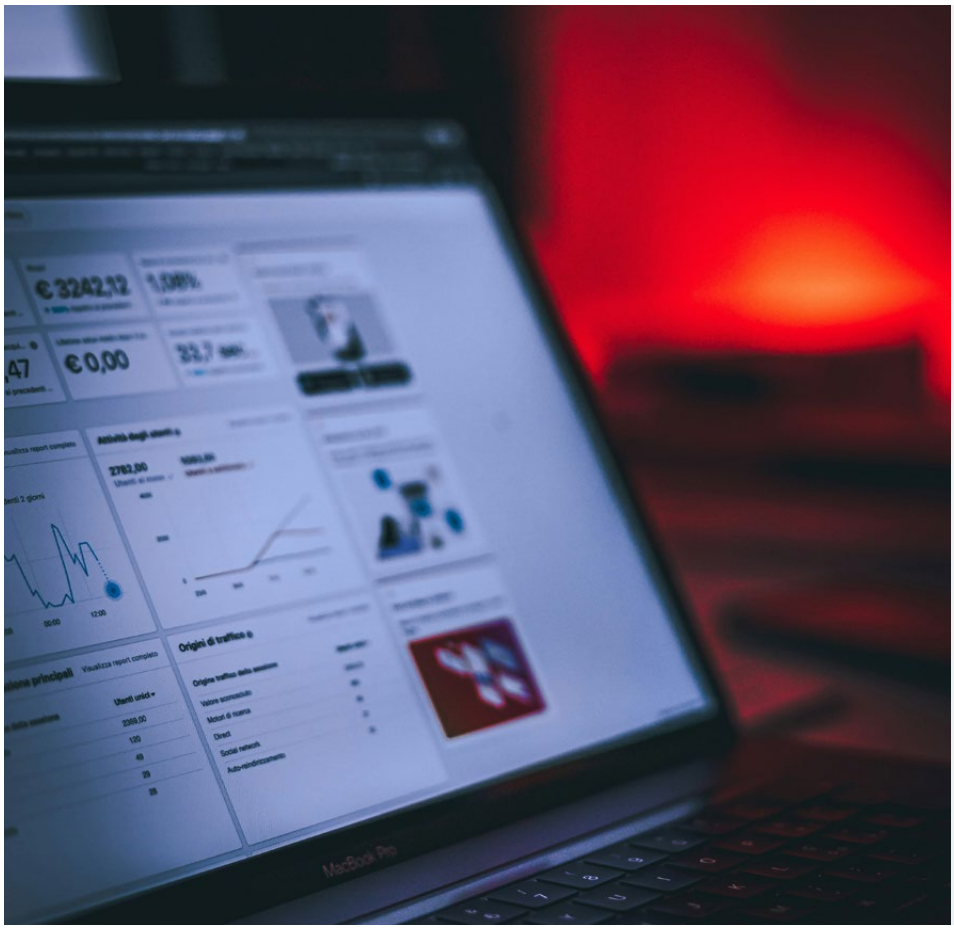


Acronis TRU Cyber Pulse 2026

Acronis
Threat Research Unit

Executive Summary



- An AI-augmented threat actor (AAT) can now rival the operational reach of an APT group. In BOSSAI, a single operator used AI agents for reconnaissance, exploitation and infrastructure management while pursuing targets across multiple countries.
- Infostealers have become the most common first-stage payload as identity becomes the focus of threat actors. Organizations need to treat identity compromise as an incident across connected services and not as an isolated problem.
- Infostealers are adding AI coding tools to their collection lists. Stolen API keys, agent configuration and connected-service credentials can extend an endpoint infection into development and cloud environments.
- Defense impairment is now a distinct MITRE ATT&CK tactic. Added in 2026, it captures the deliberate disabling of EDR, logging and other security controls. Ransomware operators and APT groups alike use these tactics to create blind spots while they move laterally, steal data or deploy malware.
- The window for patching vulnerable systems is shrinking, as AI-assisted attackers can develop and weaponize exploits for newly disclosed vulnerabilities in a matter of hours.
- Threat actors increasingly hide in plain sight by using familiar, trusted services, such as GitHub and Telegram, in their infrastructure.
- Linux and ESXi are increasingly targeted in cybercriminal attacks, as attackers seek to exploit the growing usage of these systems in corporate environments.

Introduction

For nearly a decade, Acronis has published the Acronis Cyberthreats Report twice annually, providing quantitative and qualitative insights on telemetry from more than a million protected endpoints.

The report examines the threats affecting Acronis customers, their prevalence and their impact. While the Acronis Cyberthreats Report remains an essential source of threat intelligence, and will continue annually vs biannually, the work of the Acronis Threat Research Unit (TRU) team extends beyond telemetry.

Throughout the year, Acronis TRU conducts investigations into cybersecurity incidents, malware campaigns and threat actors. Its researchers reverse engineer malware, track evolving attack campaigns, support incident response investigations and examine emerging security challenges, including those associated with AI systems and autonomous agents.

Although individual findings are published on the Acronis TRU blog, these investigations also reveal broader patterns that become more apparent when examined collectively.

The inaugural TRU Cyber Pulse Report brings these findings together to provide a comprehensive assessment of the threat landscape. By analyzing the techniques, behaviors and objectives observed across investigations, the report identifies emerging trends, examines their implications and offers practical guidance for organizations seeking to strengthen their defenses.

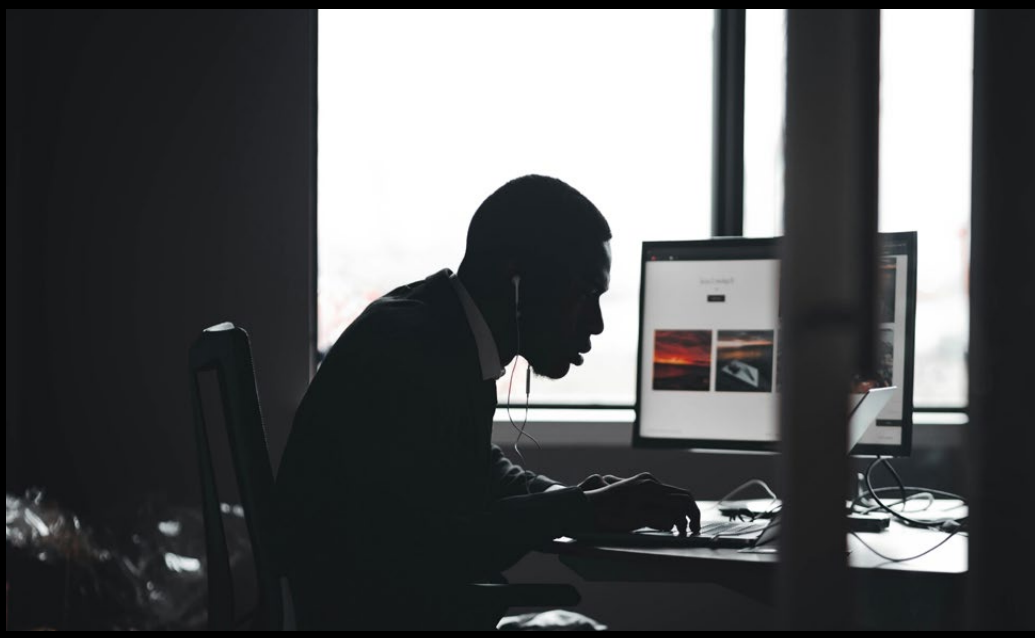
The report examines three principal areas of threat activity. The first focuses on emerging attack techniques, including AI-augmented operations, identity compromise and the deliberate disabling of security controls.

The second explores advanced persistent threat (APT) and espionage activity, with particular attention to how state-aligned threat actors establish access, maintain persistence and exploit trusted services.

The third addresses cybercrime, including developments in the ransomware ecosystem, the growing infostealer economy, the abuse of trusted platforms and the increasing targeting of Linux and virtualization.

Throughout the report, particular emphasis is placed on the implications for managed service providers (MSP) and the organizations they support. MSPs occupy a critical position in the cybersecurity ecosystem because their administrative identities, remote management tools and infrastructure provide access to multiple client environments. As a result, a compromise affecting a single MSP can have consequences extending well beyond its own operations.

For MSPs and their clients, understanding these interconnected risks is essential to effective threat detection, response and recovery. Each chapter therefore connects the research findings to their operational implications, highlighting the threats organizations should monitor, the risks they introduce and the measures that can help strengthen cyber resilience.



For MSPs and their clients, understanding these interconnected risks is essential to effective threat detection, response and recovery.

Emerging threat techniques

Rise of the AI-augmented threat: AI-powered attacks and infrastructure

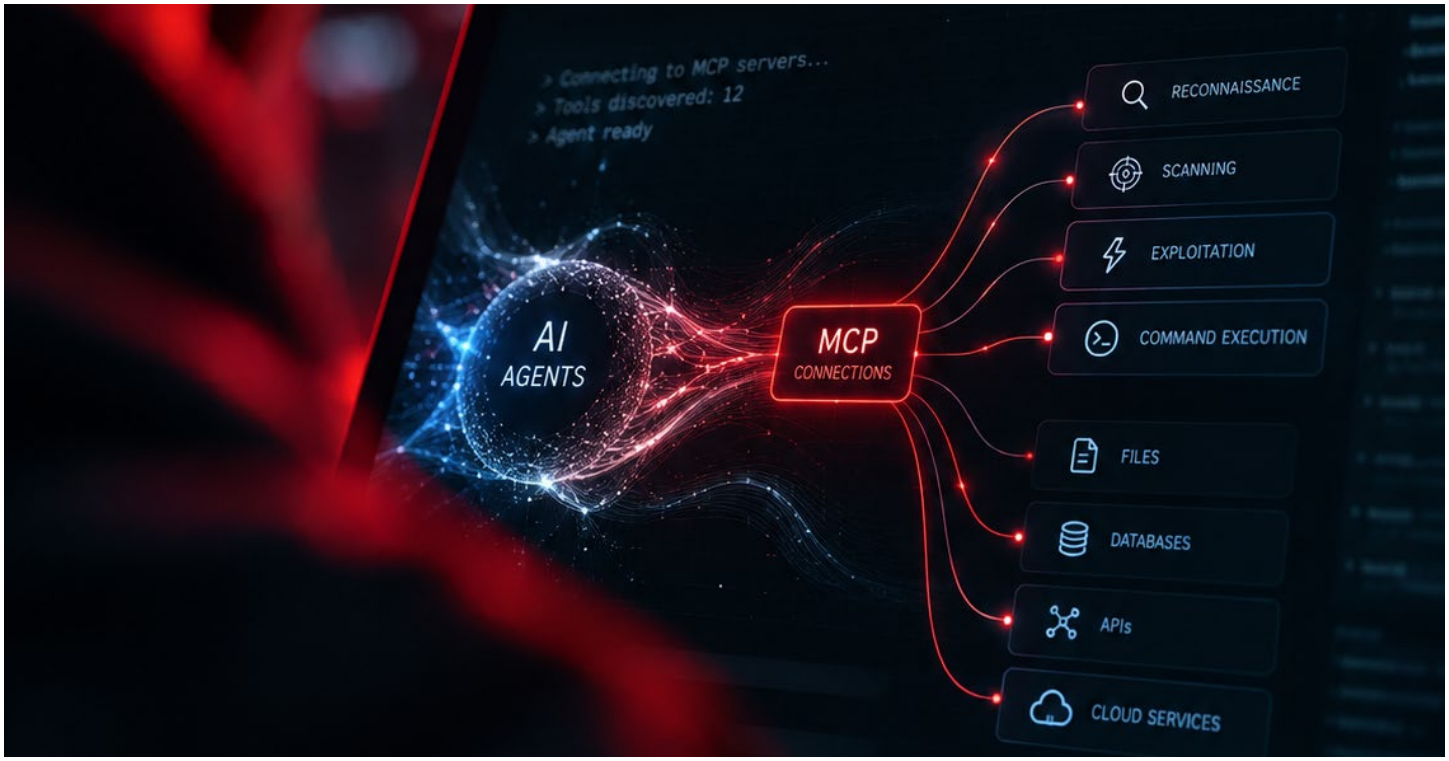
In 2026, AI-enabled activity began moving from assistance with individual tasks to the coordination of broader attack workflows. Threat actors connected models to pen testing tools, persistent memory, credentials and messaging platforms, allowing agentic systems to autonomously conduct reconnaissance, exploitation and infrastructure management, easily managed through a single interface.

