 this year – they still emerged as the top actor in this space for the third year in a row.

Finally, in May 2024, Europol launched the largest-ever operation against malware loaders and botnets that support first-stage ransomware deployment, including IcedID, Smokeloader, SystemBC, Pikabot, and Bumblebee. This operation included the arrest of relevant targets, taking down of criminal infrastructure, and freezing of illegal proceeds. Nevertheless, we did not observe any notable dip in ransomware activity or in the overall volume of posts made to data leak sites, suggesting affiliates pivoted to using other tools and/or the malware's infrastructure was rebuilt. Further, the prevalent use of valid accounts this past year could mean ransomware operators are no longer relying on tools such as these for ransomware deployment.



THE SITE IS NOW UNDER CONTROL OF LAW ENFORCEMENT

This site is now under the control of The National Crime Agency of the UK, working in close cooperation with the FBI and the international law enforcement task force, 'Operation Cronos'.

We can confirm that Lockbit's services have been disrupted as a result of International Law Enforcement action – this is an ongoing and developing operation.

Return here for more information at:

11:30 GMT on Tuesday 20th Feb.

Identity-based threats

Identity attacks dominated the threat landscape in 2024

Identity was a common through line in 2024 across much of the data we looked at for this report. From initial access vectors to operational techniques further down the attack chain, threat actors relied heavily on identity-based attacks to power their operations. Adversaries are increasingly opting to compromise networks and accounts by simply logging in, rather than using more complex methods like exploiting vulnerabilities or deploying malware.

Identity-based attacks are attractive to threat actors because they can allow an adversary to carry out a range of malicious operations, often with minimal effort or without meeting much resistance from a security standpoint. This is due in large part to the activity being difficult to detect because it emanates from seemingly legitimate user accounts.

In addition to these types of operations being highly effective, there's also a major market for stolen credentials — which are often used in the early stages of an operation — with valid password and username combinations frequently traded on the dark web. This means that there is a strong financial incentive for cybercriminals to steal credentials for future sale, and it also underscores the ease at which bad actors can obtain access to stolen credentials for use in their own operations.

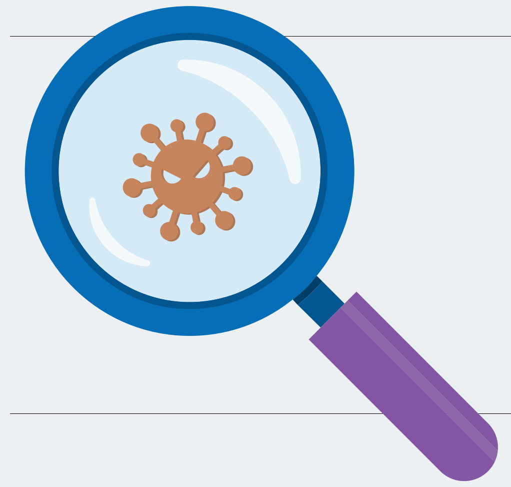
In addition to credentials or personally identifiable information (PII), illicit marketplaces on the dark web also offer tools-as-a-service specifically for performing identity-based attacks, as well as outsourced services to obtain specific data or accesses to certain victim networks. Here are some other findings we've seen in this space, based on our dark web research:

- Marketed stolen data includes plaintext credentials, particularly for email accounts; SSH credentials; financial data, like bank identification numbers (BIN) or credit card numbers; session tokens from browser caches; addresses; and more. Cyber actors may have acquired this data using one or several TTPs we outline later in this section, or by using malware like infostealers.
- Software and infrastructure are sold as-a-service, commonly in tiered subscriptions ranging from less than $50 to around $750 for tools specially geared towards credential theft, like phishing kits and infostealers. These tools have user-friendly interfaces and offer customer assistance, lowering the barrier to entry for novice cyber actors.
- Experienced actors advertise their services and can be hired to perform specific functions. They also auction off access to high-profile companies, which on average sell between $1,000 and $3,000.
- Bulk lists of credentials commonly sell for as little as $10 to $15 on dark web marketplaces.

Why are we seeing more identity-based attacks?

Growing attack surface

The use of web applications, cloud-based environments, BYOD policies, and SSO solutions have been on the rise in recent years, especially with the normalization of remote work. This, in turn, has increased the number of credential-enabled access points within a network that could be exploited by attackers.

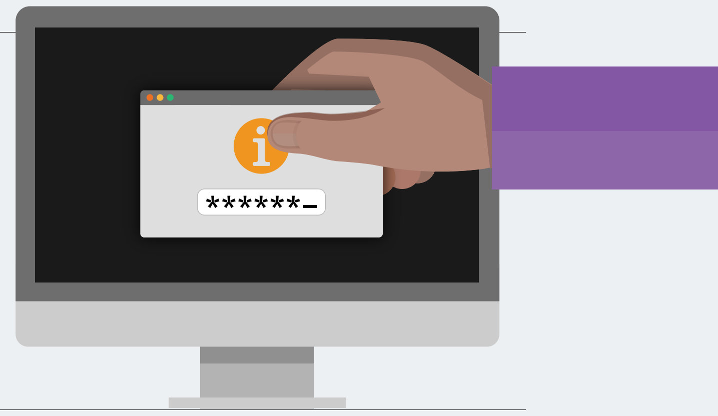


Hard to detect

Many of these attacks leverage legitimate authentication processes, making them hard to detect at the network perimeter. Moreover, once an attacker gains access, malicious activity emanating from a valid user's compromised account is more likely to go unnoticed.

Easy to carry out

Attackers can easily obtain stolen credentials, often via the dark web and previous data breaches. Additionally, identity-based attacks largely rely on social engineering rather than technically sophisticated means.

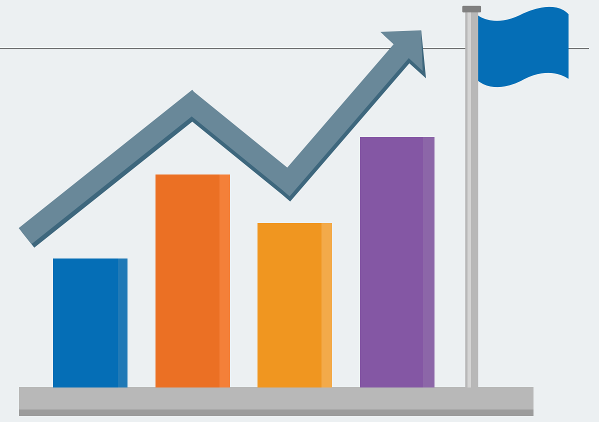


Enables other operations

In addition to gaining initial access to a target device, threat actors can continue to use identity attacks throughout their operations to escalate privileges, move laterally, conduct internal social engineering attacks, and more.

Achieves significant access

Using relatively simple means, actors can beat identity-based security challenges and gain access to the Active Directory, where an entire organization's access and permissions are managed; cloud applications that power daily operations; or even IT networks and operational technology (OT) systems-crucial components of any organization's cybersecurity.



What is an identity attack?

An identity-based attack targets the unique digital identity of a user, organization, or machine to access data or networks. Digital identities encompass much more than just usernames and passwords for valid accounts. To obtain initial access, actors exploit a range of identifiers, like digital certificates, API keys, encryption keys, session tokens, and more.

Login credentials

Cleartext and plaintext passwords, usernames

API key

A unique identifier used to authenticate and authorize a user or calling program to an API. Used for security purposes and for monitoring/limiting usage.

Digital certificate

An electronic file that verifies the identity of a user, device, or server

Encryption key

A string of random bits that scrambles and unscrambles data. Used in secure connections like SSH, HTTP, and Telnet.

