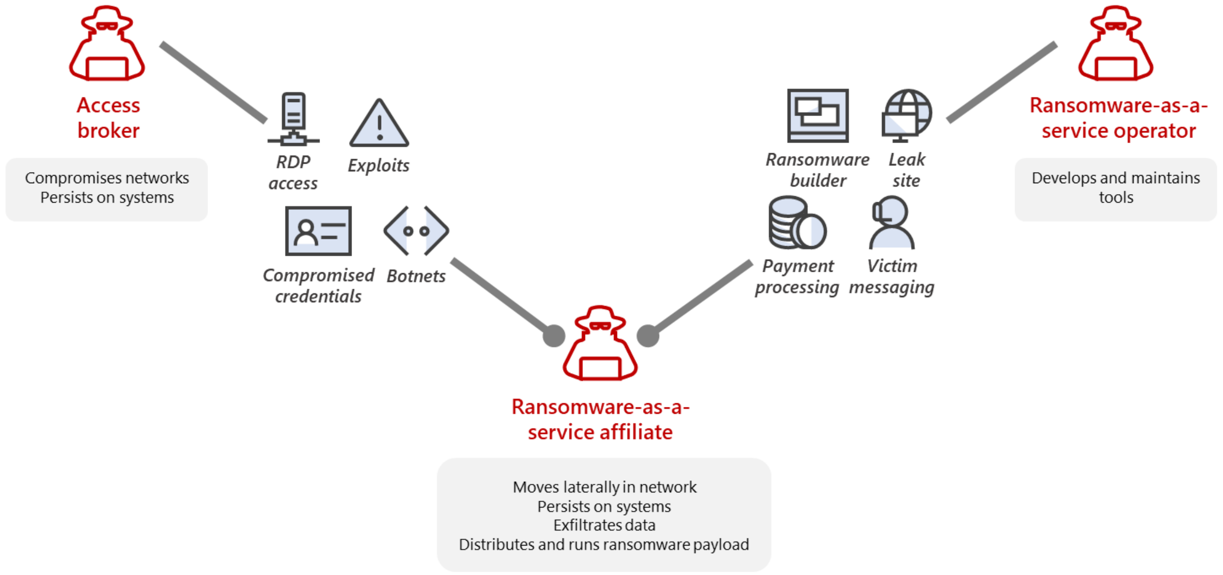
Ransomware-as-a-service: Understanding the cybercrime gig economy and how to protect yourself

microsoft.com/security/blog/2022/05/09/ransomware-as-a-service-understanding-the-cybercrime-gig-economy-and-how-to-protect-yourself/

May 9, 2022



Microsoft processes 24 trillion signals every 24 hours, and we have blocked billions of attacks in the last year alone. Microsoft Security tracks more than 35 unique ransomware families and 250 unique threat actors across observed nation-state, ransomware, and criminal activities.

That depth of signal intelligence gathered from various domains—identity, email, data, and cloud—provides us with insight into the gig economy that attackers have created with tools designed to lower the barrier for entry for other attackers, who in turn continue to pay dividends and fund operations through the sale and associated “cut” from their tool’s success.

The cybercriminal economy is a continuously evolving connected ecosystem of many players with different techniques, goals, and skillsets. In the same way our traditional economy has shifted toward gig workers for efficiency, criminals are learning that there’s less work and less risk involved by renting or selling their tools for a portion of the profits than performing the attacks themselves. This industrialization of the cybercrime economy has made it easier for attackers to use ready-made penetration testing and other tools to perform their attacks.

Within this category of threats, Microsoft has been tracking the trend in the ransomware-as-a-service (RaaS) gig economy, called human-operated ransomware, which remains one of the most impactful threats to organizations. We coined the industry term “human-operated ransomware” to clarify that these threats are driven by humans who make decisions at every stage of their attacks based on what they find in their target’s network.

Unlike the broad targeting and opportunistic approach of earlier ransomware infections, attackers behind these human-operated campaigns vary their attack patterns depending on their discoveries—for example, a security product that isn‘t configured to prevent tampering or a service that’s running as a highly privileged account like a domain admin. Attackers can use those weaknesses to elevate their privileges to steal even more valuable data, leading to a bigger payout for them—with no guarantee they’ll leave their target environment once they’ve been paid. Attackers are also often more determined to stay on a network once they gain access and sometimes repeatedly monetize that access with additional attacks using different malware or ransomware payloads if they aren’t successfully evicted.