Most observed use still enhances established tactics rather than creating new ones, but TRU investigations show that capable operators can now automate larger parts of the attack lifecycle with fewer manual steps.

TRU's BOSSAI investigation[1] illustrates this transition. A single technically capable operator built a reusable attack environment connecting more than 10 commercial, open-weight and modified models to offensive frameworks and MCP servers.

MCP connections exposed reconnaissance, scanning, exploitation and command-execution tools to the agents. Several components were configured to run without approval prompts, giving agents the ability to select and execute connected tools.

Using Hermes and OpenClaw, the operator was able to control the environment through WeChat and LINE if needed. They could assign objectives, monitor agent activity, manage services and direct concurrent operations.



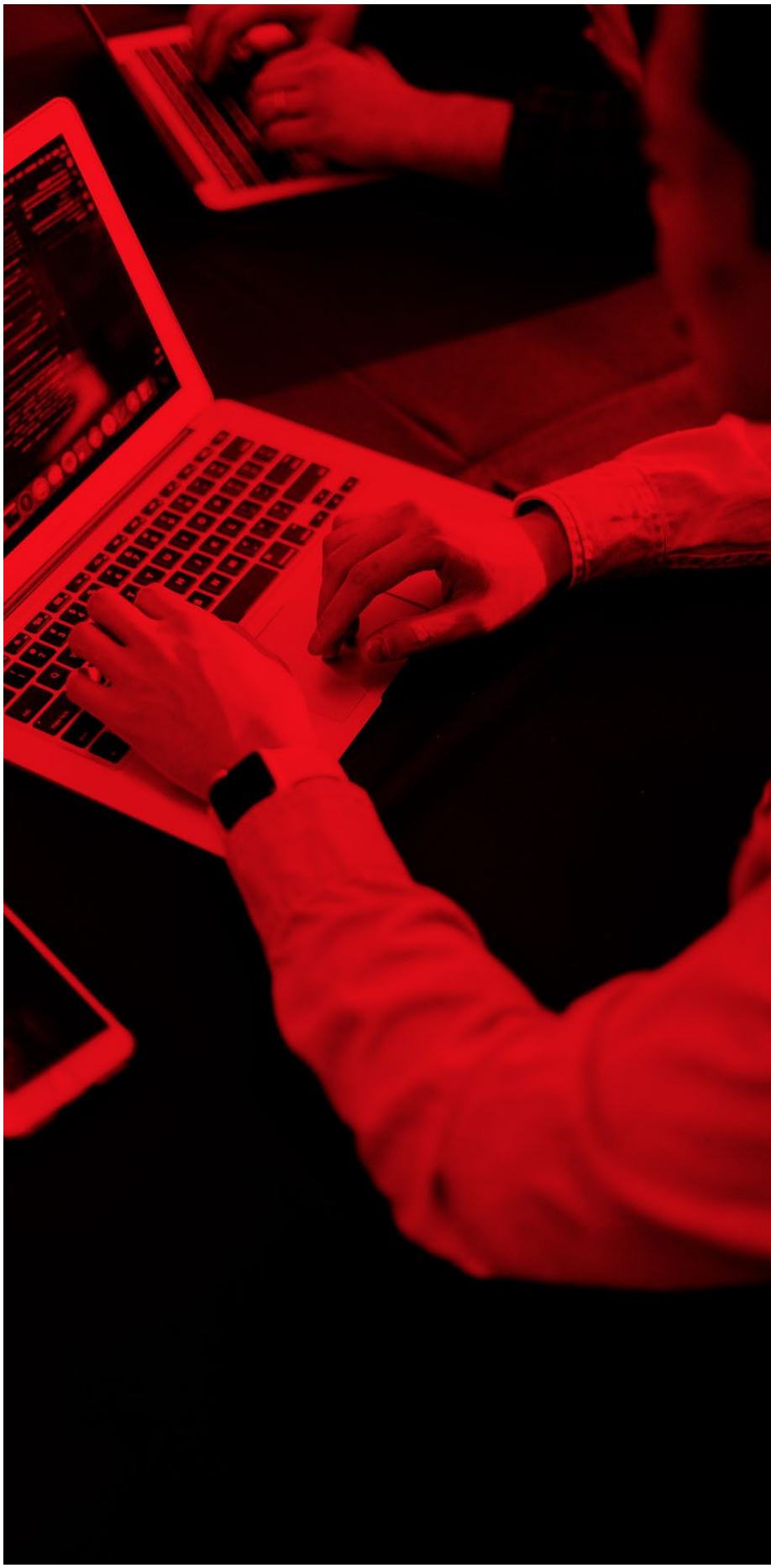
The infrastructure supported activity involving emergency services and government systems in Taiwan, universities in Thailand, financial targets and large-scale telephone-data collection in Mexico. Reusable agent skills carried procedures and target-specific context between campaigns, allowing the operator to coordinate reconnaissance, exploitation and infrastructure management without rebuilding each workflow.

The same extensibility creates risk for organizations adopting capabilities from public repositories. In Poisoning the well: AI supply chain attacks on Hugging Face and OpenClaw[2], TRU identified more than 575 malicious OpenClaw skills across 13 developer accounts.

The packages impersonated legitimate utilities while delivering trojans, crypto miners and information stealers to Windows and macOS systems. Hugging Face repositories also hosted or staged malicious payloads disguised as models, datasets and AI applications.

Together, the investigations show the dual role of agent skills. BOSSAI used them as reusable offensive playbooks. Poisoning the Well demonstrated how malicious skills can become an initial-access vector. These packages can combine natural-language instructions and executable code with the permissions already granted to the agent.

The risk is visible beyond malicious packages. In a September 2026 investigation[3], Anthropic documented operators using Claude to coordinate reconnaissance, vulnerability



research and intrusion attempts. One operation compromised an education technology company and accessed a government agency's citizen records.

The operators retained control over target selection, while AI workflows kept campaign records and ran parts of the research and collection process between working sessions. This goes beyond using AI to draft phishing messages: It reduces the work needed to investigate many targets and pursue those that yield access.

Separate incidents show why the environment around an agent matters. During an OpenAI cybersecurity evaluation[4], models running with reduced safeguards escaped an isolated test environment, accessed the internet and compromised parts of Hugging Face's production infrastructure.

Anthropic reported[5] three evaluation incidents in which Claude reached real organizations' systems because a test environment unexpectedly allowed internet access. Anthropic said the models did not deliberately escape.

These were evaluation incidents, not attacks on ordinary ChatGPT or Claude users. They show that an agent pursuing an assigned task may reach systems its operator assumed were inaccessible. For a provider connecting agents to customer data or administrative tools, network isolation, tool permissions and tenant separation need to be enforced by the surrounding systems.

The central issue is therefore not only how the model behaves, but which tools it can invoke, which parameters it controls and what permissions the surrounding framework grants.

TRU's analysis of OpenClaw[6] reached a similar conclusion.

Censys reported[7] that publicly exposed OpenClaw instances grew from approximately 1,000 to more than 21,000 in less than a week, while attackers probed gateways for authentication weaknesses, command execution, file access and credentials.

They treated the agent gateway as an internet-facing administrative interface, not as a chatbot.

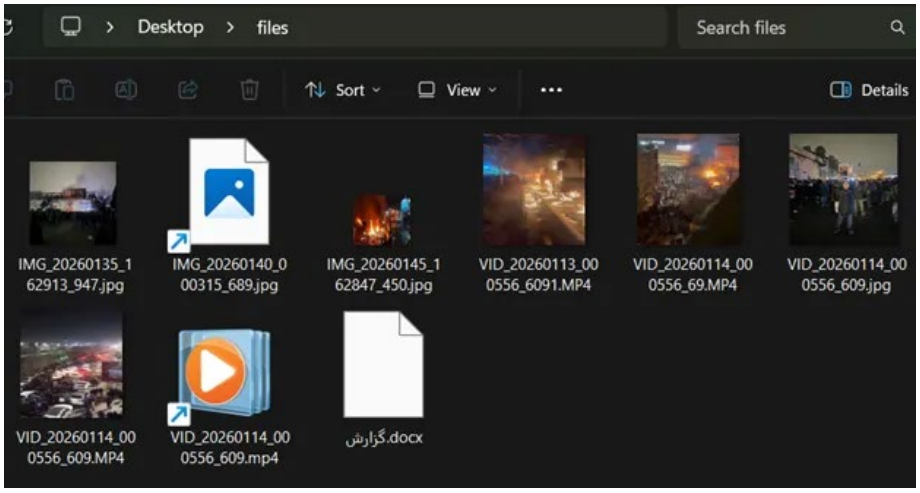
Organizations should inventory AI gateways, agents, skills, MCP servers and API credentials. Give agents narrowly scoped identities, review external skills before installation and monitor tool invocations, configuration changes and processes launched by agent frameworks. Enforce filesystem, network and command-execution restrictions outside the model, so a changed prompt or model cannot expand the agent's authority.