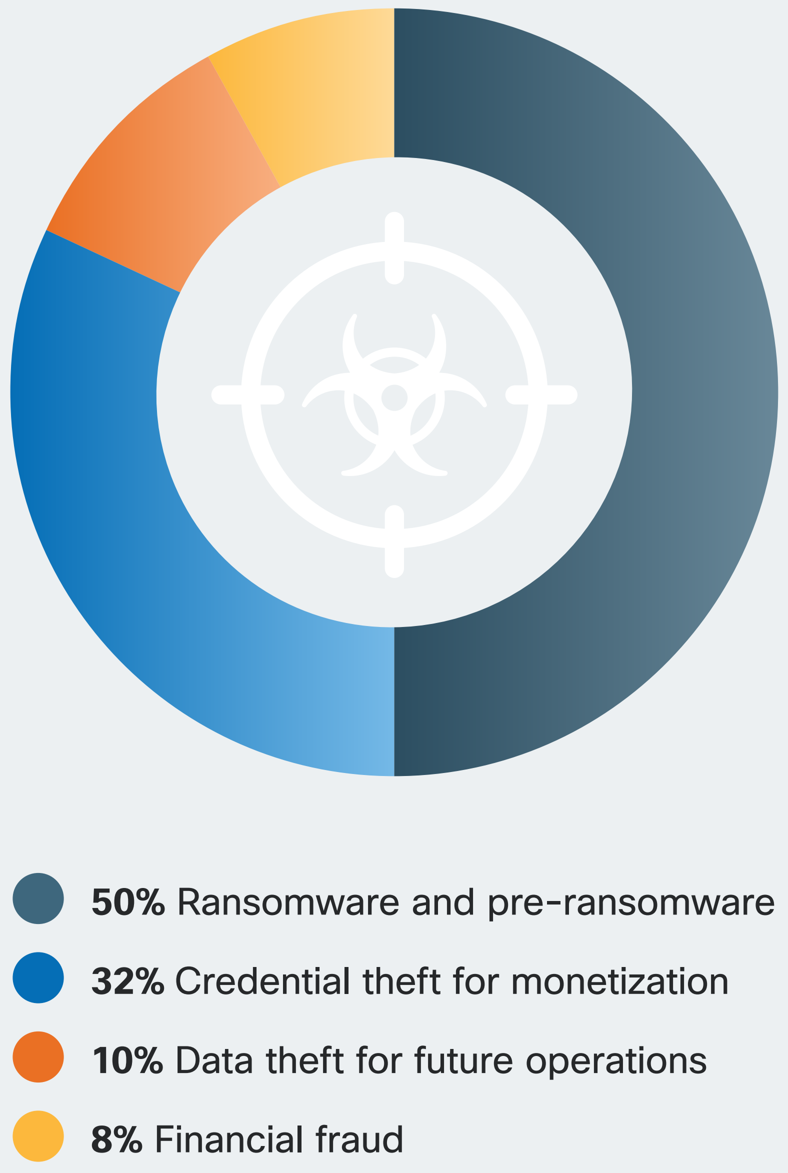
Session ID

Identifies a user's session on a website or application. If stolen, an attacker can access resources as a legitimate authenticated user.

Identity attacks in 2024

We have seen a strong shift toward identity-based attacks in Talos IR incidents. In 2024, the most common technique used to gain initial access was valid accounts, making this the top access vector for the second consecutive year.

Adversaries’ goals in identity attacks



60%

More than half of Talos IR cases had an identity attack component in 2024.

44%

Nearly half of all identity attacks targeted the Active Directory. Another 20% targeted cloud applications.

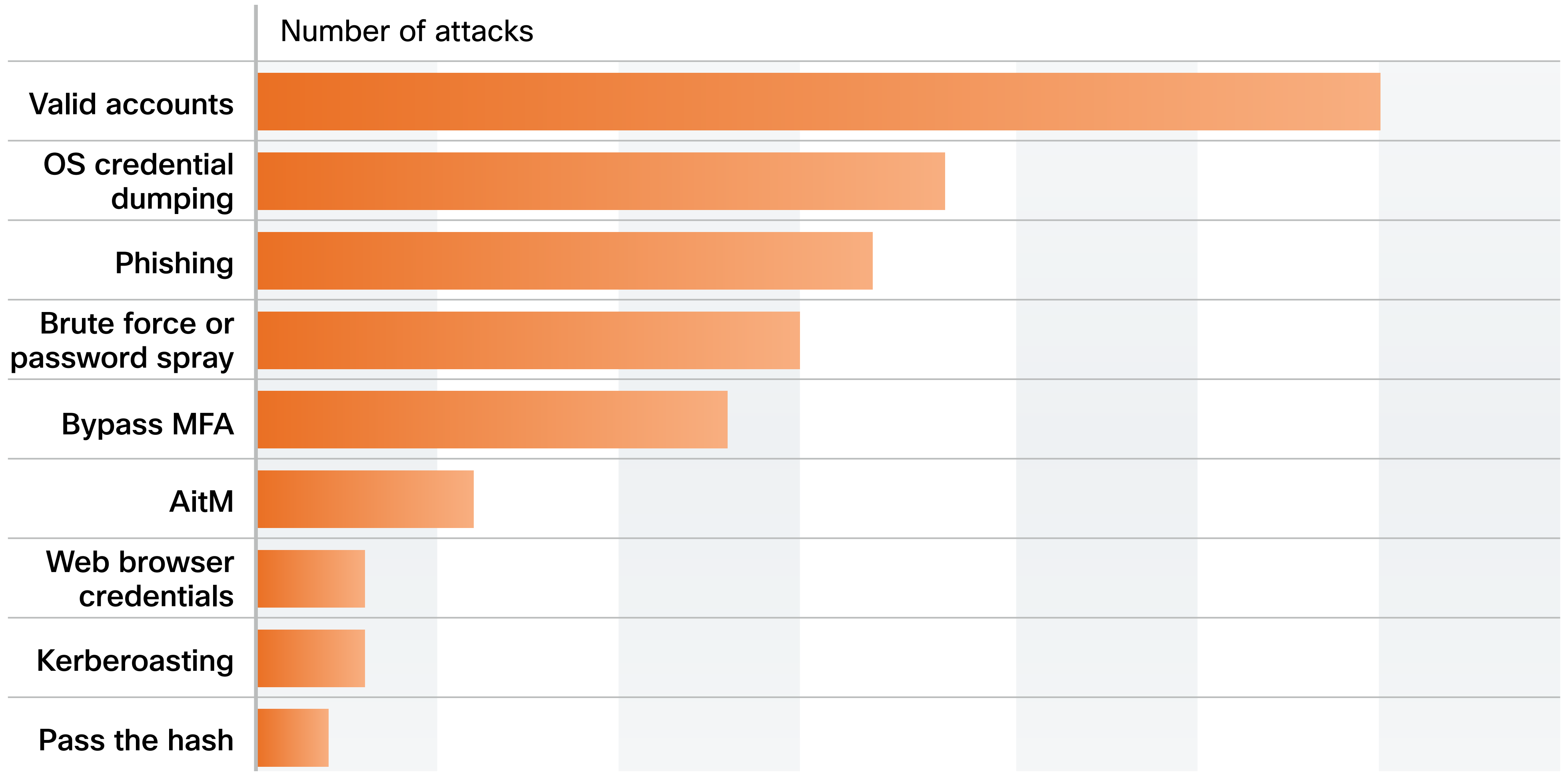
Identity attacks omnipresent throughout the attack chain

The most common tactic we observed in Talos IR engagements was the use of valid accounts – typically seen in the initial access phase – where adversaries obtained and abused credentials of existing accounts to carry out various phases of their operations. OS credential dumping was also extremely common. While the majority of actors targeted credentials in LSASS memory and Active Directory, we also saw a variety of other techniques in this threat category, including attempts to extract credentials from the Security Account Manager (SAM) database, attempts to access cached domain credentials, the use of a technique called DCSync to abuse a Windows Domain Controller’s API, and attempts to access Local Security Authority (LSA) secrets–which can contain a variety of different credential materials–which adversaries can obtain with system access to a host.