Ransomware attacks have become even more impactful in recent years as more ransomware-as-a-service ecosystems have adopted the double extortion monetization strategy. All ransomware is a form of extortion, but now, attackers are not only encrypting data on compromised devices but also exfiltrating it and then posting or threatening to post it publicly to pressure the targets into paying the ransom. Most ransomware attackers opportunistically deploy ransomware to whatever network they get access to, and some even purchase access to networks from other cybercriminals. Some attackers prioritize organizations with higher revenues, while others prefer specific industries for the shock value or type of data they can exfiltrate.

All human-operated ransomware campaigns—all human-operated attacks in general, for that matter—share common dependencies on security weaknesses that allow them to succeed. Attackers most commonly take advantage of **an organization’s poor credential hygiene and legacy configurations or misconfigurations to find easy entry and privilege escalation points in an environment.**

In this blog, we detail several of the ransomware ecosystems using the RaaS model, the importance of cross-domain visibility in finding and evicting these actors, and best practices organizations can use to protect themselves from this increasingly popular style of attack. We also offer security best practices on credential hygiene and cloud hardening, how to address security blind spots, harden internet-facing assets to understand your perimeter, and more. Here’s a quick table of contents:

How RaaS redefines our understanding of ransomware incidents

With ransomware being the preferred method for many cybercriminals to monetize attacks, human-operated ransomware remains one of the most impactful threats to organizations today, and it only continues to evolve. This evolution is driven by the “human-operated” aspect of these attacks—attackers make informed and calculated decisions, resulting in varied attack patterns tailored specifically to their targets and iterated upon until the attackers are successful or evicted.

In the past, we've observed a tight relationship between the initial entry vector, tools, and ransomware payload choices in each campaign of one strain of ransomware. The RaaS affiliate model, which has allowed more criminals, regardless of technical expertise, to deploy ransomware built or managed by someone else, is weakening this link. As ransomware deployment becomes a gig economy, it has become more difficult to link the tradecraft used in a specific attack to the ransomware payload developers.

Reporting a ransomware incident by assigning it with the payload name gives the impression that a monolithic entity is behind all attacks using the same ransomware payload and that all incidents that use the ransomware share common techniques and infrastructure. However, focusing solely on the ransomware stage obscures many stages of the attack that come before, including actions like data exfiltration and additional persistence mechanisms, as well as the numerous detection and protection opportunities for network defenders.

We know, for example, that the underlying techniques used in human-operated ransomware campaigns haven't changed very much over the years—attacks still prey on the same security misconfigurations to succeed. Securing a large corporate network takes disciplined and sustained focus, but there's a high ROI in implementing critical controls that prevent these attacks from having a wider impact, even if it's only possible on the most critical assets and segments of the network.

Without the ability to steal access to highly privileged accounts, attackers can't move laterally, spread ransomware widely, access data to exfiltrate, or use tools like Group Policy to impact security settings. Disrupting common attack patterns by applying security controls also reduces alert fatigue in security SOCs by stopping the attackers before they get in. This can also prevent unexpected consequences of short-lived breaches, such as exfiltration of network topologies and configuration data that happens in the first few minutes of execution of some trojans.

In the following sections, we explain the RaaS affiliate model and disambiguate between the attacker tools and the various threat actors at play during a security incident. Gaining this clarity helps surface trends and common attack patterns that inform defensive strategies focused on preventing attacks rather than detecting ransomware payloads. Threat intelligence and insights from this research also enrich our solutions like Microsoft 365 Defender, whose comprehensive security capabilities help protect customers by detecting RaaS-related attack attempts.

The RaaS affiliate model explained

The cybercriminal economy—a connected ecosystem of many players with different techniques, goals, and skillsets—is evolving. The industrialization of attacks has progressed from attackers using off-the-shelf tools, such as Cobalt Strike, to attackers being able to purchase access to networks and the payloads they deploy to them. This means that the impact of a successful ransomware and extortion attack remains the same regardless of the attacker's skills.