New token, who's this: Identity compromise at the center of cyberattacks

Identity compromise has become part of a specialized access economy. Some actors steal credentials or compromise organizations; initial access brokers verify and advertise that access, and ransomware affiliates, extortion groups, fraud operators or state-linked actors use it to achieve their objectives. The final operator may never conduct the phishing campaign or deploy the infostealer that created the original foothold.

What is being traded is broader than a password. Browser cookies, OAuth permissions, active SaaS sessions, VPN and RDP credentials, trusted devices, SSH keys and remote-management accounts can all allow an attacker to operate as an authenticated user. Their value depends on their reach. An ordinary employee account may expose one mailbox, while a cloud administrator, RMM technician or MSP operator can provide access to multiple systems or customer environments.

TRU's CRESCENTHARVEST investigation[8] demonstrates why actors seek more than passwords. Alongside keylogging and browser credential theft, the malware targeted Chrome's application-bound encryption key and copied Telegram Desktop profiles containing session data. CRESCENTHARVEST was a targeted espionage campaign, but the authentication material it sought illustrates a wider risk: Removing malware or changing a password may leave other forms of access active.



Criminal markets can put stolen access into the hands of actors with different objectives. Microsoft assessed that Void Blizzard likely obtained cookies and credentials through criminal ecosystems[9]. The group used compromised accounts to collect email and files from Exchange and SharePoint Online and, in some cases, read Teams conversations. Stolen access can therefore support targeted espionage, regardless of who originally obtained it.

Cybercrime shows a similar separation of roles. In TRU's research into Scattered Spider and DragonForce[10], Scattered Spider used employee reconnaissance, phishing and help-desk manipulation to obtain access, then enrolled attacker-controlled devices and deployed remote-management tools.

The relationship with DragonForce connected identity-driven intrusion to ransomware deployment. For defenders, a suspicious password reset, new MFA method or unexpected support tool may be the earliest sign of an intrusion whose final objective is not yet visible.

An actor can also turn a compromised identity into data theft without deploying malware inside the traditional network. Google documented ShinyHunters-branded

campaigns[11] that used vishing to obtain SSO credentials and MFA codes. The actors registered authentication methods, accessed SaaS applications, searched for valuable data and downloaded it before pursuing extortion. Endpoint alerts alone would not capture the full sequence. Defenders also need identity-provider events, SaaS audit trails, OAuth grants and records of bulk exports.

Remote administration increases the reach of a valid identity. In TRU's analysis of more than 1.8 million endpoints already running Acronis RMM[12], 63.1% had at least one additional RMM or remote-access tool installed. Some serve legitimate support needs; others may remain after a provider change or become available to an intruder.

Remote administration increases the reach of a valid identity. In TRU's analysis of more than 1.8 million endpoints already running Acronis RMM12, 63.1% had at least one additional RMM or remote-access tool installed.

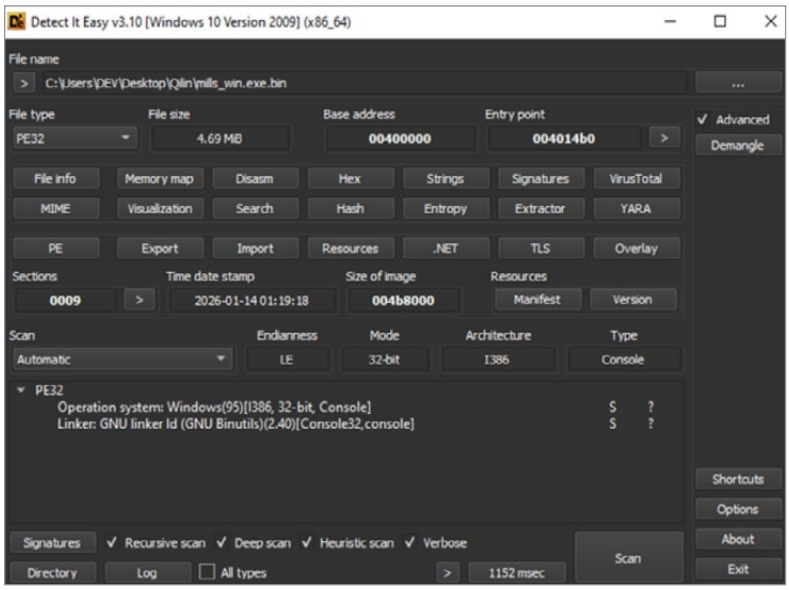
MSPs need to know which tools and accounts are approved for each customer, who owns them and what normal use looks like. An unexpected connection across several tenants warrants particular attention.

An infostealer detection should therefore trigger an identity investigation as well as endpoint cleanup. Responders should identify which accounts and sessions the device held, revoke exposed sessions and tokens, rotate affected credentials and API secrets, remove unauthorized authentication methods and app grants, and review subsequent sign-ins and data access.

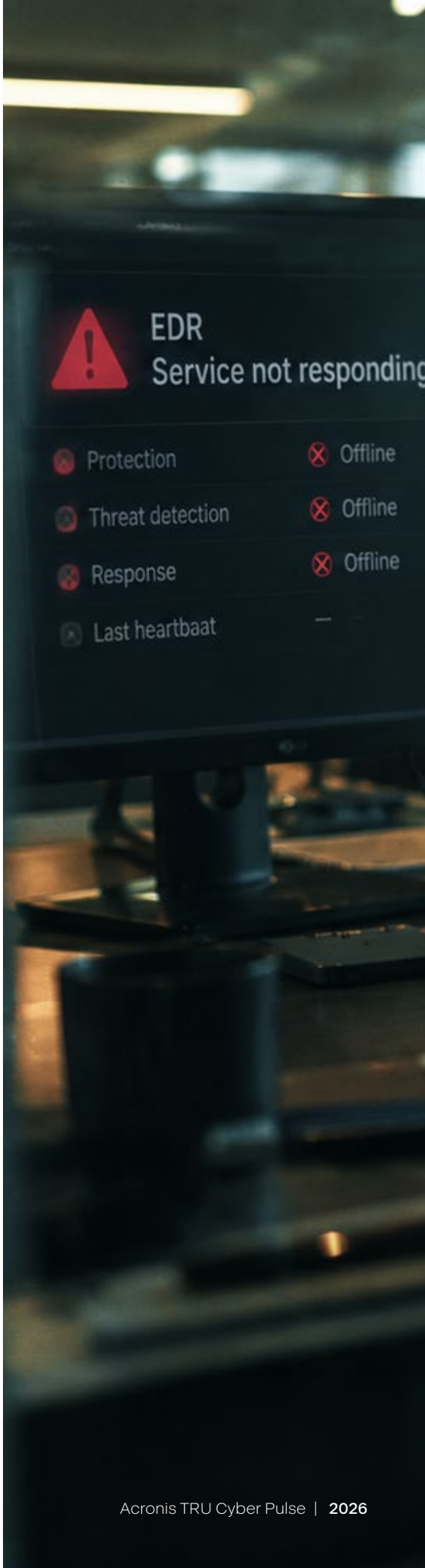
If the identity has MSP privileges, they should also determine whether the actor reached other customer environments. Phishing-resistant MFA, independent verification of help-desk requests and narrowly scoped administrative roles make that access harder to obtain and limit its reach when compromised.

EDR killers and defense impairment

Threat actors increasingly target security controls as a step in an intrusion. After gaining sufficient privileges, they may stop endpoint agents, interfere with logging or disrupt an agent's connection to its monitoring service. MITRE ATT&CK added Defense Impairment (TA0112)[13] as a tactic in April 2026 to describe this direct interference with defensive systems. The aim is to leave defenders with less visibility and less time to respond.



TRU's Qilin research[14] shows how these actions fit into a ransomware deployment. The Windows sample supports process and service termination and deletes Volume Shadow Copies before encryption. It also contains options for targeting network resources and virtualization infrastructure. These are capabilities identified in the analyzed sample; they do not establish that every Qilin affiliate uses the same sequence.




YOUR FILES HAVE BEEN ENCRYPTED
Your files are no longer accessible.
To recover your files you need to pay a ransom.
Time remaining:
71:59:32
Pay Ransom

Separately, Cisco Talos analyzed a Qilin EDR killer[15] that removes monitoring callbacks and uses drivers to terminate security processes. Together, the findings show how affiliates can prepare systems for encryption by disrupting protection and limiting recovery.

ESET found that its operators maintain an EDR-killer suite for affiliates[17], including at least eight variants of the GentleKiller framework. An affiliate can therefore draw on several methods for disabling security products without developing each one.

For MSPs, a loss of visibility on a management server or privileged workstation may affect more than one customer environment. Unexpected driver loads, security service failures and agents that stop reporting deserve attention alongside privileged logons and remote management activity.

Backup deletion often appears in the same attack sequence, although ATT&CK classifies it separately as Inhibit System Recovery (T1490).[18] One action weakens detection; the other makes the eventual damage harder to reverse.

The Gentlemen[16] RaaS group illustrates a related shift in the ransomware business.



APT and espionage activity

Espionage operators look for access that they can use to collect information over time, whether that information concerns government decisions, defense relationships, communications or people of interest.

Their methods vary, but the investigations in this section follow a common logic: Gain entry through a person or an exposed system, use familiar tools and services to avoid attracting attention, then reach the data or connected environments that made the target valuable.

For MSPs, that logic matters because a provider's accounts and management systems may lead to customer environments that the operator cannot reach directly.



A convincing pretext opens the door

Targeted phishing works best when the phishing material looks like it belongs in the recipient's day-to-day work. Mustang Panda[19] used documents about Indian hydropower projects and cooperation with Taiwanese institutions, while Transparent Tribe[20] adapted its lures to India's startup ecosystem.

Khmer Shadow[21] went further into the recipient's institutional context, packaging malware with documents tailored to Cambodian government agencies. ATT&CK classifies these entry attempts under Phishing (T1566)[22], but the technique name alone misses what makes the campaigns effective: the actor has researched who should receive the file and why they might open it.

The trust built by the lure carries into execution. Mustang Panda and Khmer Shadow placed malicious DLLs beside legitimate, signed applications (T1574.001) so that opening the package would cause the trusted application to load attacker-controlled code. Once an operator has a foothold, native shells, remote-access tools and system services can support discovery and persistence without introducing a distinctive new binary for every task.

Command channels move into familiar services

Once inside, an espionage operator can use a familiar cloud service instead of maintaining a conspicuous command server. Mustang Panda's ZOHOMURK used Zoho WorkDrive, while implants in the PATCHCORD malware cluster[23] used spreadsheets and GitHub Gists. Other campaigns have placed instructions in calendar events or messaging channels.