Based on our assessments of threat actor intentions, we found that half of all identity-based attacks were related to ransomware and pre-ransomware operations. Actors were also frequently motivated by their intent to sell stolen credentials for a profit, such as with initial access brokers (32%), stealing credentials for espionage purposes or to enable future operations (10%), and financial fraud, such as stealing credit card data or conning victims into sending money (8%).

Figure 19

Types of identity attacks observed in Talos IR



Identity in-the-wild: Compromising Active Directory

As mentioned above, 44% of all identity-based attacks seen in Talos IR incidents targeted Active Directory, a widely used Microsoft service for Windows. Active Directory holds critical user information like usernames, passwords, and access permissions, making it a gold mine of high-value data for attackers. Moreover, according to a recent government report, Active Directory is the most widely used authentication and authorization solution in enterprise IT networks globally.

Adding to the risks around Active Directory being such a high-value target for attackers, organizations often fail to properly secure these environments. In many of the Talos IR cases involving compromised Active Directory, successful attacks occurred in enterprise environments that had misconfigured security products and/or policies inconsistent with industry-recommended best practices.

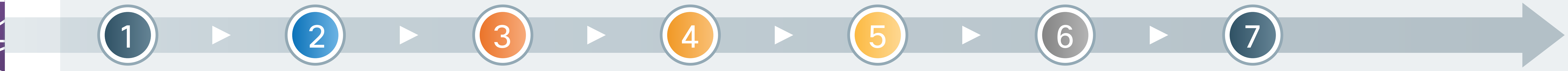
Active Directory holds critical enterprise user information and is also the most widely used identity and access management (IAM) solution globally, underscoring why actors targeted this service in nearly 50% of all identity-based attacks seen in Talos IR cases.

Case study: How adversaries leverage AD to disrupt data centers and critical services

In August 2024, a Cisco customer in the manufacturing sector reported that multiple endpoint detection and response (EDR) solutions had unexpectedly been uninstalled from servers hosted in the organization's managed data center, including two domain controllers, potentially indicating threat actors had full Active Directory domain access. In this investigation, Talos IR observed evidence suggesting the actor had compromised the Active Directory in preparation for deploying ransomware. The adversary leveraged ADExplorer, a utility that is part of the suite of the Sysinternals admin tools, to browse the different domains in the environment and dump the Active Directory database.

In this case, we saw the attacker use identity-based attacks in the initial stages of their operation, showing how affective these techniques can be. We also saw how initial access to the Active Directory was essential to kicking off the broader attack, and how that type of access can enable the deployment of high-impact threats like ransomware.

Case study: Active Directory attack



1. Extract local passwords and password hashes within AD database
2. Escalate to privileged admin accounts
3. Reset passwords; create new accounts to maintain persistence
4. Lateral movement to domain controllers and use of Mimikatz
5. Installation of backdoors and other software to maintain persistent access
6. Access backup systems
7. Result in pre-ransomware TTPs which could have eventually led to deployment of ransomware if not actioned swiftly

We frequently observe accounts (i.e., user, admin, and service) with excessive or incorrect privileges, accounts with weak or default passwords, flat network architectures, and missing or misconfigured MFA. Our recommendations for mitigating Active Directory compromises are in line with [CISA's strategies] to mitigate the 17 most common techniques used by adversaries and malicious actors to compromise Active Directory.

Identity in-the-wild: Compromising cloud services providers' APIs

Attacks targeting the cloud are also on the rise, with 20% of identity-based compromises impacting cloud applications, according to Talos IR findings. Cloud APIs are necessary to facilitate seamless communication, integration, and data transfers between a wide range of cloud services and between cloud and on-premise applications.

APIs are attractive targets because they can provide direct access to sensitive data and critical application functionalities, as they are used in software designed to support users and companies across all business verticals. Cloud APIs are also inherently difficult to manage due to their sheer number, diverse functionalities across different cloud providers, and the need to constantly monitor and update them to keep pace with evolving cloud services. Moreover, many cloud APIs are publicly accessible, making it easy for attackers to discover and test potential vulnerabilities.

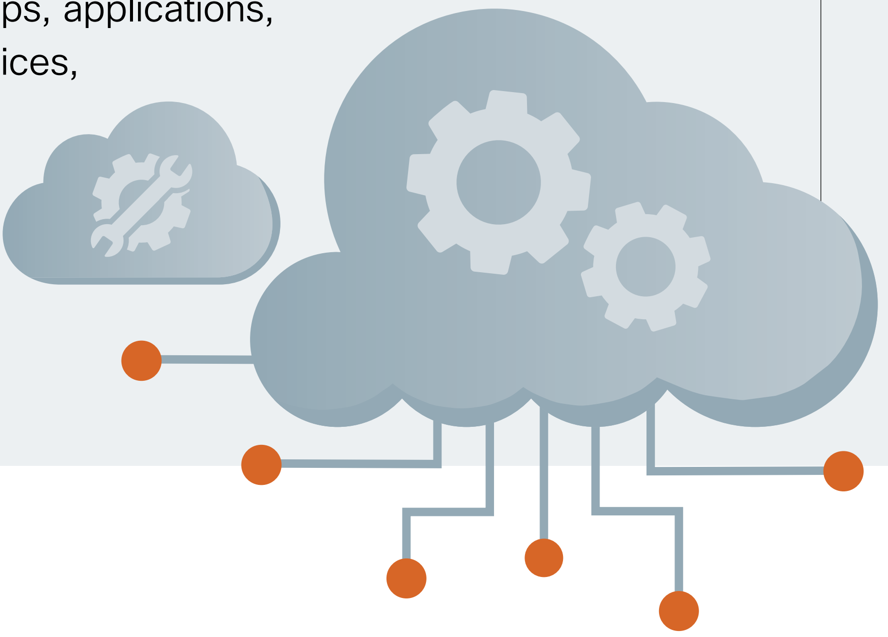
Here is an example of how threat actors could leverage many of the identity-based attack techniques to compromise cloud APIs, based on our experience in responding to these types of incidents.

Step 1: Access cloud API using compromised digital IDs

- For initial access, a threat actor could leverage stolen API keys or session cookies to bypass MFA and pivot to the cloud environment. In the case of Microsoft Entra ID, a cloud-based IAM service, a stolen Primary Refresh token can be leveraged to maintain access and sign in to services across the Microsoft cloud.
- An actor could also steal credentials via phishing, deploying infostealers, or, if the API relies on weak or easily guessable credentials, using brute force techniques like password spraying or credential stuffing.
- Note that threat actors have found success in repurposing traditional techniques to compromise cloud environments.

Step 2: Use a cloud-specific tool to enumerate data

- Skilled actors have created tooling that is freely available on the open web, easy to deploy, and designed to specifically target cloud environments.
- Some examples include ROADtools and AAAInternals, publicly available frameworks designed to enumerate Microsoft Entra ID environments. These tools can collect data on users, groups, applications, service principals, and devices, and execute commands. ROADtools can also work with custom plug-ins to query and analyze data.



Step 3: Execute commands for post-compromise activity

- With proper permissions, actors may abuse cloud APIs to execute commands by leveraging the cloud provider's command line interface, which is intended to be used to manage cloud resources.
- Actors could execute commands to setup backdoors or create reverse-shell connections for persistence or to exfiltrate data.

Step 4: With this level of access, there is the potential for widespread disruptive and destructive attacks.

- Actors could steal data that is handled by APIs, which includes PII like social security numbers, financial data like credit card information, health-related data like medical records, intellectual property, private correspondences, or user activity data like browsing history or physical locations.
- Potential attacks:
  - Use stolen identities to impersonate victims for financial theft.
  - Use stolen accounts to conduct business email compromise (BEC) attacks against third parties, including customers and trusted business partners.
  - Disrupt business operations, possibly leading to delays in critical services.
  - Remove users' access to their accounts.
  - Conduct ransomware or data theft extortion operations, possibly leading to financial loss, reputational damage, and government compliance violations.