RaaS is an arrangement between an operator and an affiliate. The RaaS operator develops and maintains the tools to power the ransomware operations, including the builders that produce the ransomware payloads and payment portals for communicating with victims. The RaaS program may also include a leak site to share snippets of data exfiltrated from victims, allowing attackers to show that the exfiltration is real and try to extort payment. Many RaaS programs further incorporate a suite of extortion support offerings, including leak site hosting and integration into ransom notes, as well as decryption negotiation, payment pressure, and cryptocurrency transaction services

RaaS thus gives a unified appearance of the payload or campaign being a single ransomware family or set of attackers. However, what happens is that the RaaS operator sells access to the ransom payload and decryptor to an affiliate, who performs the intrusion and privilege escalation and who is responsible for the deployment of the actual ransomware payload. The parties then split the profit. In addition, RaaS developers and operators might also use the payload for profit, sell it, and run their campaigns with other ransomware payloads—further muddying the waters when it comes to tracking the criminals behind these actions.

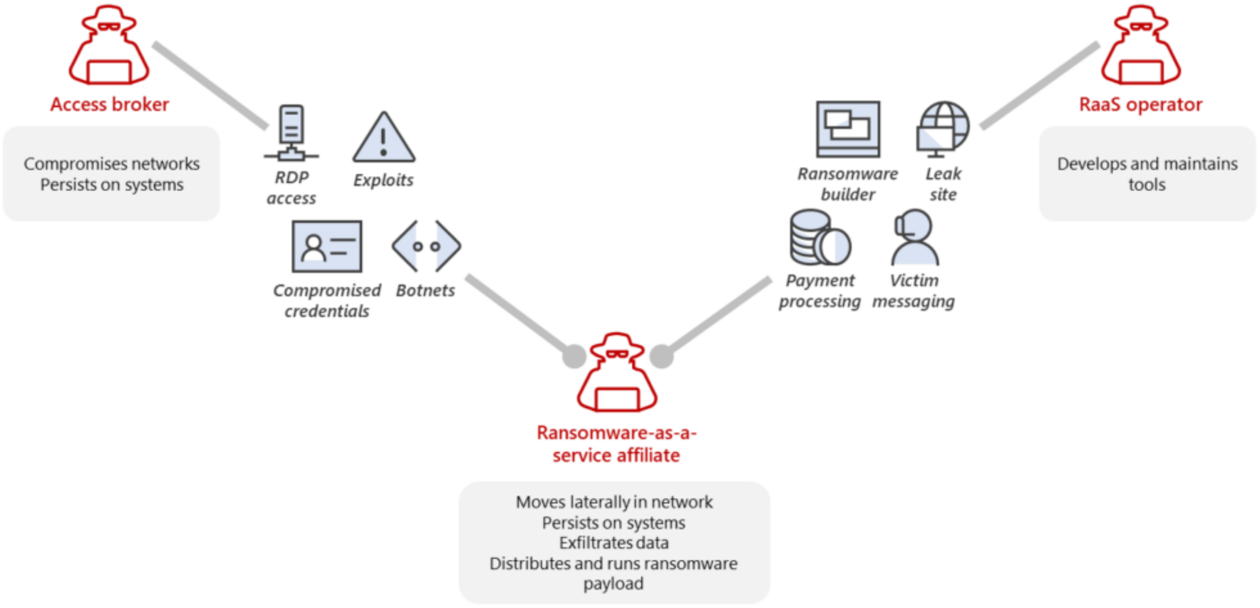


Figure 1. How the RaaS affiliate model enables ransomware attacks

Access for sale and mercurial targeting

A component of the cybercriminal economy is selling access to systems to other attackers for various purposes, including ransomware. Access brokers can, for instance, infect systems with malware or a botnet and then sell them as a “load”. A load is designed to install other malware or backdoors onto the infected systems for other criminals. Other access brokers scan the internet for vulnerable systems, like exposed Remote Desktop Protocol (RDP) systems with weak passwords or unpatched systems, and then compromise them *en masse* to “bank” for later profit. Some advertisements for the sale of initial access specifically cite that a system isn’t managed by an antivirus or endpoint detection and response (EDR) product and has a highly privileged credential such as Domain Administrator associated with it to fetch higher prices.

Most ransomware attackers opportunistically deploy ransomware to whatever network they get access to. Some attackers prioritize organizations with higher revenues, while some target specific industries for the shock value or type of data they can exfiltrate (for example, attackers targeting hospitals or exfiltrating data from technology companies). In many cases, the targeting doesn’t manifest itself as specifically attacking the target’s network, instead, the purchase of access from an access broker or the use of existing malware infection to pivot to ransomware activities.