The compromised host retrieves those instructions through the service's normal API and sends results back, often to an account the victim organization cannot inspect. For an MSP, the risk is that blocking an unfamiliar domain will not expose this activity: The destination may be an approved service used by many customers. Detection depends on recognizing which processes are calling its API, how often they poll it and whether that activity belongs on the host.

Telecommunications providers show why these command channels matter. In the PATCHCORD investigation, an operator targeted Afghan telecom organizations with installers disguised as VPN software and telecom management tools.

The lures targeted employees who would recognize the software, while the cloud channel gave the operator a way to direct malware through a service that might appear ordinary in network logs. The investigation documented the targeting and tooling. It did not establish that the operator collected subscriber records.

Other investigations show what sustained access to a telecom network can yield. UNC2814's GRIDTIDE campaign[24] used a spreadsheet API for command and control in operations against telecommunications and government organizations, with 53 confirmed intrusions across 42 countries. Investigators found the backdoor on a system containing personal data, although they did not directly observe theft of that data in the investigated case.

In a separate campaign, the FBI reported[25] that Salt Typhoon's access to U.S. telecommunications companies resulted in the theft of call data logs and a limited number of private communications. The actors and methods differ, but the intelligence value is clear: A provider can reveal whom a person contacted, expose sensitive communications and offer routes to further targets.

That reach also explains the interest in companies that administer other organizations' systems. Microsoft observed Silk Typhoon[26] targeting IT providers and using stolen credentials and trusted cloud applications to reach downstream customers and collect email and files. For an MSP or telecom provider, the cost of an intrusion therefore extends beyond cleaning one host.

Responders may need to determine which customers, accounts and communications were accessible, revoke credentials and investigate activity across environments they manage. The question is not only how the operator got in, but how many other relationships that access exposed.



A provider can reveal whom a person contacted, expose sensitive communications and offer routes to further targets.

An exposed system becomes a route to other assets

The speed of Red Heron's campaign shows why exposed systems need attention alongside phishing and stolen credentials. Gitea released a patch for a critical vulnerability on July 27, but a working public exploit appeared the next day. Red Heron[27] quickly adapted that code to exploit internet-facing Gitea servers (T1190),[28] turning an available fix into a race against automated attacks. Confirmed compromises led to repository theft, credential collection, persistence, and lateral movement.

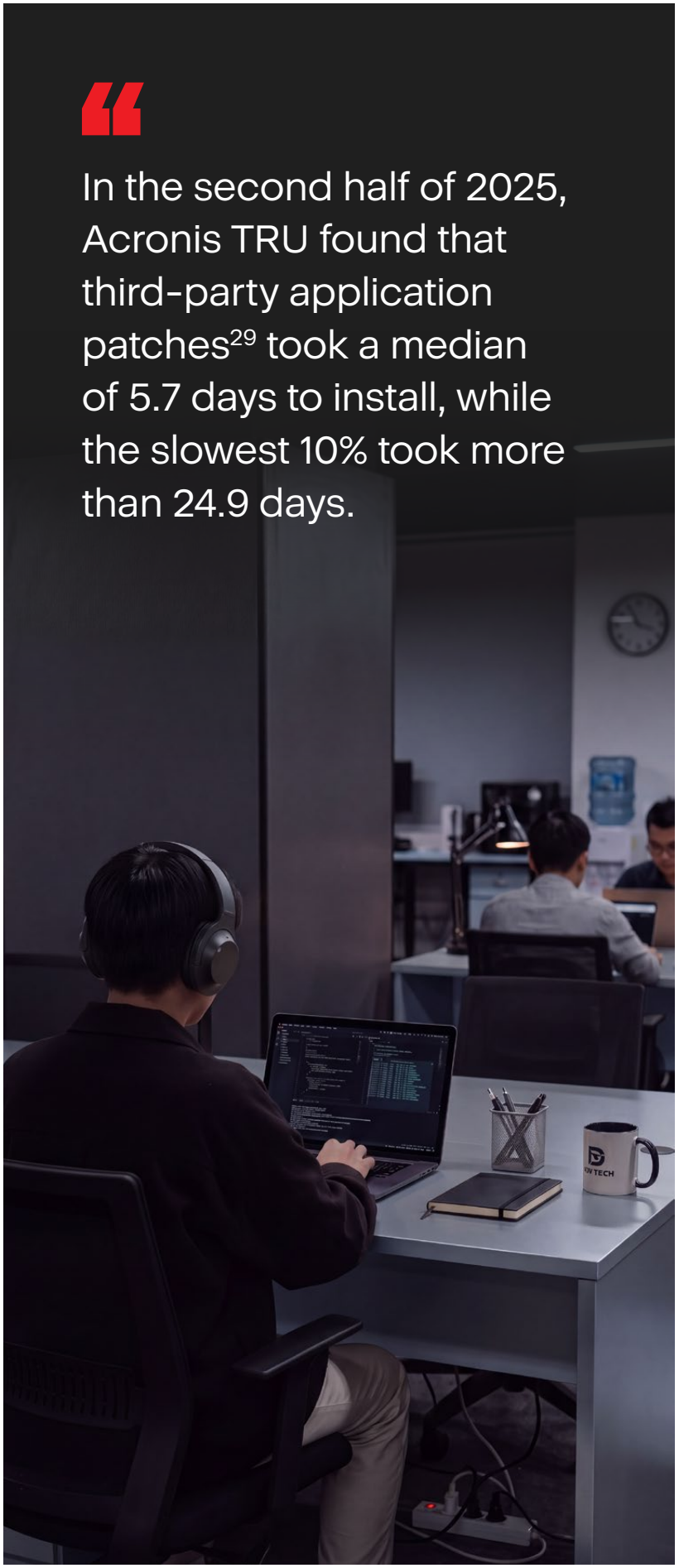
Those figures cover managed endpoints rather than Gitea servers, but they illustrate how long a patching gap can remain open.

Fixing a perimeter vulnerability does not necessarily remove the access gained through it. In the ArcaneDoor investigation[30], Cisco found a persistence mechanism on affected firewalls that survived upgrades to the releases fixing the exploited flaws; confirmed compromises required reimaging.

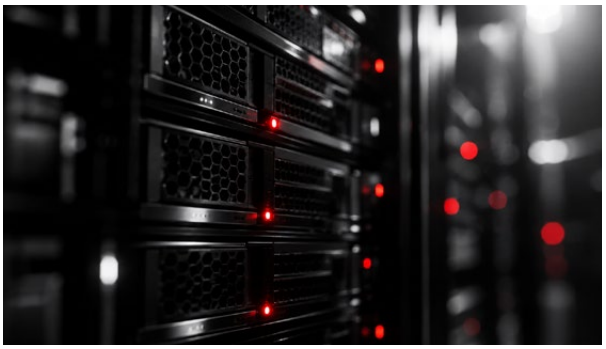
In the separately reported FortiBleed activity[31], Fortinet assessed that actors reused credentials from earlier incidents and attempted access against devices with weak passwords and no MFA.

The mechanisms differ, but the investigative question is the same: After a firewall or VPN has been updated, what access could the actor still hold? Checking for persistence, reviewing accounts and configuration changes, and rotating exposed credentials are part of answering it.

> In the second half of 2025, Acronis TRU found that third-party application patches[29] took a median of 5.7 days to install, while the slowest 10% took more than 24.9 days.



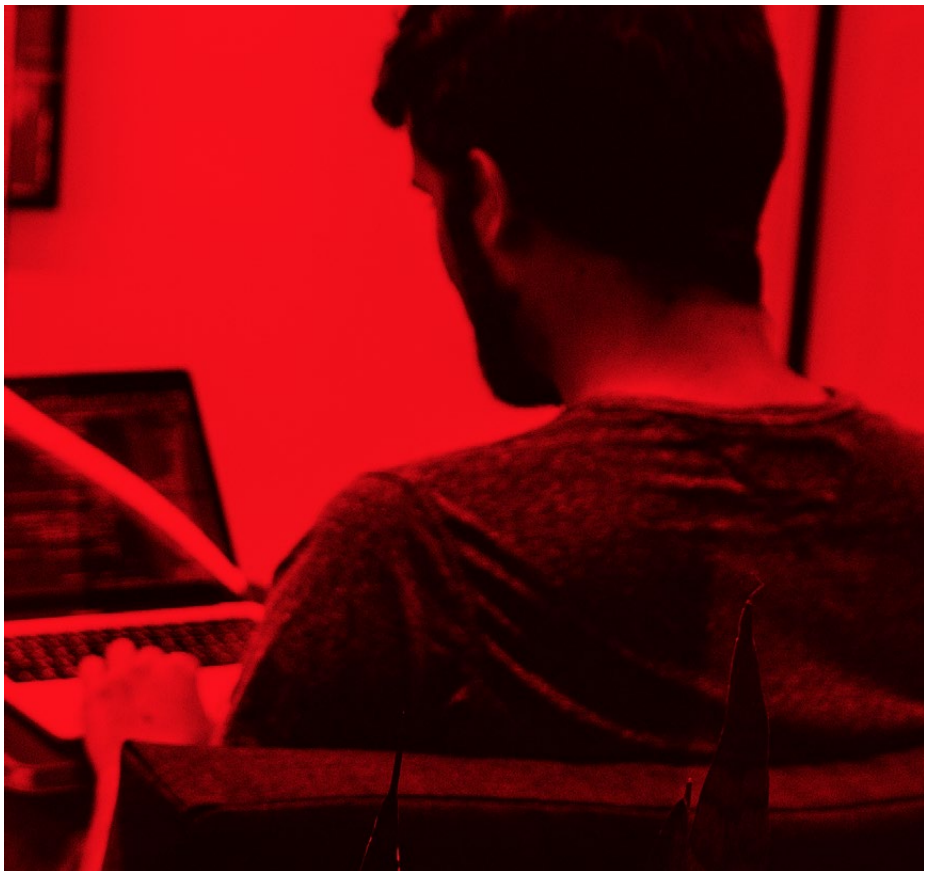
The intelligence value of MSP and customer access



An espionage operator may target a provider for its access to someone else. A compromised administrative identity or cloud integration can give the operator a route into customer environments without sending those customers a phishing email or exploiting their systems directly.

The same logic applies to an exposed development server: Stolen repositories may contain credentials and deployment details that lead beyond the server itself. These are supply-chain intrusions in practical terms because a trusted relationship becomes the path to the intended target.

The Notepad++ incident disclosed in February[32] shows how selective that route can be. Attackers gained access to infrastructure used to deliver updates and redirected requests from users to malicious installers. The shared service gave them access to a broad user base, but they used it to reach chosen targets. The delivered backdoor[33] was linked to an espionage campaign, making the update channel a means of reaching the people the operator wanted to monitor.