Step 5: Evade detection

- Attackers can attempt to evade detection by deleting or modifying logs, burying malicious activity by mimicking legitimate API traffic, reverting changes they made to a cloud instance, and more.
- Evasive measures can be taken at any point in the attack lifecycle.

Attacks against MFA

In partnership with
CISCO DUO

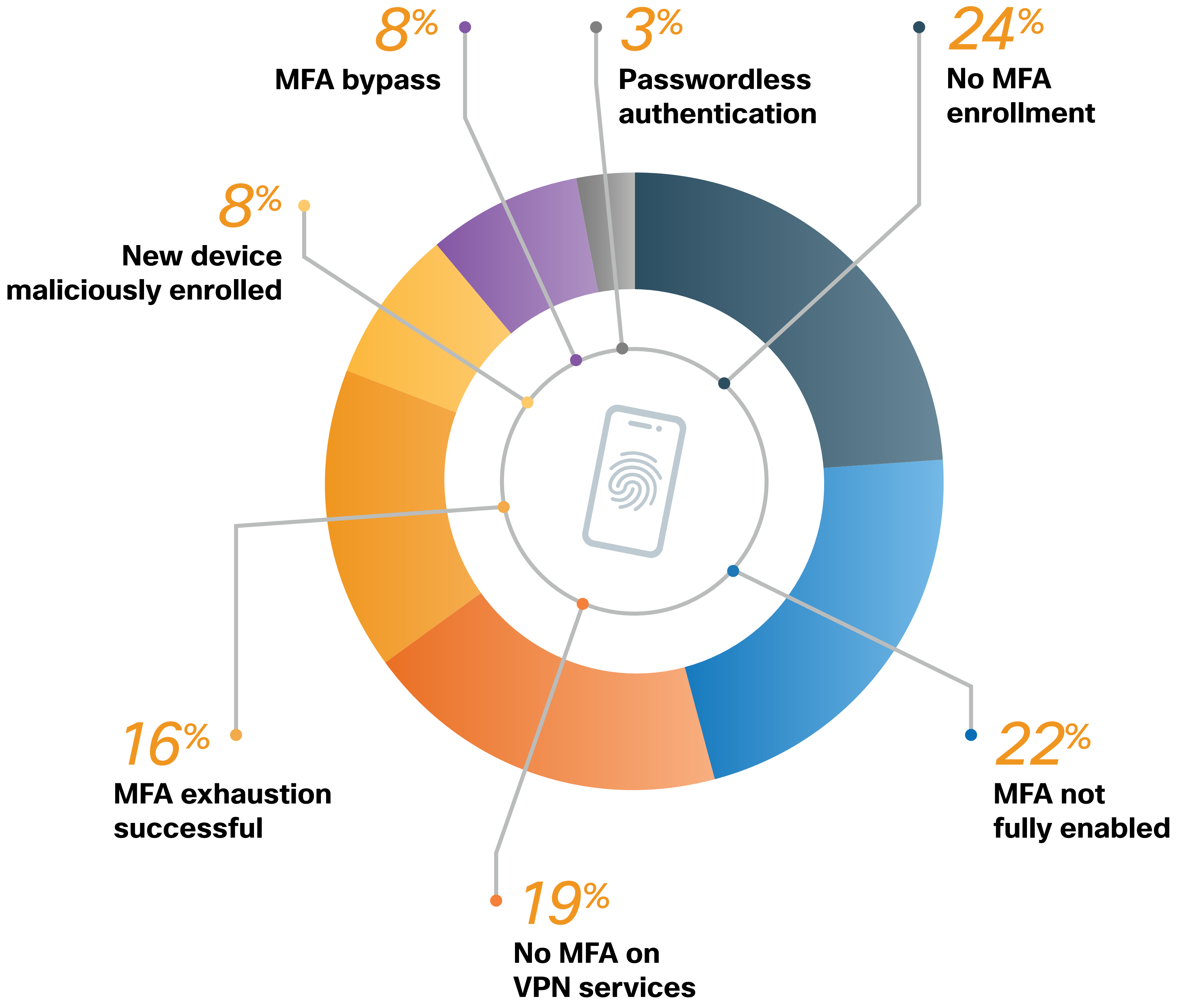
Threat actors capitalize on a variety of MFA weaknesses

A key way in which actors can compromise a user's identity, as mentioned earlier, is by targeting the multi-factor authentication (MFA) process. Given the amount of activity in this space, and the role that we see it play in attacks that could have been prevented, it made sense to devote an entire section of the report specifically to MFA attacks. Here, we explore the threats facing MFA, a key component of the rise in identity attacks we saw in 2024.

MFA weakness was the leading security weakness in Talos IR data this year, an enduring trend year over year. Lack of MFA enrollment made up a quarter of the MFA issues observed; however, we saw a variety of other ways in which MFA was insufficiently deployed this year, enabling threat actors to gain access to key resources and establish persistence in targeted networks.

Figure 20

Observed MFA weaknesses in Talos IR cases



Though MFA is proven to provide strong security, some organizations may choose not to employ it, given the cost and complexity of knowing which systems and resources to defend with it.

Talos IR observations point to four top security practices to guide MFA deployment:

1. Enable MFA on VPN services: We consistently observed threat actors taking advantage of a lack of MFA on VPN services. Organizations, particularly those whose employees use VPNs to access corporate networks, should prioritize requiring MFA to access their VPNs.

2. Enact user education and MFA prompt thresholds: Threat actors also leveraged MFA exhaustion attacks, also known as MFA fatigue, by sending repeated requests to authenticate until the user finally accepted one. This attack can be mitigated by improving user education as well as limiting the number of failed MFA requests from a single IP address or device.

3. Implement higher security factors: Organizations should implement additional security measures, such as "challenge-response authentication," where a user must provide a valid answer to a question. Examples of this include security questions, such as "what is your maiden name," or "where did you go to high school;" and CAPTCHA, where an image is presented to the user who then must enter the characters they see to verify they are human. This creates another layer of challenges one must pass to gain access to the desired information or digital assets.

4. Conduct robust monitoring of device registrations: Security teams should continuously monitor and log when new devices are enrolled in MFA, and/or require new users to authenticate through an additional method before the new device is enrolled. Some MFA products, like Duo, provide logging for organizations. In many cases, we saw actors successfully add their own authentication device to victims' MFA systems, allowing them to operate under the radar.