In some ransomware attacks, the affiliates who bought a load or access may not even know or care how the system was compromised in the first place and are just using it as a “jump server” to perform other actions in a network. Access brokers often list the network details for the access they are selling, but affiliates aren’t usually interested in the network itself but rather the monetization potential. As a result, some attacks that seem targeted to a specific industry might simply be a case of affiliates purchasing access based on the number of systems they could deploy ransomware to and the perceived potential for profit.

“Human-operated” means human decisions

Microsoft coined the term “human-operated ransomware” to clearly define a class of attacks driven by expert human intelligence at every step of the attack chain and culminate in intentional business disruption and extortion. Human-operated ransomware attacks share commonalities in the security misconfigurations of which they take advantage and the manual techniques used for lateral movement and persistence. However, the human-operated nature of these actions means that variations in attacks—including objectives and pre-ransom activity—evolve depending on the environment and the unique opportunities identified by the attackers.

These attacks involve many reconnaissance activities that enable human operators to profile the organization and know what next steps to take based on specific knowledge of the target. Many of the initial access campaigns that provide access to RaaS affiliates perform automated reconnaissance and exfiltration of information collected in the first few minutes of an attack.

After the attack shifts to a hands-on-keyboard phase, the reconnaissance and activities based on this knowledge can vary, depending on the tools that come with the RaaS and the operator’s skill. Frequently attackers query for the currently running security tools, privileged users, and security settings such as those defined in Group Policy before continuing their attack. The data discovered via this reconnaissance phase informs the attacker’s next steps.

If there’s minimal security hardening to complicate the attack and a highly privileged account can be gained immediately, attackers move directly to deploying ransomware by editing a Group Policy. The attackers take note of security products in the environment and attempt to tamper with and disable these, sometimes using scripts or tools provided with RaaS purchase that try to disable multiple security products at once, other times using specific commands or techniques performed by the attacker.

This human decision-making early in the reconnaissance and intrusion stages means that even if a target’s security solutions detect specific techniques of an attack, the attackers may not get fully evicted from the network and can use other collected knowledge to attempt to continue the attack in ways that bypass security controls. In many instances, attackers test their attacks “in production” from an undetected location in their target’s environment, deploying tools or payloads like commodity malware. If these tools or payloads are detected and blocked by an antivirus product, the attackers simply grab a different tool, modify their payload, or tamper with the security products they encounter. Such detections could give SOCs a false sense of security that their existing solutions are working. However, these could merely serve as a smokescreen to allow the attackers to further tailor an attack chain that has a higher probability of success. Thus, when the attack reaches the active attack stage of deleting backups or shadow copies, the attack would be minutes away from ransomware deployment. The adversary would likely have already performed harmful actions like the exfiltration of data. This knowledge is key for SOCs responding to ransomware: prioritizing investigation of alerts or detections of tools like Cobalt Strike and performing swift remediation actions and incident response (IR) procedures are critical for containing a human adversary before the ransomware deployment stage.

Exfiltration and double extortion

Ransomware attackers often profit simply by disabling access to critical systems and causing system downtime. Although that simple technique often motivates victims to pay, it is not the only way attackers can monetize their access to compromised networks. Exfiltration of data and “double extortion,” which refers to attackers threatening to leak data if a ransom hasn’t been paid, has also become a common tactic among many RaaS affiliate programs—many of them offering a unified leak site for their affiliates. Attackers take advantage of common weaknesses to exfiltrate data and demand ransom without deploying a payload.

This trend means that focusing on protecting against ransomware payloads via security products or encryption, or considering backups as the main defense against ransomware, instead of comprehensive hardening, leaves a network vulnerable to all the stages of a human-operated ransomware attack that occur before ransomware deployment. This exfiltration can take the form of using tools like Rclone to sync to an external site, setting up email transport rules, or uploading files to cloud services. With double extortion, attackers don’t need to deploy ransomware and cause downtime to extort money. Some attackers have moved beyond the need to deploy ransomware payloads and are shifting straight to extortion models or performing the destructive objectives of their attacks by directly deleting cloud resources. One such extortion attackers is DEV-0537 (also known as LAPSUS$), which is profiled below.

Persistent and sneaky access methods

Paying the ransom may not reduce the risk to an affected network and potentially only serves to fund cybercriminals. Giving in to the attackers’ demands doesn’t guarantee that attackers ever “pack their bags” and leave a network. Attackers are more determined to stay on a network once they gain access and sometimes repeatedly monetize attacks using different malware or ransomware payloads if they aren’t successfully evicted.