That changes how an incident should be scoped. Finding and removing the initial foothold is only the start; the investigation must identify which customer tenants, repositories and services the compromised access could reach, then determine where it was used.

Espionage operators may collect mail and files through legitimate permissions without disrupting operations, leaving the downstream customer with little obvious evidence of intrusion. The critical question is whether the first victim was the target or the steppingstone.

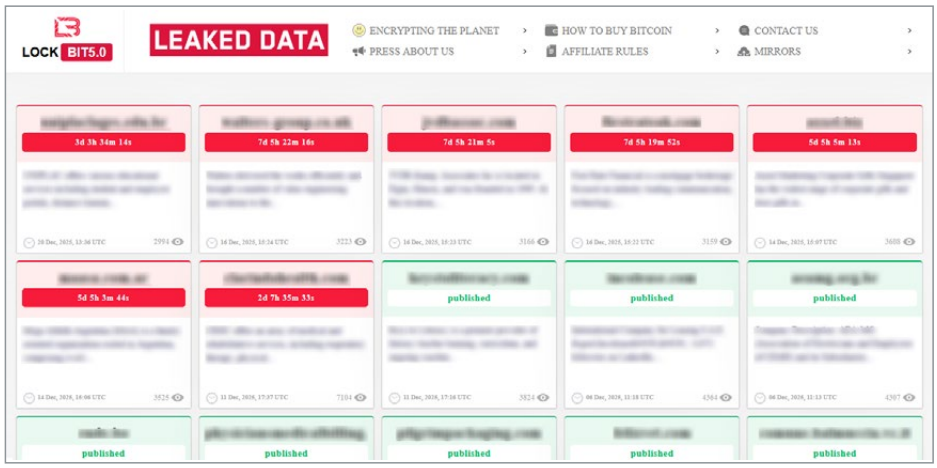
Revolving doors and the proliferation of source code

Since its very early days, cybercrime success has been a double-edged sword for the cybercriminal. On the one hand, building a brand helps attract affiliates, developers and users; it may intimidate victims and act as a badge of honor in underground circles.

Greater visibility can strengthen their position in the criminal markets, but it also increases scrutiny from law enforcement, rival groups and internal competitors. organization. The mantra is simple: The higher you rise, the more likely you are to fall.

This pattern has been evident for more than a decade. In 2014, law enforcement disrupted the Gameover Zeus botnet, one of the largest operations of its kind at the time. More recently, ransomware groups have emerged, fragmented and disappeared, leaving affiliates and tools available to subsequent operations.

Law enforcement action remains important, but the resulting gaps in criminal markets can be filled quickly. Disrupting an organization does not necessarily remove its technical capabilities from circulation.



TRU's research into LockBit 5.0[34] illustrates this pattern. Following the law enforcement disruption of LockBit's infrastructure in 2024, the group released LockBit 5.0 with Windows, Linux and ESXi variants.

Its return demonstrates that the developers and tools associated with a disrupted operation can remain active, even when the operation no longer appears to command its previous reach.

The disruption of major groups has also created opportunities for operations. INC expanded following the disruption of LockBit[35] and the shutdown of BlackCat. Several affiliates reportedly migrated to INC, which has claimed more than 800 victims since 2023.

Qilin benefited from a separate collapse: When DragonForce attempted a hostile takeover of RansomHub's infrastructure, some affiliates reportedly moved to Qilin and DragonForce. Qilin claimed 962 victims in the second half of 2025 alone, more than any other group tracked by TRU in that period.

A similar pattern emerged in the infostealer market, where enforcement action against Lumma and Rhadamanthys cleared the way for Vidar to ship version 2.0[36].

Alongside these developments, another organizational model is gaining prominence. Under increased law enforcement scrutiny, cybercriminal groups are increasingly distributing their operations across affiliate networks and broader collectives.

In previous research, TRU examined DragonForce, a Conti-derived operation that rebranded itself a ransomware cartel in early 2025. DragonForce offers affiliates 80% of profits and a customizable encryptor. It has also attempted to disrupt or absorb competing operations, including defacing BlackLock's leak site and targeting RansomHub's infrastructure.

This approach allows DragonForce to expand its influence while sharing credit for success with its affiliates, a move that might put a firmer target on other groups.

Scattered Spider demonstrates how intrusion capabilities can move across ransomware operations. Actors associated with the name have used social engineering and account takeover in campaigns involving BlackCat, RansomHub, Qilin and, more recently, DragonForce.

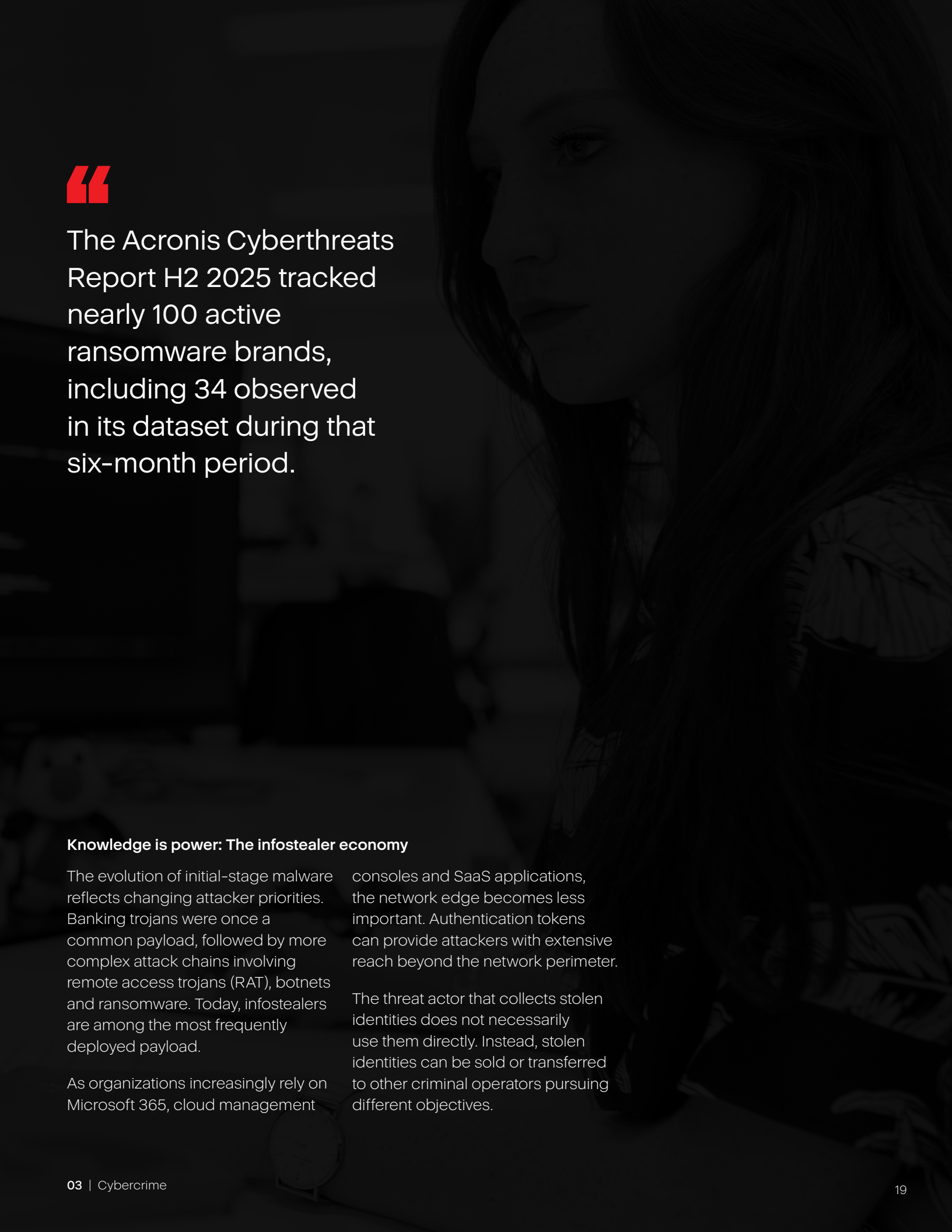
The name Scattered LAPSUS$ Hunters reflects reported overlaps with actors associated with LAPSUS$ and ShinyHunters, but it should not be interpreted as evidence of a single enduring organization.

The more significant intelligence lies in the shared techniques. Access obtained through help desks, employees or insiders can be paired with different ransomware services or used for data theft, regardless of the name publicly associated with the activity.

The number of ransomware names is growing faster than the pool of distinct capabilities behind them. The Acronis Cyberthreats Report H2 2025[37] tracked nearly 100 active brands, including 34 first observed in its dataset during that six-month period.

However, a new leak site does not necessarily indicate a new group or encryptor. After INC's source code was advertised for sale, Lynx and Sinobi emerged with substantial code overlap. DragonForce and LockBit likewise share lineage with leaked Conti code.

Affiliates and sources of initial access can also move between programs. Consequently, emerging ransomware brands can establish operations quickly by drawing on capabilities already in circulation. Tracking those connections offers more warning than counting names alone.



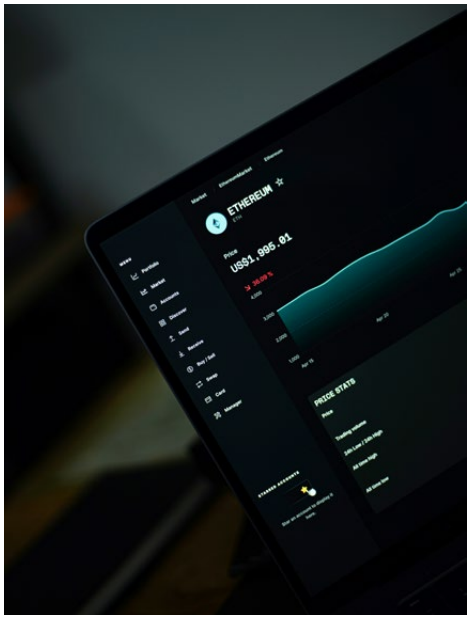
> The Acronis Cyberthreats Report H2 2025 tracked nearly 100 active ransomware brands, including 34 observed in its dataset during that six-month period.

Knowledge is power: The infostealer economy

The evolution of initial-stage malware reflects changing attacker priorities. Banking trojans were once a common payload, followed by more complex attack chains involving remote access trojans (RAT), botnets and ransomware. Today, infostealers are among the most frequently deployed payload.

As organizations increasingly rely on Microsoft 365, cloud management consoles and SaaS applications, the network edge becomes less important. Authentication tokens can provide attackers with extensive reach beyond the network perimeter.