MFA attackers go straight for IAM applications

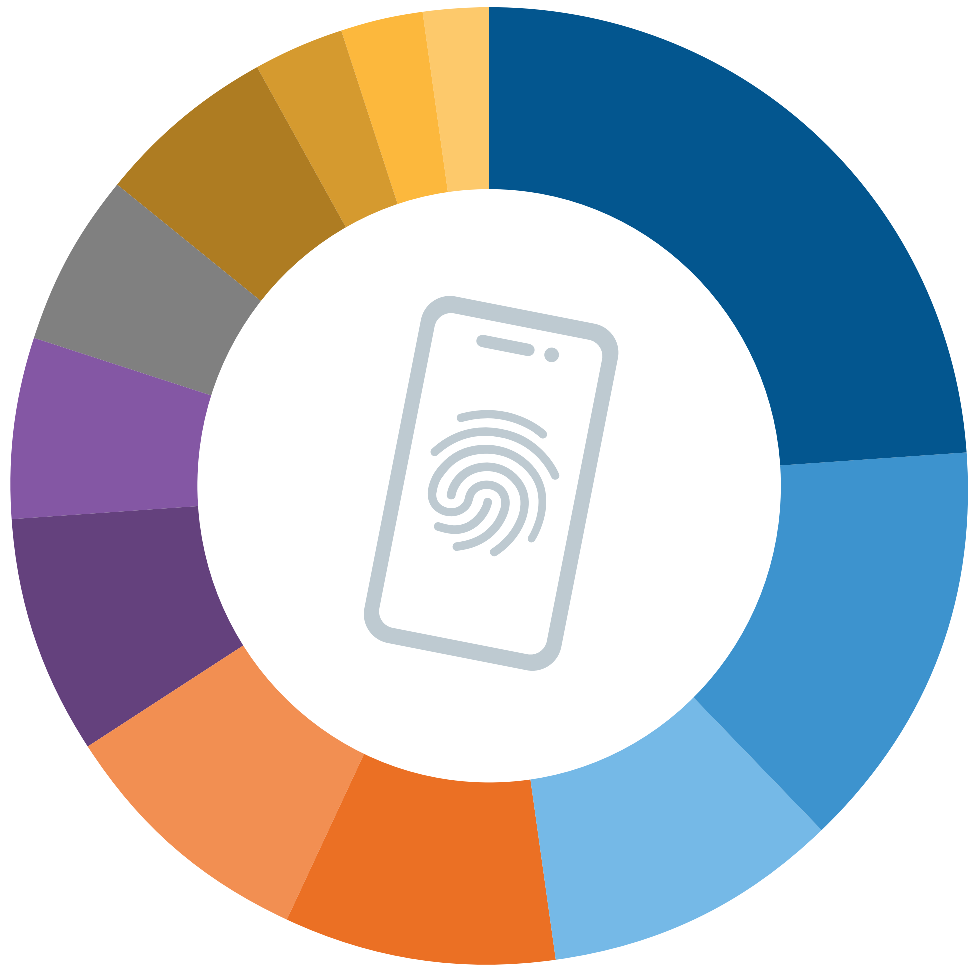
Based on Cisco Duo data, IAM applications were most frequently targeted in MFA attacks, accounting for nearly a quarter of related incidents. IAM applications, combined with network security, authentication and networking, and remote access applications, accounted for more than 50% of incidents where attackers targeted MFA deployments (see figure 21).

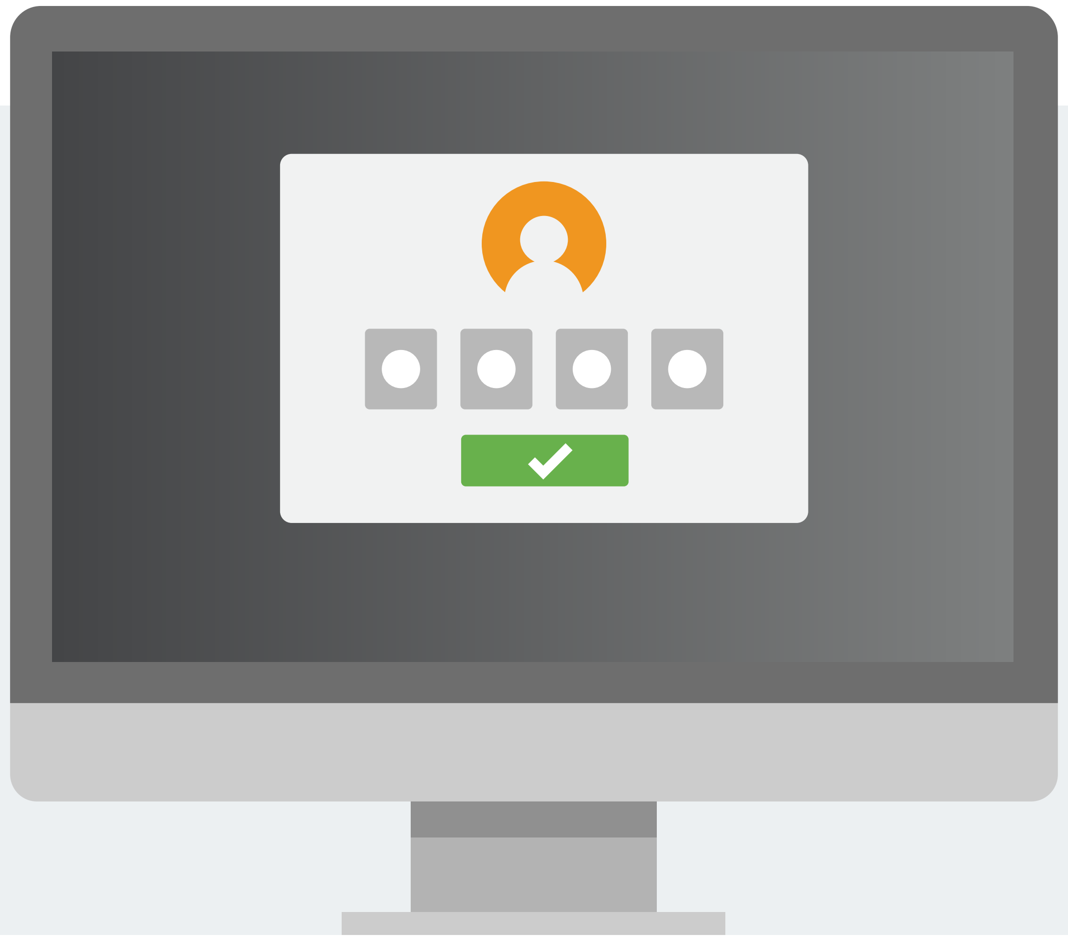
The most commonly targeted IAM applications are listed below. A variety of vendors were targeted, including Citrix, Microsoft, Fortinet, Palo Alto Networks, Cisco, and F5, which is not surprising given the widespread use of these companies' products globally.

Figure 21

Types of applications targeted in MFA attacks

- ● **24%** Identity and access management (IAM)
- ● **14%** Network security
- ● **10%** Authentication and networking
- ● **9%** Remote access
- ● **9%** Application delivery and security
- ● **8%** Software development kit (SDK)
- ● **6%** Authentication protocol
- ● **6%** Authentication and network access control (NAC)
- ● **6%** Cloud security and authentication
- ● **3%** Operating system
- ● **3%** API communication protocol
- ● **2%** Email and collaboration





Applications most frequently targed in IAM attacks

Shibboleth

An open-source SSO solution widely used in education and research institutions.

Central Authentication Service (CAS)

An open-source SSO solution that provides secure authentication for web applications.

Active Directory Federation Service (ADFS)

A Microsoft solution for SSO and identity federation.

Duo Central

Duo's centralized web portal that users can visit to get access to their organization's applications.