The handoff between different attackers as transitions in the cybercriminal economy occur means that multiple attackers may retain persistence in a compromised environment using an entirely different set of tools from those used in a ransomware attack. For example, initial access gained by a banking trojan leads to a Cobalt Strike deployment, but the RaaS affiliate that purchased the access may choose to use a less detectable remote access tool such as TeamViewer to maintain persistence on the network to operate their broader series of campaigns. Using legitimate tools and settings to persist versus malware implants such as Cobalt Strike is a popular technique among ransomware attackers to avoid detection and remain resident in a network for longer.

Some of the common enterprise tools and techniques for persistence that Microsoft has observed being used include:

- AnyDesk
- Atera Remote Management
- ngrok.io
- Remote Manipulator System
- Splashtop
- TeamViewer

Another popular technique attackers perform once they attain privilege access is the creation of new backdoor user accounts, whether local or in Active Directory. These newly created accounts can then be added to remote access tools such as a virtual private network (VPN) or Remote Desktop, granting remote access through accounts that appear legitimate on the network. Ransomware attackers have also been observed editing the settings on systems to enable Remote Desktop, reduce the protocol’s security, and add new users to the Remote Desktop Users group.

The time between initial access to a hands-on keyboard deployment can vary wildly depending on the groups and their workloads or motivations. Some activity groups can access thousands of potential targets and work through these as their staffing allows, prioritizing based on potential ransom payment over several months. While some activity groups may have access to large and highly resourced companies, they prefer to attack smaller companies for less overall ransom because they can execute the attack within hours or days. In addition, the return on investment is higher from companies that can't respond to a major incident. Ransoms of tens of millions of dollars receive much attention but take much longer to develop. Many groups prefer to ransom five to 10 smaller targets in a month because the success rate at receiving payment is higher in these targets. Smaller organizations that can't afford an IR team are often more likely to pay tens of thousands of dollars in ransom than an organization worth millions of dollars because the latter has a developed IR capability and is likely to follow legal advice against paying. In some instances, a ransomware associate threat actor may have an implant on a network and never convert it to ransom activity. In other cases, initial access to full ransom (including handoff from an access broker to a RaaS affiliate) takes less than an hour.