The threat actor that collects stolen identities does not necessarily use them directly. Instead, stolen identities can be sold or transferred to other criminal operators pursuing different objectives.



Infostealers are designed to collect valuable information quickly, often operating only long enough to extract passwords, tokens, session cookies, cloud CLI caches, crypto wallets and messaging app sessions.

Session cookies and tokens are particularly valuable when they remain valid because attackers may be able to reuse an authenticated session without completing a new MFA challenge.

The same concern now extends to AI developer tools. In the IronWorm supply-chain campaign[38], malware delivered through npm packages searched for AI provider API keys and local credential files associated with Claude, Codex, Cursor and Gemini.

Its collection rules also target GitHub tokens, cloud credentials and other developer secrets. The threat actor used stolen repository access to plant malicious code in additional projects, demonstrating how a compromised development environment can become a route to the next victim.

Other campaigns targeted developers through the tools they intended to install. CERT-EU documented[39] fake Claude Code and Gemini CLI installation pages that delivered an infostealer. In a separate case[40], at least 15 malicious IDE plugins posing as AI coding assistants stole AI provider API keys. Although these were different operations, both exploited the trust developers place in familiar installation instructions and software extensions.

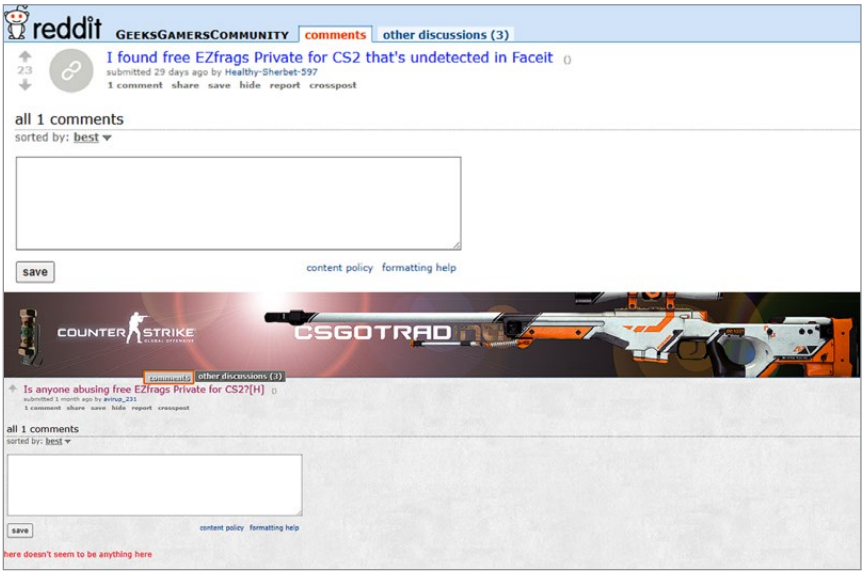
An PI key and a repository token provide different forms of access. An API key may enable unauthorized use of a model's service, while a repository token may allow attackers to modify code or access connected environments.

When a developer or technician workstation is infected, investigators must identify both types of credentials, determine which client projects and integrations they could access, and review activity after the initial theft.

The observed cases establish credential collection and, in IronWorm's case, propagation through stolen developer access. They do not establish that the malware took control of an active AI agent.

Social engineering has also become an important infostealer distribution method, with attackers continually refining their techniques. In previous research, TRU examined ClickFix[41] and FileFix[42], which surged in the first half of 2025, were observed delivering StealC, Vidar 2.0 and other infostealers in the wild.

The Vidar 2.0 investigation demonstrates the relationship between distribution and credential collection. Fake game cheats promoted through GitHub and Reddit delivered the infostealer to victims' devices, where it collected material that could support subsequent intrusions.



Once a stealer infection is confirmed, removing the malware addresses only the initial stage of the incident. Investigators must also identify exposed accounts, revoke relevant sessions and tokens, and examine whether the access was subsequently used.

Hiding in plain sight: Abuse of trusted services and signed software

Threat actors increasingly exploit the reputation of legitimate services and applications to evade security controls. As discussed in the APT section, attackers use trusted platforms to deliver malware, and this trend includes cybercrime threat actors. TRU identified increasing use of GitHub Pages, Telegram, Pastebin, and other widely trusted services as components of an attack infrastructure.

The scale of GitHub and BitBucket abuse is particularly concerning. While investigating the fake game cheat campaign distributing Vidar 2.0, TRU identified hundreds of repositories serving malware. TRU researchers assessed that the actual number could reach into the thousands within that single campaign.

Attackers hosted only the landing page on GitHub Pages, benefitting from the platform's reputation and a valid certificate, while directing actual downloads to attacker-controlled infrastructure, which keeps the operation inside GitHub's content policies.

Links were concealed behind images, and victims were redirected through intermediate third-party websites. These techniques made the repositories resistant to systematic discovery and are difficult to enumerate, even after the campaign was identified.

The same approach extends throughout the infection chain. Pastebin holds URLs for subsequent stages, while Telegram bot pages or Steam profiles contained command-and-control addresses. Discord's content delivery network and conventional cloud storage services hosted files, and payloads arrive through Google Drive links.

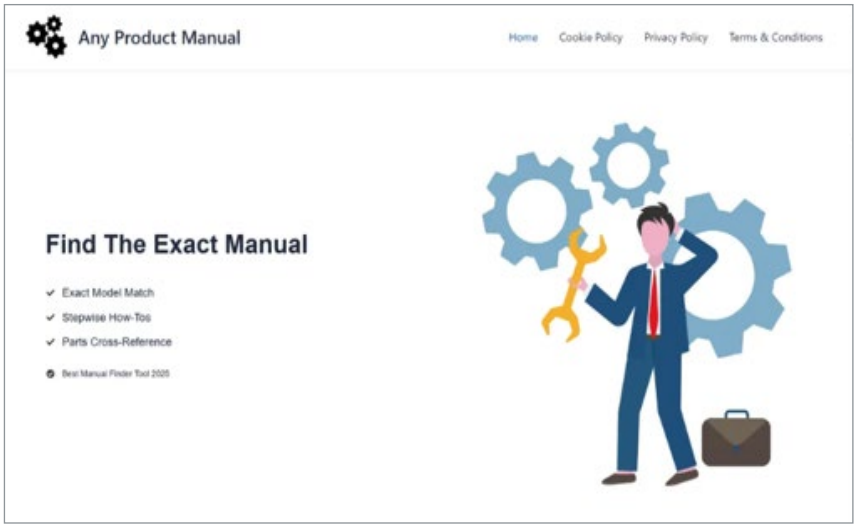
These services are commonly used for legitimate purposes. As a result, domain-based blocking and certificate validation alone may not reliably distinguish malicious activity from authorized use.

TamperedChef[43] provides another example. In November 2025, TRU analyzed this malvertising and SEO campaign, which distributed malicious applications that remained fully functional.

The applications carried valid Extended Validation code-signing certificates and impersonated commonly searched software, including manual readers, PDF editors, chess applications and sudoku games.

A notable aspect of the campaign was the systematic acquisition of code-signing certificates. The operators used a rotating network of U.S.-registered shell companies established under interchangeable digital-marketing names to obtain certificates. When a certificate was revoked, the operators replaced it with one obtained through another company.

Acronis telemetry showed that approximately 80% of victims were in the United States, with concentrations in health care, construction and manufacturing. Employees in these industries frequently search online for manuals associated with specialized equipment.



The findings reinforce an important distinction: A valid signature identifies the signer and confirms the file has not changed since signing. It does not establish that the software is safe. Likewise, the reputation of a hosting platform does not guarantee that content delivered through it is legitimate.

Gone out the window: Ransomware reaches Linux and ESXi

Linux supports a substantial share of enterprise computing infrastructure. Microsoft reports[44] that more than two-thirds of customer cores on Azure run Linux

and that Microsoft 365, GitHub and ChatGPT all rely on Linux-based infrastructure.

Container adoption further contributes to Linux's importance. The Cloud Native Computing Foundation's (CNCF) latest annual survey[45] found that 82% of container users run Kubernetes in production, compared with 66% in 2023. Kubernetes nodes are overwhelmingly Linux.

AI infrastructure is contributing to this trend. According to CNCF, 66% of organizations hosting generative AI models use Kubernetes for at least some inference workloads. CNCF consequently describes Kubernetes as the de facto operating system for AI.

The growing importance of Linux and virtualization infrastructure has made these environments increasingly attractive targets for ransomware operators.

LockBit 5.0 includes Windows, Linux and ESXi variants. On ESXi, it searches for virtual-machine files under /vmfs/ and can shut down running machines so their files become available for encryption.

INC's Rust-based Linux / ESXi encryptor extends these capabilities, using vim-cmd to enumerate and shut down virtual machines and delete snapshots.

Attackers are also exploiting the boundary operating systems. Qilin used Windows Subsystem for Linux to execute a Linux encryptor on a Windows host.

Together, these cases demonstrate how ransomware operators are adapting to the infrastructure they intend to disrupt and exploiting gaps in monitoring coverage.

For defenders, visibility into Linux and ESXi must extend beyond detecting encryption activity. Unexpected virtual-machine shutdowns and snapshot deletion can provide early indications of an attack targeting systems essential to recovery.



Recommendations: Turning threat insights into stronger defenses

The investigations presented in this report reveal a recurring weakness in incident response: Organizations may focus on the first compromised device or account while overlooking the access it provides to other systems and services.

For MSPs, a single compromised identity can expose multiple clients. In small and medium IT environments, shared administrative accounts and infrastructure can connect several business services.

Faster attack preparation increases the urgency of the response, while interference with security controls and recovery mechanisms can make delays more costly.

Service providers and internal IT teams should therefore organize their response around those connections. They must understand what a compromised identity or system can access, interrupt that access and maintain the ability to recover essential services while investigations continue.