Microsoft 365 SSO

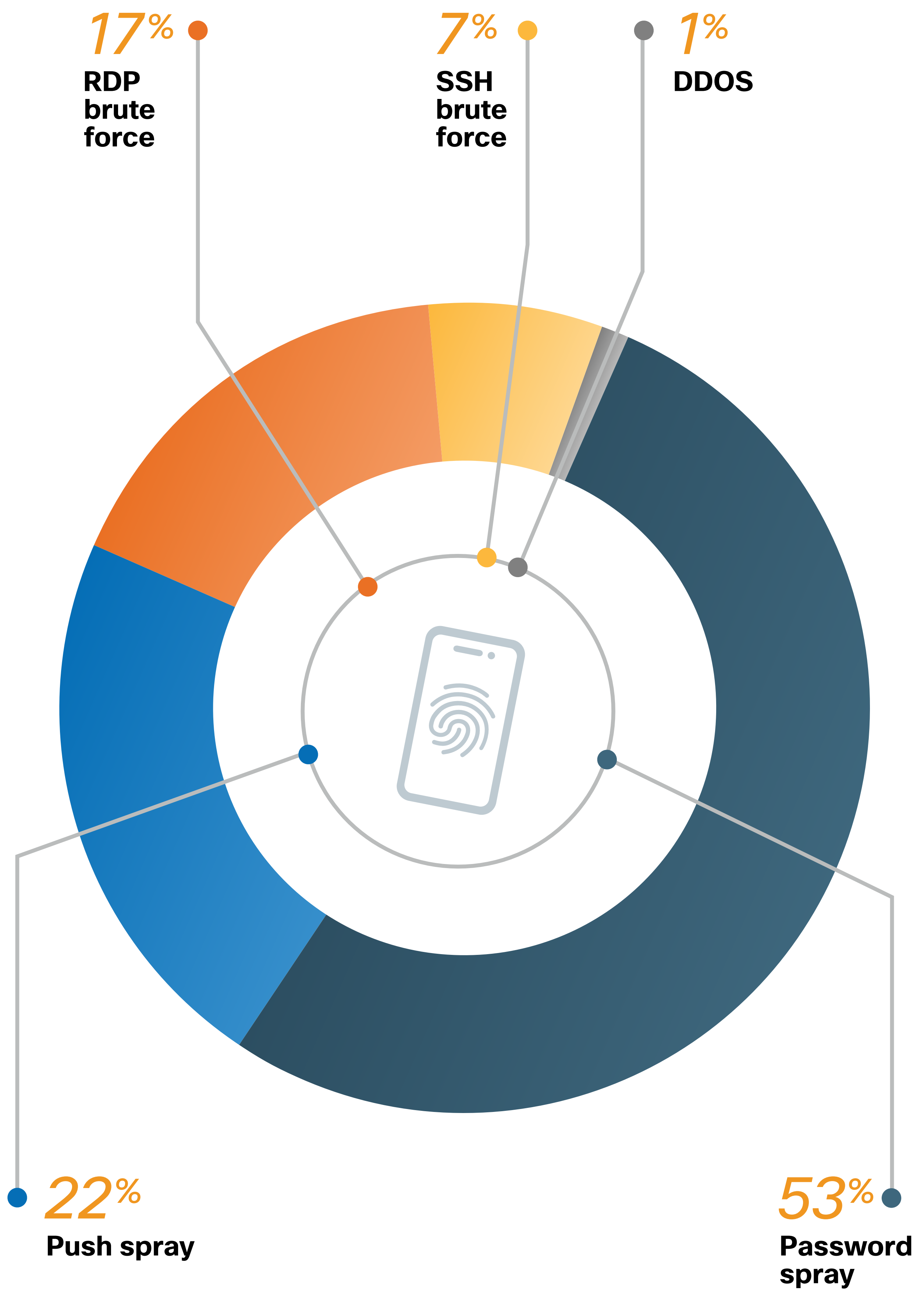
SSO for Microsoft 365, enabling unified authentication for Microsoft services.

High volume, easily preventable spray attacks are most common

Attackers are probing for organizations lacking MFA or with MFA incorrectly configured. Password spray attacks, in which an adversary tries common passwords to access many accounts, were the most frequent type of threat we observed against MFA-protected applications. However, MFA is highly effective at mitigating brute force and password spray attacks due to the additional authentication measure that is required, which often results in lower success rates for these types of campaigns.

Push spray was the second most common attack type. This technique, also known as MFA "bombing" or "fatigue," goes beyond the simple password guessing approach represented by spray attacks. Threat actors flood a victim's device with MFA push notifications prompting them to confirm/accept the login request in hopes that the victim will eventually relent and unwittingly grant the adversary access.

Figure 22
Types of MFA attacks

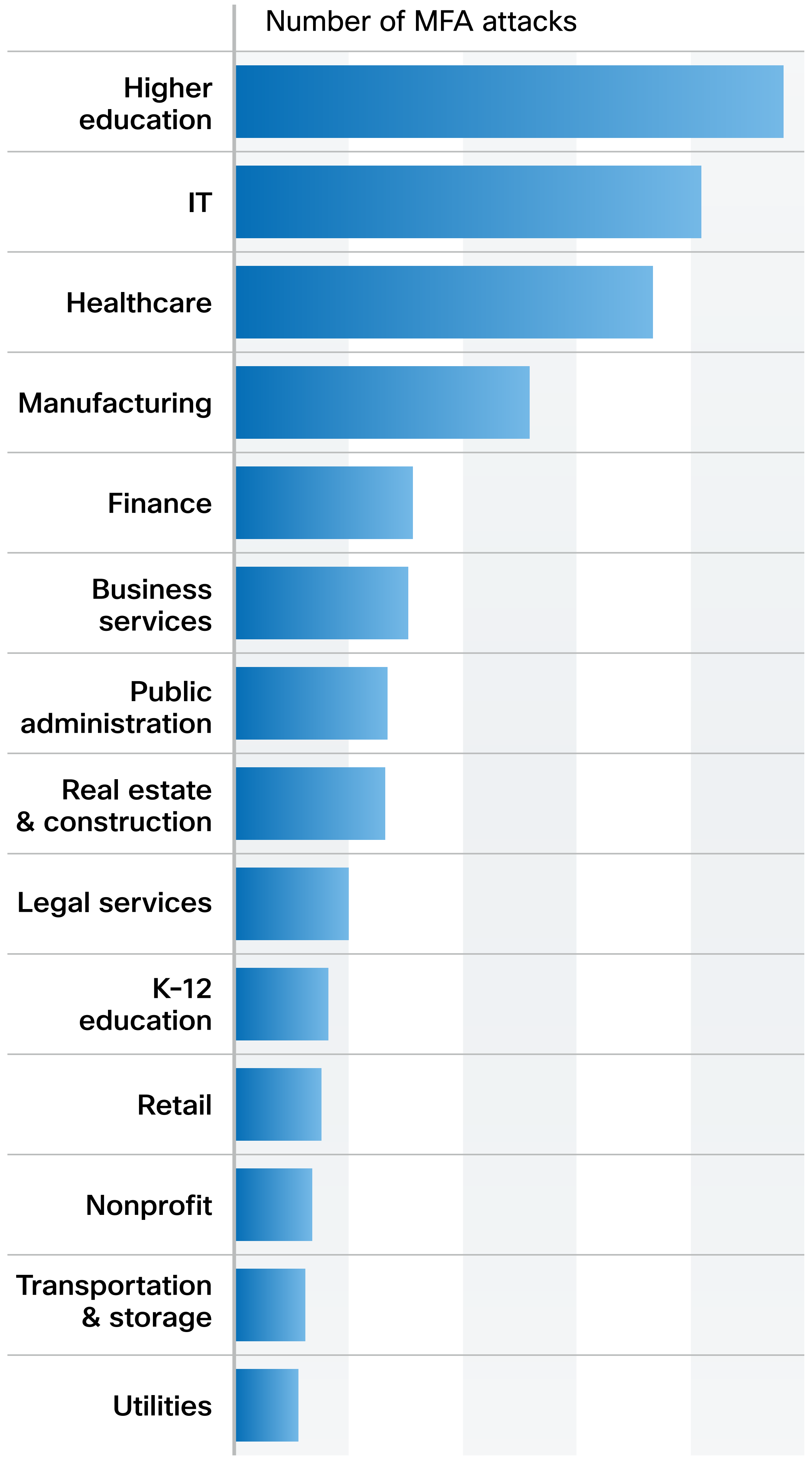


Top-targeted verticals align with broader attack trends

It is no surprise that the most targeted industries in MFA attacks were education and healthcare, as these sectors have consistently been among the top-targeted in our Talos IR findings across all attack types as well as ransomware campaigns specifically (detailed earlier in this report).

When looking at the types of impacted education entities, colleges and universities were targeted six times more frequently than K-12 schools. Higher education institutions can be ideal targets for data theft and password harvesting for several reasons: 1) There is a wider attack surface, as high volumes of students are accessing university resources via their mobile devices; 2) Students with access to this information are likely easy targets because they may not employ good cybersecurity practices; and 3) The applications colleges and universities use are more likely to store sensitive information about their students, including social security numbers, billing and payment information, driver's licenses, and other PII that is typically not collected by primary, secondary, and high schools.