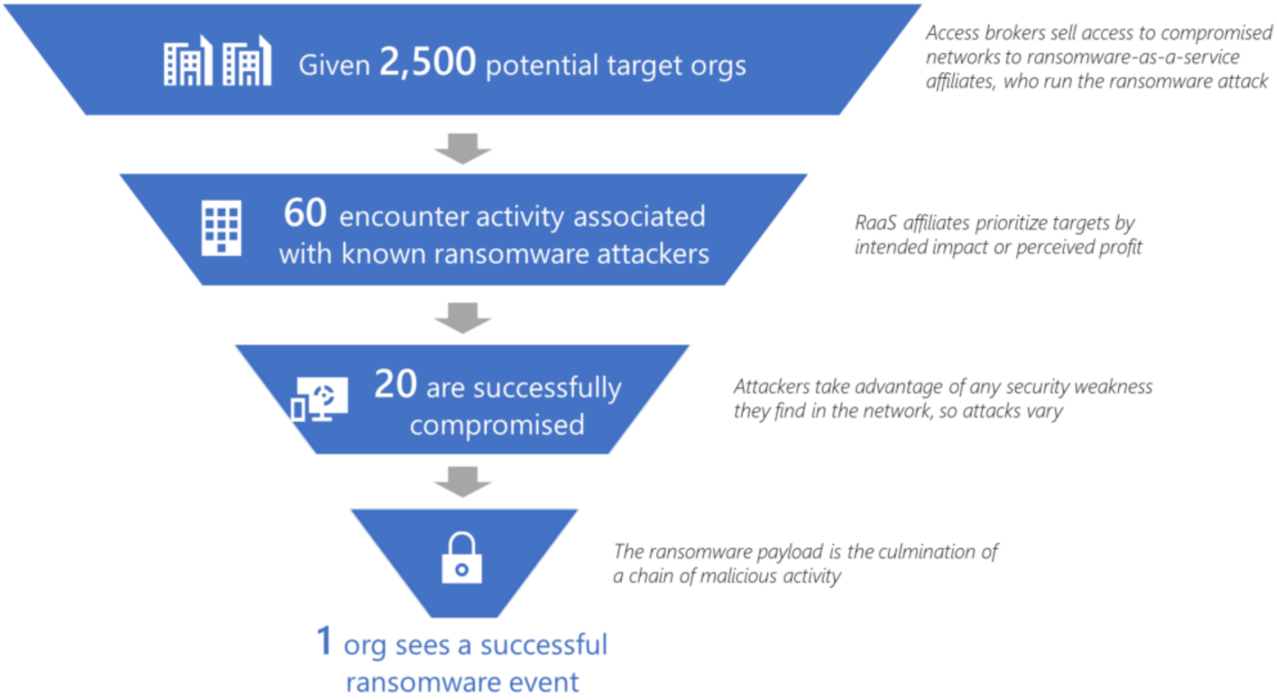

Given 2,500 potential target orgs
Access brokers sell access to compromised networks to ransomware-as-a-service affiliates, who run the ransomware attack
60 encounter activity associated with known ransomware attackers
RaaS affiliates prioritize targets by intended impact or perceived profit
20 are successfully compromised
Attackers take advantage of any security weakness they find in the network, so attacks vary
The ransomware payload is the culmination of a chain of malicious activity
1 org sees a successful ransomware event

Figure 2. Human-operated ransomware targeting and rate of success, based on a sampling of Microsoft data over six months between 2021 and 2022

The human-driven nature of these attacks and the scale of possible victims under control of ransomware-associated threat actors underscores the need to take targeted proactive security measures to harden networks and prevent these attacks in their early stages.

Threat actors and campaigns deep dive: Threat intelligence-driven response to human-operated ransomware attacks

For organizations to successfully respond to evict an active attacker, it’s important to understand the active stage of an ongoing attack. In the early attack stages, such as deploying a banking trojan, common remediation efforts like isolating a system and resetting exposed credentials may be sufficient. As the attack progresses and the attacker performs reconnaissance activities and exfiltration, it’s important to implement an incident response process that scopes the incident to address the impact specifically. Using a threat intelligence-driven methodology for understanding attacks can assist in determining incidents that need additional scoping.

In the next sections, we provide a deep dive into the following prominent ransomware threat actors and their campaigns to increase community understanding of these attacks and enable organizations to better protect themselves:

- DEV-0193 cluster (Trickbot LLC): The most prolific ransomware group today
- ELBRUS: (Un)arrested development
- DEV-0504: Shifting payloads reflecting the rise and fall of RaaS programs
- DEV-0237: Prolific collaborator
- DEV-0206 and DEV-0243: An “evil” partnership
- DEV-0401: China-based lone wolf turned LockBit 2.0 affiliate
- DEV-0537: From extortion to destruction

Microsoft threat intelligence directly informs our products as part of our commitment to track adversaries and protect customers. Microsoft 365 Defender customers should prioritize alerts titled “Ransomware-linked emerging threat activity group detected”. We also add the note “Ongoing hands-on-keyboard attack” to alerts that indicate a human attacker is in the network. When these alerts are raised, it’s highly recommended to initiate an incident response process to scope the attack, isolate systems, and regain control of credentials attackers may be in control of.

A note on threat actor naming: as part of Microsoft’s ongoing commitment to track both nation-state and cybercriminal threat actors, we refer to the unidentified threat actors as a “development group”. We use a naming structure with a prefix of “DEV” to indicate an emerging threat group or unique activity during investigation. When a nation-state group moves out of the DEV stage, we use chemical elements (for example, PHOSPHOROUS and NOBELIUM) to name them. On the other hand, we use volcano names (such as ELBRUS) for ransomware or cybercriminal activity groups that have moved out of the DEV state. In the cybercriminal economy, relationships between groups change very rapidly. Attackers are known to hire talent from other cybercriminal groups or use “contractors,” who provide gig economy-style work on a limited time basis and may not rejoin the group. This shifting nature means that many of the groups Microsoft tracks are labeled as DEV, even if we have a concrete understanding of the nature of the activity group.

DEV-0193 cluster (Trickbot LLC): The most prolific ransomware group today

A vast amount of the current cybercriminal economy connects to a nexus of activity that Microsoft tracks as DEV-0193, also referred to as Trickbot LLC. DEV-0193 is responsible for developing, distributing, and managing many different payloads, including Trickbot, Bazaloader, and AnchorDNS. In addition, DEV-0193 managed the Ryuk RaaS program before the latter's shutdown in June 2021, and Ryuk's successor, Conti as well as Diavol. Microsoft has been tracking the activities of DEV-0193 since October 2020 and has observed their expansion from developing and distributing the Trickbot malware to becoming the most prolific ransomware-associated cybercriminal activity group active today.