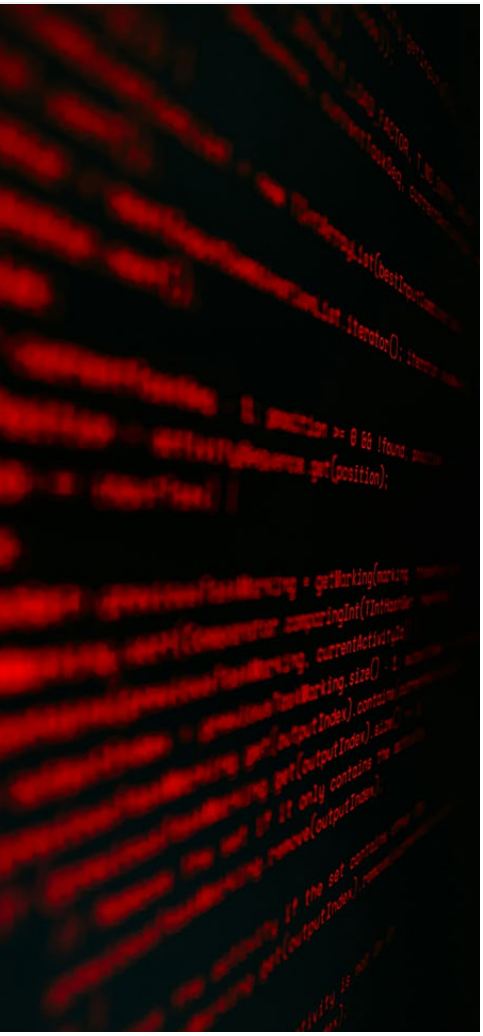
Treat identity compromise as an incident across connected services

When a compromised device contains administrative sessions, responders must trace those identities into cloud applications, remote management tools and client environments.

Endpoint remediation should not close an incident while access obtained from that device remains active elsewhere. Endpoint, identity and cloud investigations should



Organizations should separate routine and privileged accounts and restrict each role to the environments it requires and remove unused integrations.



have a shared owner responsible for establishing the scope of the incident, revoking access and checking for subsequent misuse.

Organizations should separate routine and privileged accounts, restrict each role to the environments it requires and remove unused integrations.

Phishing-resistant authentication also requires a recovery process that cannot be bypassed through an unverified help-desk request.

Organizations should practice revoking sessions and application access before an incident occurs, so responders understand which actions take effect immediately and which require further intervention.

Match remediation urgency to the access behind the vulnerability

A vulnerability affecting a system that holds deployment secrets or provides access to client environments can have consequences beyond the compromised host.

Organizations should prioritize exposed systems according to the access they provide, the severity of the vulnerability and evidence of exploitation.

They should establish who can authorize emergency changes outside scheduled maintenance windows and determine what access can be restricted when patching must be delayed.

Following suspected exploitation, applying a patch addresses the software vulnerability but may not eliminate access the attacker has already established.

Responders should review accounts, configuration changes and persistence; identify secrets stored on the system; and investigate their use elsewhere.

Credentials that could provide access across multiple clients should be rotated before the security investigation is closed.

Evaluate legitimate tools based on their use

An approved remote tool or familiar cloud service can still facilitate unauthorized activity. Organizations should establish which operators normally use each tool, which systems and clients they access, and how the tools are deployed. New installations, unfamiliar administrators or identical scripts appearing across multiple tenants should be evaluated against authorized activity.

Obsolete agents and accounts should be removed following provider changes, and organizations should confirm that access has been terminated.

On sensitive systems, unexpected processes calling cloud APIs or launching downloaded content warrant investigation, even when the destination or application is trusted. A valid signature or familiar domain cannot establish that the activity is authorized.

Make the loss of protection actionable

A security control may not be able to report that it has been disabled. Organizations should independently monitor for agents that stop communicating, unexpected changes to logging and modifications to protected settings.

An agent failure accompanied by a privileged login, new driver or remote execution should be escalated from routine support triage to a security investigation.

On management systems, responders should assess which client environments could have been accessed while visibility was unavailable.

Relevant logs should be retained outside affected hosts. Organizations should restrict permissions to modify security settings and review exceptions to driver restrictions. If visibility cannot be restored promptly, responders need predefined criteria for restricting the affected system's access while determining what occurred.

Design recovery for situations where normal administration is compromised

Recovery testing should not assume that the usual administrator account, management server or virtualization platform remains trustworthy. Multiple applications may depend on the same identity service, host or network configuration. Restoring their data individually does not necessarily mean the business can resume operations.

Providers should agree with each client which services must be restored first and who can authorize each stage of recovery.

Backup administration should remain separate from routine endpoint access, with stronger approval requirements for deletion or retention changes.

Protected copies, recovery credentials and essential configuration records must remain accessible during a production compromise.

Organizations should test the complete recovery process, from



authorization through restoration of a functioning service. They should document delays and use the results to validate recovery commitments.

Investigations into potential data theft must continue during restoration. Bringing a service back online does not establish what information an intruder may have copied.

Keep automation within acceptable authority boundaries

AI agents can exercise existing permissions quickly and across multiple systems.

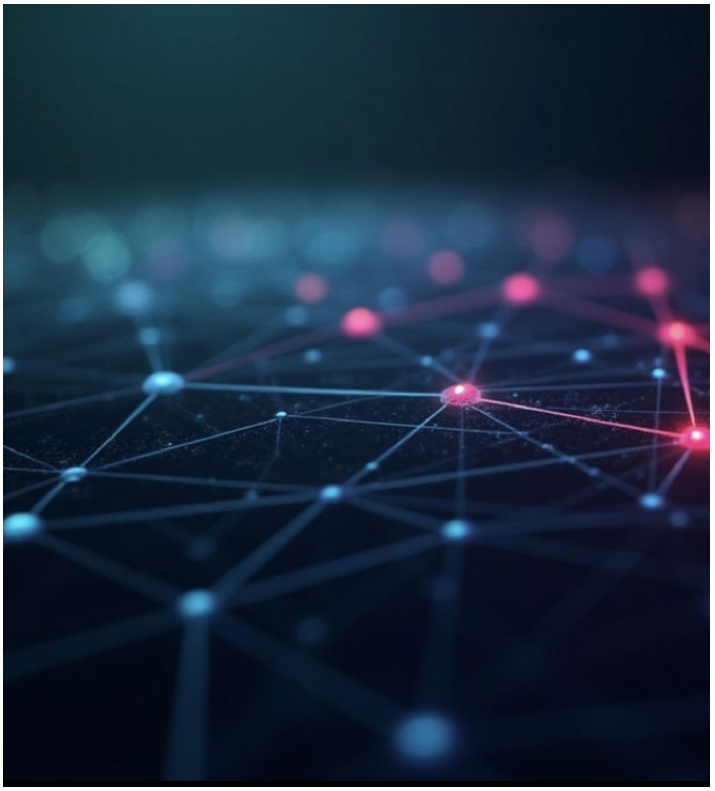
Access should reflect the agent's assigned task. Summarizing support tickets, reading a repository and executing scripts require different levels of authority.

For MSPs, client separation must extend to retrieved data, retained context, credentials and tool execution.

These boundaries should be enforced through connected systems and tool permissions.

Organizations should require independent approval for broad deployments, identity changes and backup deletion, and review external skills before adoption.

Logs should connect each request to the data retrieved, the tools invoked and the identity authorizing the action.



Establish response decisions before an incident

Across the threats examined in this report, effective response depends on decisions that organizations can make in advance.

Providers and clients should agree on who can revoke delegated access, isolate systems, preserve evidence and initiate recovery. They also need communication channels that remain available if normal email or ticketing systems are compromised.

Organizations should test these decisions together through scenarios involving stolen administrative access, a loss of security visibility and attempts to interfere with recovery.

Exercises should measure how long it takes to establish the scope of an incident, terminate access and restore critical services.

The results can reveal dependencies that allow one compromise to spread and identify where additional controls will deliver the greatest practical benefit.

Conclusion

The investigations described in this report demonstrate how threat actors are adapting their techniques, infrastructure and operating models to exploit increasingly interconnected IT environments.

Although their objectives vary, several recurring patterns emerge. Red Heron exploited vulnerable development servers to obtain repositories and credentials. Espionage operators used familiar cloud services to direct malware and trusted relationships to reach downstream targets. Cybercriminal groups transformed stolen sessions and credentials into access that could be traded, exploited for extortion or used to support further intrusions.

These activities share a common characteristic: An initial compromise can provide access far beyond the original entry point.

AI is accelerating this process by enabling threat actors to automate portions of the attack lifecycle and coordinate operations across multiple agents.

The BOSSAI investigation demonstrated how a single operator could connect models, offensive tools and messaging platforms to coordinate activity across targets. Reusable agent skills preserved procedures and operational context.

At the same time, malicious skills and exposed agent gateways created additional attack paths into organizations adopting agentic technologies.

The resulting risk depends not only on the model behavior but also on the permissions granted to agents, the tools they can invoke, and the boundaries enforced by their surrounding infrastructure.

The latter stages of intrusion are also changing. Identity compromise is becoming equally significant. Infostealers and vishing attacks provide attackers with credentials, authentication tokens and active sessions that can be used to access cloud applications and connected services.

Threat actors are also developing new ways to undermine security controls and recovery capabilities. EDR killers and other defense impairment techniques reduce visibility during an intrusion, while interference with backups and recovery infrastructure complicates restoration.

The expansion of ransomware capabilities to Linux and ESXi further increases the potential impact by exposing infrastructure that hosts multiple workloads and critical business services.

A working backup or an installed endpoint agent offers limited reassurance if an intruder can disable the agent, delete snapshots or access the backup platform through a compromised administrative account.

The growing abuse of legitimate services introduces another challenge. Both espionage operators and cybercriminal groups are exploiting trusted platforms such as GitHub and Telegram to distribute malware, maintain command-and-control channels and conceal malicious activity.

Being able to completely circumvent detection by abusing a potentially allowlisted service or site allows attackers to stay under the radar and requires additional vigilance from security teams.



Organizations need clear visibility into the identities, integrations and remote management tools that connect client environments.

For MSPs, their customers, and the cybersecurity community, these findings reinforce the importance of limiting the reach of trusted access and detecting its misuse.

When compromise is confirmed, responders must trace the affected access into every environment it could have reached, determine where it was used and verify that recovery capabilities remain intact.