Figure 23
Top-targeted industries in MFA attacks



MFA in-the-wild: Phishing and device compromise lead to major breach at large university

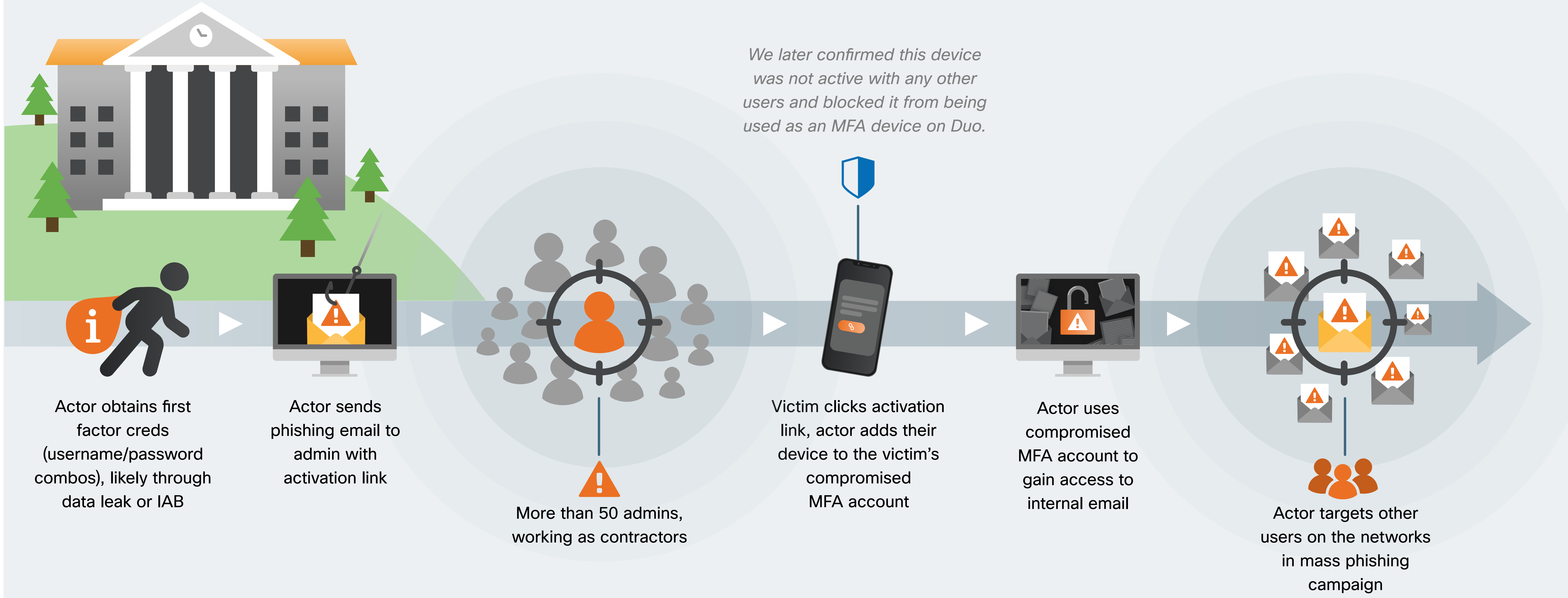
The following case study is an example of how we see the above trends play out in everyday scenarios. In this incident, a large university with more than 100,000 users was the victim of both phishing and device compromise.

Based on our investigation, the threat actor already had stolen credentials (login/password combinations) for targets at this organization, which they had likely purchased from an initial access broker (IAB) or obtained from a separate data leak. The actor sent phishing emails to a system administrator and tricked them into clicking an authentication link that added the attacker's device to the victim's MFA account. From there, the adversary was able to send internal phishing emails to several other users on the network.

One major security pitfall for the organization was that it had 50 administrator accounts — a significantly high number given the sensitive access admins have. Moreover, all of the admins were contractors, which is not a security best practice, and we know from our own research that threat actors often prefer to target contractors over account holders with comparable privileges.

Case study: MFA attack against university

Victim: Large university
(100,000+ users)


Actor obtains first factor creds (username/password combos), likely through data leak or IAB
Actor sends phishing email to admin with activation link
More than 50 admins, working as contractors
Victim clicks activation link, actor adds their device to the victim's compromised MFA account
We later confirmed this device was not active with any other users and blocked it from being used as an MFA device on Duo.
Actor uses compromised MFA account to gain access to internal email
Actor targets other users on the networks in mass phishing campaign

2024

YEAR IN REVIEW

AI threats

Overview of the AI threat landscape in 2024 and 2025

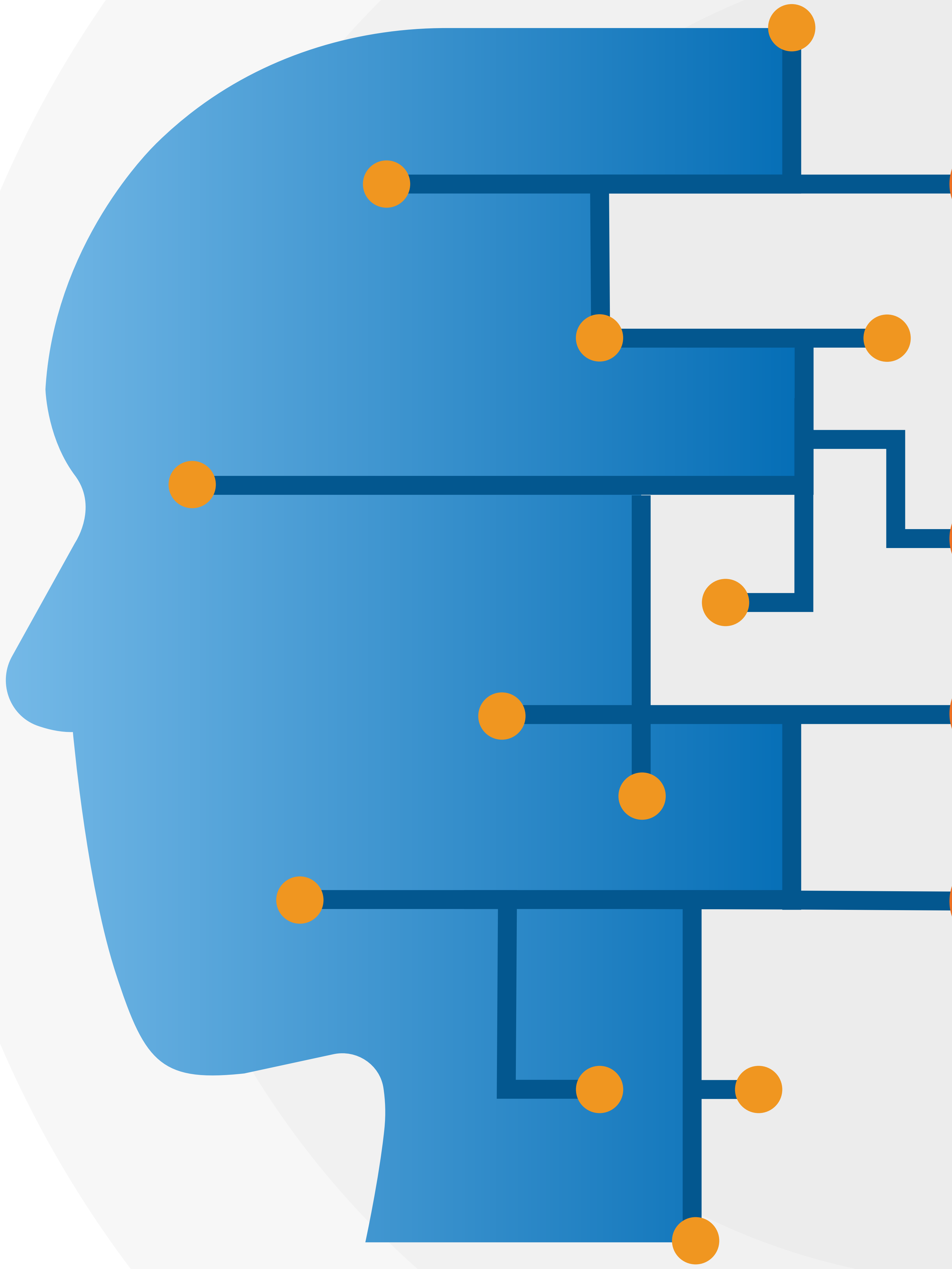
2024 brought the continued proliferation of artificial intelligence and machine learning (AI/ML) applications, as well as various business integrations and tools. Meanwhile, in cybersecurity, service providers have increasingly integrated AI into their products and workflows to enhance threat and vulnerability detection, automate responses, and bolster organizations' overall security postures. While the advancement and adoption of AI/ML technology has paved the way for copious new business opportunities, it also complicates risk and threat environments. Cisco's Robust Intelligence team – the threat researchers and developers behind Cisco's new AI Defense security solution – is watching this space closely. Here are the potential AI-based cyber attacks they are most worried about as we look ahead:

- Cybersecurity risk to AI systems, applications, and infrastructure;
- Data exfiltration, tampering, accessibility risk from AI models; and
- Use of AI to automate and professionalize threat actor cyber operations, particularly in social engineering

While these types of threats might be on the horizon for 2025 and beyond, 2024 mainly saw AI enhance existing malicious tactics, rather than aid in the creation of new ones.