DEV-0193's actions and use of the cybercriminal gig economy means they often add new members and projects and utilize contractors to perform various parts of their intrusions. As other malware operations have shut down for various reasons, including legal actions, DEV-0193 has hired developers from these groups. Most notable are the acquisitions of developers from Emotet, Qakbot, and IcedID, bringing them to the DEV-0193 umbrella.

A subgroup of DEV-0193, which Microsoft tracks as DEV-0365, provides infrastructure-as-a-service for cybercriminals. Most notably, DEV-0365 provides Cobalt Strike Beacon-as-a-service. These DEV-0365 Beacons have replaced unique C2 infrastructure in many active malware campaigns. DEV-0193 infrastructure has also been implicated in attacks deploying novel techniques, including exploitation of CVE-2021-40444.

The leaked chat files from a group publicly labeled as the "Conti Group" in February 2022 confirm the wide scale of DEV-0193 activity tracked by Microsoft. Based on our telemetry from 2021 and 2022, Conti has become one of the most deployed RaaS ecosystems, with multiple affiliates concurrently deploying their payload—even as other RaaS ecosystems (DarkSide/BlackMatter and REvil) ceased operations. However, payload-based attribution meant that much of the activity that led to Conti ransomware deployment was attributed to the "Conti Group," even though many affiliates had wildly different tradecraft, skills, and reporting structures. Some Conti affiliates performed small-scale intrusions using the tools offered by the RaaS, while others performed weeks-long operations involving data exfiltration and extortion using their own techniques and tools. One of the most prolific and successful Conti affiliates—and the one responsible for developing the "Conti Manual" leaked in August 2021—is tracked as DEV-0230. This activity group also developed and deployed the FiveHands and HelloKitty ransomware payloads and often gained access to an organization via DEV-0193's BazaLoader infrastructure.

ELBRUS: (Un)arrested development

ELBRUS, also known as FIN7, has been known to be in operation since 2012 and has run multiple campaigns targeting a broad set of industries for financial gain. ELBRUS has deployed point-of-sale (PoS) and ATM malware to collect payment card information from in-store checkout terminals. They have also targeted corporate personnel who have access to sensitive financial data, including individuals involved in SEC filings.

In 2018, this activity group made headlines when three of its members were arrested. In May 2020, another arrest was made for an individual with alleged involvement with ELBRUS. However, despite law enforcement actions against suspected individual members, Microsoft has observed sustained campaigns from the ELBRUS group itself during these periods.

ELBRUS is responsible for developing and distributing multiple custom malware families used for persistence, including JSSLoader and Griffon. ELBRUS has also created fake security companies called “Combi Security” and “Bastion Security” to facilitate the recruitment of employees to their operations under the pretense of working as penetration testers.

In 2020 ELBRUS transitioned from using PoS malware to deploying ransomware as part of a financially motivated extortion scheme, specifically deploying the MAZE and Revil RaaS families. ELBRUS developed their own RaaS ecosystem named DarkSide. They deployed DarkSide payloads as part of their operations and recruited and managed affiliates that deployed the DarkSide ransomware. The tendency to report on ransomware incidents based on payload and attribute it to a monolithic gang often obfuscates the true relationship between the attackers, which is very accurate of the DarkSide RaaS. Case in point, one of the most infamous DarkSide deployments wasn’t performed by ELBRUS but by a ransomware-as-a-service affiliate Microsoft tracks as DEV-0289.

ELBRUS retired the DarkSide ransomware ecosystem in May 2021 and released its successor, BlackMatter, in July 2021. Replicating their patterns from DarkSide, ELBRUS deployed BlackMatter themselves and ran a RaaS program for affiliates. The activity group then retired the BlackMatter ransomware ecosystem in November 2021.

While they aren’t currently publicly observed to be running a RaaS program, ELBRUS is very active in compromising organizations via phishing campaigns that lead to their JSSLoader and Griffon malware. Since 2019, ELBRUS has partnered with DEV-0324 to distribute their malware implants. DEV-0324 acts as a distributor in the cybercriminal economy, providing a service to distribute the payloads of other attackers through phishing and exploit kit vectors. ELBRUS has also been abusing CVE-2021-31207 in Exchange to compromise organizations in April of 2022, an interesting pivot to using a less popular authenticated vulnerability in the ProxyShell cluster of vulnerabilities. This abuse has allowed them to target organizations that patched only the unauthenticated vulnerability in their Exchange Server and turn compromised low privileged user credentials into highly privileged access as SYSTEM on an Exchange Server.