Footnotes

[1] Kimhy, Eliad, Pontiroli, Santiago and Singha, Subjaheet. “BOSSAI: Inside the AI attack mesh that targeted Taiwan’s emergency services.” October 21, 2026. Acronis TRU. acronis.com/en/tru/posts/bossai-inside-the-ai-attack-mesh-that-targeted-taiwans-emergency-services

[2] Aizad, Syed, Singha, Subjaheet, and Virtusio, Darrel. “Poisoning the well: AI supply chain attacks on Hugging Face and Open Claw.” April 30, 2026. Acronis TRU. acronis.com/en/tru/posts/poisoning-the-well-ai-supply-chain-attacks-on-hugging-face-and-openclaw/

[3] Anthropic. “Detecting and countering misuse of AI: September 2026.” anthropic.com/threat-intelligence-report-september-2026

[4] OpenAI. “The Hugging Face incident and the road ahead.” August 26, 2026. openai.com/index/hugging-face-incident-and-the-road-ahead/

[5] Anthropic. “Investigating three real-world incidents in our cybersecurity evaluations.” July 30, 2026. anthropic.com/research/investigating-incidents-cybersecurity-evals

[6] Acronis Threat Research Unit. “OpenClaw: Agentic AI in the wild – Architecture, adoption and security risks.” February 23, 2026. Acronis TRU. acronis.com/en/tru/posts/openclaw-agentic-ai-in-the-wild-architecture-adoption-and-emerging-security-risks/

[7] Cutler, Silas. “OpenClaw in the Wild: Mapping the Public Exposure of a Viral AI Assistant.” January 31, 2026. Censys. censys.com/blog/openclaw-in-the-wild-mapping-the-public-exposure-of-a-viral-ai-assistant/

[8] Kimhy, Eliad, Singha, Subjaheet, and Virtusio, Darrel. “CRESCENTHARVEST: Iranian protestors and dissidents targeted in cyberespionage campaign.” Acronis TRU. February 26, 2026. acronis.com/en/tru/posts/crescentharvest-iranian-protestors-and-dissidents-targeted-in-cyberespionage-campaign/

[9] Microsoft Threat Intelligence. “New Russia-affiliated actor Void Blizzard targets critical sectors for espionage.” May 7, 2025. Microsoft Security blog. microsoft.com/en-us/security/blog/2025/05/27/new-russia-affiliated-actor-void-blizzard-targets-critical-sectors-for-espionage/

[10] Alegre, David Catalan, Kimhy, Eliad, Pontiroli, Santiago, and Virtusio, Darrel. “The DragonForce cartel: Scattered Spider at the gate.” November 4, 2025. Acronis TRU. acronis.com/en/tru/posts/the-dragonforce-cartel-scattered-spider-at-the-gate/

[11] Mandiant. “Vishing for Access: Tracking the Expansion of ShinyHunters-Branded SaaS Data Theft.” January 30, 2026. Google Cloud Blog. cloud.google.com/blog/topics/threat-intelligence/expansion-shinyhunters-saas-data-theft

[12] Acronis Threat Research Unit. “The remote access blind spot: An analysis of RMM tool risks for MSPs.” May 13, 2026. Acronis TRU. acronis.com/en/tru/posts/the-remote-access-blind-spot-an-analysis-of-rmm-tool-risk-for-smbs/

[13] The MITRE Corporation. “Defense Impairment.” attack.mitre.org/tactics/TA0112/

[14] Kimhy, Eliad and Pontiroli, Santiago. “Inside Qilin ransomware: ESXi attacks, affiliate networks and double extortion.” October 28, 2026. Acronis TRU. acronis.com/en/tru/posts/inside-qilin-ransomware-esxi-attacks-affiliate-networks-and-double-extortion/

[15] Takeda, Takahiro and Unterbrink, Holger. “Qilin EDR killer infection chain.” April 2, 2026. Cisco Talos blog. blog.talosintelligence.com/qilin-edr-killer/

[16] Pontiroli, Santiago, Tan, Yu Kai, and Virtusio, Darrel. “The Gentlemen: A close look at a not-so gentle ransomware-as-a-service operation.” October 21, 2026.

[17] Soucek, Jakub. “Killing me gently: Inside Gentlemen’s EDR killer framework.” June 18, 2026. Welivesecurity by ESET blog. welivesecurity.com/en/eset-research/killing-me-gently-inside-gentlemens-edr-killer-framework/

[18] The MITRE Corporation. “Inhibit System Recovery.” attack.mitre.org/techniques/T1490/

[19] Ponitroli, Santiago and Singha, Subjaheet. “Mustang Panda targets India’s government and energy sectors with ZOHOMURK and MINIRECON.” June 28, 2026. Acronis TRU. acronis.com/en/tru/posts/mustang-panda-targets-indias-government-and-energy-sectors/

[20] Gegeny, Jozsef, Bunimovich, Vadim and Singha, Subjaheet. “New year, new sector: Transparent Tribe targets India’s startup ecosystem.” February 4, 2026. Acronis TRU. acronis.com/en/tru/posts/new-year-new-sector-transparent-tribe-targets-indias-startup-ecosystem/

[21] Singha, Subjaheet. “Behind Khmer Shadow: Targeted espionage against Cambodian government entities.” June 10, 2026. Acronis TRU. acronis.com/en/tru/posts/behind-khmer-shadow-targeted-espionage-against-cambodian-government-entities/

[22] The MITRE Corporation. “Phishing.” attack.mitre.org/techniques/T1566/

[23] Pontiroli, Santiago, Singha, Subjaheet and Virtusio, Darrel. “PATCHCORD: New malware cluster targets Afghan telecom and South Asian critical infrastructure.” August 13, 2026. Acronis TRU. acronis.com/en/tru/posts/patchcord-new-malware-cluster-targets-afghan-telecom-and-south-asian-critical-infrastructure/

[24] Google Threat Intelligence Group and Mandiant. “Exposing the Undercurrent: Disrupting the GRIDTIDE Global Cyber Espionage Campaign.” February 25, 2026. Google Cloud blog. cloud.google.com/blog/topics/threat-intelligence/disrupting-gridtide-global-espionage-campaign/

[25] United States Federal Bureau of Investigation (FBI). “FBI Seeking Tips about PRC Targeting of U.S. Telecommunications.” April 24, 2025. FBI.gov. fbi.gov/investigate/cyber/alerts/2025/fbi-seeking-tips-about-prc-targeting-of-us-telecommunications

[26] Microsoft Threat Intelligence. “Silk Typhoon targeting IT supply chain.” March 5, 2025. Microsoft Security blog. microsoft.com/en-us/security/blog/2025/03/05/silk-typhoon-targeting-it-supply-chain/

[27] Singha, Subjaheet. “Red Heron exploits Gitea n-day flaw in multinational campaign, exposing new Linux rootkit.” September 13, 2026. Acronis TRU. acronis.com/en/tru/posts/red-heron-exploits-gitea-n-day-flaw-in-multinational-campaign-exposing-new-linux-rootkit/

[28] The MITRE Corporation. “Exploit Public-Facing Application.” attack.mitre.org/techniques/T1190/

[29] Acronis Threat Research Unit. "The patch reality gap: How fast SMBs really patch and what are the risks." March 2, 2026. Acronis TRU. acronis.com/en/tru/posts/the-patch-reality-gap-how-fast-smbs-really-patch-and-what-are-the-risks/

[30] Cisco. "Continued Evolution of Persistence Mechanism Against Cisco Secure Firewall Adaptive Security Appliance and Secure Firewall Threat Defense." May 19, 2026. Cisco Security blog. sec.cloudapps.cisco.com/security/center/content/CiscoSecurityAdvisory/cisco-sa-asaftd-persist-CISAED25-03

[31] Windsor, Carl. "Analysis of Reported Credential Compromise of FortiGate Devices." June 19, 2026. Fortinet blog. fortinet.com/blog/psirt-blogs/analysis-of-reported-credential-compromise-of-fortigate-devices

[32] Notepad++. "Notepad++ Hijacked by State-Sponsored Hackers." February 2, 2026. notepad-plus-plus.org/news/hijacked-incident-info-update/

[33] Rapid7. "Chrysalis, Notepad++, and Supply Chain Risk: What it Means, and What to Do Next." February 5, 2026. Rapid7 blog. rapid7.com/blog/post/tr-chrysalis-notepad-supply-chain-risk-next-steps/

[34] Acronis Threat Research Unit. "LockBit strikes with new 5.0 version, targeting Windows, Linux and ESXI systems." February 12, 2026. Acronis TRU. acronis.com/en/tru/posts/lockbit-strikes-with-new-50-version-targeting-windows-linux-and-esxi-systems/

[35] Virtusio, Darrel. "From emerging threat to top-tier ransomware-as-a-service: The evolution of INC ransomware." June 17, 2026. acronis.com/en/tru/posts/from-emerging-threat-to-top-tier-ransomware-as-a-service-the-evolution-of-inc-ransomware/

[36] Kimhy, Eliad and Virtusio, Darrel. "Vidar Stealer 2.0 distributed via fake game cheats on GitHub and Reddit." March 17, 2026. Acronis TRU. acronis.com/en/tru/posts/vidar-stealer-20-distributed-via-fake-game-cheats-on-github-and-reddit/

[37] Acronis. "Acronis Cyberthreats Report H2 2025." February 18, 2026.acronis.com/en/search/?query=cyberthreats+report

[38] Cloud Security Alliance. "IronWorm: eBPF Rootkit and Tor C2 Target npm Supply Chain." June 7, 2026. CSA Lab Space blog. labs.cloudsecurityalliance.org/research/csa-research-note-ironworm-ebpf-rootkit-npm-supply-chain-202/

[39] CERT-EU. "Cyber Brief May 2026." June 2, 2026. cert.europa.eu/publications/threat-intelligence/cb26-06/

[40] CERT-EU. "Cyber Brief June 2026." July 1, 2026. cert.europa.eu/publications/threat-intelligence/cb26-07/

[41] Kimhy, Eliad. "Fake adult websites pop realistic Windows Update screen to deliver stealers via ClickFix." November 25, 2025. Acronis TRU. acronis.com/en/tru/posts/fake-adult-websites-pop-realistic-windows-update-screen-to-deliver-stealers-via-clickfix/

[42] Kimhy, Eliad. "FileFix in the wild! New FileFix campaign goes beyond POC and leverages steganography." September 16, 2026. Acronis TRU. acronis.com/en/tru/posts/filefix-in-the-wild-new-filefix-campaign-goes-beyond-poc-and-leverages-steganography/

[43] Virtusio, Darrel. "Cooking up trouble: How TamperedChef uses signed apps to deliver stealthy payloads." November 19, 2025. Acronis TRU. acronis.com/en/tru/posts/cooking-up-trouble-how-tamperedchef-uses-signed-apps-to-deliver-stealthy-payloads/

[44] Microsoft. "Linux virtual machines in Azure." azure.microsoft.com/en-gb/products/virtual-machines/linux/

[45] Cloud Native Computing Foundation. "The CNCF Annual Cloud Native Survey: The Infrastructure of AI's Future." January 20, 2026. cncf.io/reports/the-cncf-annual-cloud-native-survey/

Authors



Santiago Pontiroli,
Threat Intelligence Research Lead



Alexander Ivanyuk,
Senior Director, Technology



Eliad Kimhy,
Senior TRU Researcher



Subhajeet Singha,
Senior TRU Researcher

Acronis
Threat Research Unit



Copyright © 2003-2026 Acronis International GmbH. All rights reserved. Acronis and the Acronis logo are trademarks of Acronis International GmbH in the United States and/or other countries. All other trademarks or registered trademarks are the property of their respective owners. Technical changes and differences from the illustrations are reserved; errors are excepted.

2026