Threat actor use of AI off to a slow start in 2024

Generative AI is powerful and its potential to influence the threat landscape is staggering, but in 2024, threat actors' use of AI did not significantly enhance attackers' TTPs. Although threat actors have the potential to harness AI and develop novel capabilities, we have not yet observed those capabilities deployed at scale in-the-wild. In the meantime, we have observed both state-sponsored adversaries and cybercriminals use AI for 1) social engineering, and 2) task automation and other productivity improvements in the threat actors' attack lifecycle.



How threat actors could leverage AI in 2025

We predict the following developments in 2025:

The rise of agentic AI:

Agentic AI, "AI systems and models that can act autonomously to achieve goals without the need for constant human guidance," could imperil organizations that are neither prepared nor equipped to handle agentic systems and their potential for compromise. As agentic systems increasingly integrate with disparate services and vendors, the opportunity for exploitation or vulnerability is ripe. Agentic systems may also have the potential to conduct multi-stage attacks, find creative ways to access restricted data systems, chain seemingly benign actions into harmful sequences, or learn to evade detection by network and system defenders.

Continued social engineering at scale:

From social engineering to propaganda proliferation, cybercriminal and state-sponsored actors will continue to leverage AI technologies to improve the personalization and professionalization of their malicious activities.

Automated vulnerability discovery and exploitation:

Threat actors could use AI to uncover vulnerabilities, including zero-day exploits, leading to faster exploitation and increased risk across both the public and private sectors.

Capabilities that can compromise AI models, systems, and infrastructure:

Numerous areas of risk could emerge in the development of capabilities targeting AI models and systems themselves, including using adversarial inputs to trick AI-powered security filters, hijacking AI agents used in business operations workflows, as well as attacking elements of the AI supply chain (e.g., corrupting training data, compromising a model's cloud infrastructure), not to mention traditional cyber attacks that can be used to target AI models and systems.

Generative AI for social engineering

The accessibility of generative AI tools, such as large language models (LLMs) and deepfake technologies, has led to a surge in sophisticated social engineering attacks, but this increase can be broken down into two distinct parts: the use of AI for social engineering and the use of AI for automating malicious activities. By combining these two components, attackers can increase their success rates exponentially, as they can produce higher volumes of socially engineered lures that are of higher quality with the assistance of LLMs and generative AI. As such, we expect phishing and other social engineering techniques to continue improving with AI's assistance, while spam and phishing detection races to catch up.

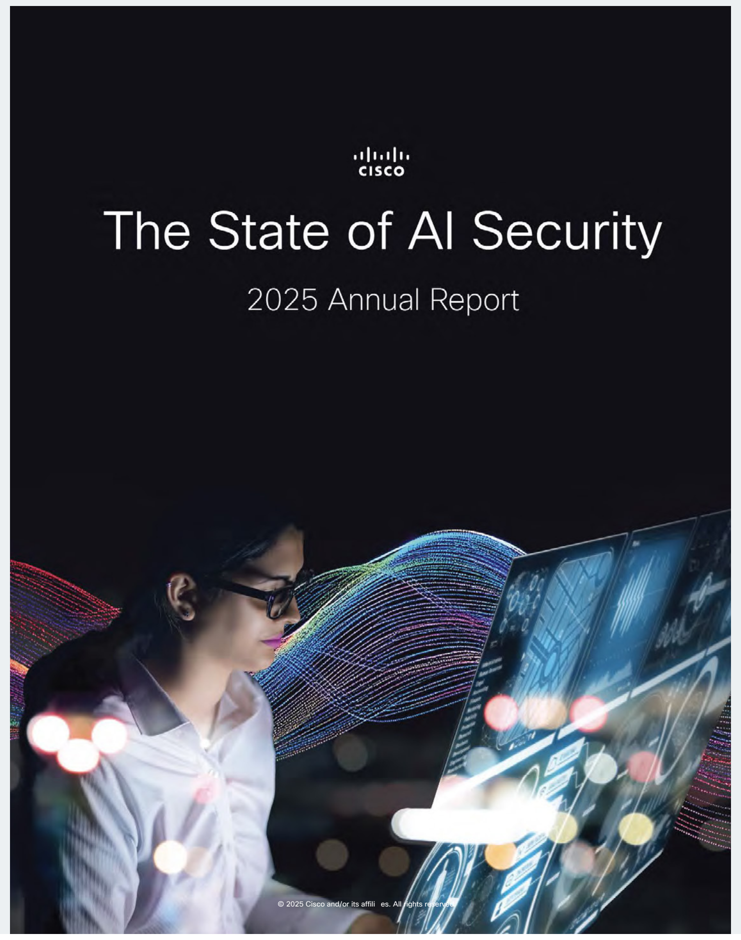
In 2024, cybercriminals leveraged these technologies to create convincing phishing campaigns and manipulate individuals into divulging sensitive information or granting unauthorized access to their organization's networks and systems.

State-sponsored advanced persistent threat (APT) groups and other sophisticated actors may leverage aspects of these features, such as deepfake video and audio for conducting interviews or phone calls or automating social engineering.

Task automation and productivity gains in the attack lifecycle

Threat actors have attempted to leverage chatbots to assist in malware development and task automation to improve their success rates. For example, malicious actors have queried chatbots as a summation tool to gather open-source intelligence on their targets.

Research has proven that LLMs can be used to exploit one-day vulnerabilities (i.e., vulnerabilities that have been disclosed but not patched in a system). Threat actors have leveraged LLMs to assist with basic scripting tasks and code debugging, but we have not yet observed threat actors deploying advanced AI capabilities for vulnerability scanning and exploitation in real-world scenarios. However, cybercriminals have allegedly developed and sold multiple tools that can aid in vulnerability research, reconnaissance, and exploit writing.



Interested in more of what Cisco has to say about AI?

A significant number of new AI policy developments occurred in 2024, largely in response to the increasing prevalence of AI-powered technologies and their market expansion. In Cisco's inaugural State of AI Security report, we provide a comprehensive overview of developments in the AI threat landscape. The report covers important developments in U.S. and international AI policy; in-depth analysis of threats to AI infrastructure, AI supply chains, and AI applications; and original research into many cutting-edge AI security topics like algorithmic jailbreaking, dataset poisoning, and data extraction.

Read here

2024
YEAR IN REVIEW

About Cisco Talos

Cisco Talos is one of the most trusted threat intelligence research teams on the globe. We are comprised of world-class researchers, analysts, incident responders and engineers. Talos powers the Cisco portfolio with comprehensive, proven and tested intelligence covering every customer environment, every event, every single day, all around the world. Talos' core mission is to protect and defend Cisco's customers by understanding the broad threat landscape and distilling the massive amount of telemetry we digest into verifiable detection, intelligence and response for our customers, users and the internet at large.

Our job is your defense.

Stay connected

View our blog: TalosIntelligence.com/blog | **Subscribe:** Threat Source newsletter | **Follow us:** LinkedIn, X, Mastadon and BlueSky