DEV-0504: Shifting payloads reflecting the rise and fall of RaaS programs

An excellent example of how clustering activity based on ransomware payload alone can lead to obfuscating the threat actors behind the attack is DEV-0504. DEV-0504 has deployed at least six RaaS payloads since 2020, with many of their attacks becoming high-profile incidents attributed to the “REvil gang” or “BlackCat ransomware group”. This attribution masks the actions of the set of the attackers in the DEV-0504 umbrella, including other REvil and BlackCat affiliates. This has resulted in a confusing story of the scale of the ransomware problem and overinflated the impact that a single RaaS program shutdown can have on the threat environment.

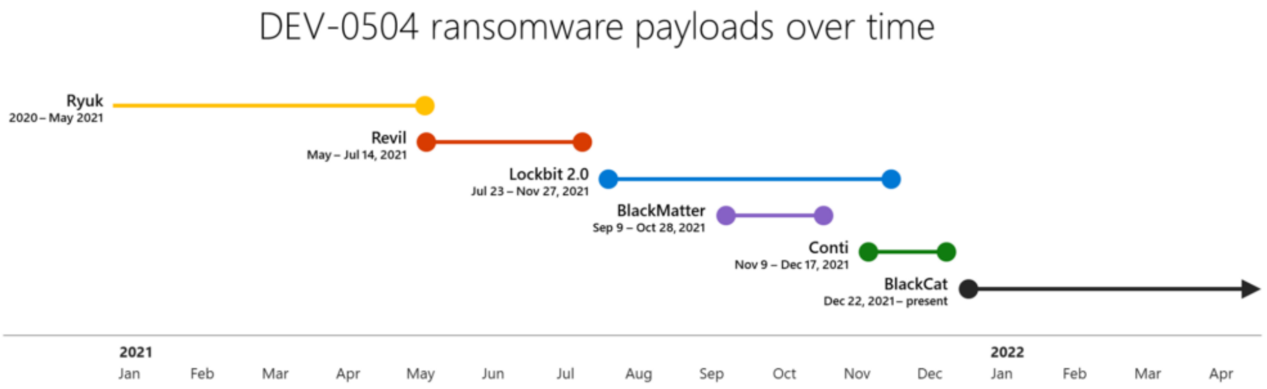


Figure 3. Ransomware payloads distributed by DEV-0504 between 2020 and April 2022
DEV-0504 shifts payloads when a RaaS program shuts down, for example the deprecation of REvil and BlackMatter, or possibly when a program with a better profit margin appears. These market dynamics aren't unique to DEV-0504 and are reflected in most RaaS affiliates. They can also manifest in even more extreme behavior where RaaS affiliates switch to older "fully owned" ransomware payloads like Phobos, which they can buy when a RaaS isn't available, or they don't want to pay the fees associated with RaaS programs.

DEV-0504 appears to rely on access brokers to enter a network, using Cobalt Strike Beacons they have possibly purchased access to. Once inside a network, they rely heavily on PsExec to move laterally and stage their payloads. Their techniques require them to have compromised elevated credentials, and they frequently disable antivirus products that aren't protected with tamper protection.

DEV-0504 was responsible for deploying BlackCat ransomware in companies in the energy sector in January 2022. Around the same time, DEV-0504 also deployed BlackCat in attacks against companies in the fashion, tobacco, IT, and manufacturing industries, among others.

DEV-0237: Prolific collaborator

Like DEV-0504, DEV-0237 is a prolific RaaS affiliate that alternates between different payloads in their operations based on what is available. DEV-0237 heavily used Ryuk and Conti payloads from Trickbot LLC/DEV-0193, then Hive payloads more recently. Many publicly documented Ryuk and Conti incidents and tradecraft can be traced back to DEV-0237.

After the activity group switched to Hive as a payload, a large uptick in Hive incidents was observed. Their switch to the BlackCat RaaS in March 2022 is suspected to be due to public discourse around Hive decryption methodologies; that is, DEV-0237 may have switched to

BlackCat because they didn't want Hive's decryptors to interrupt their business. Overlap in payloads has occurred as DEV-0237 experiments with new RaaS programs on lower-value targets. They have been observed to experiment with some payloads only to abandon them later.

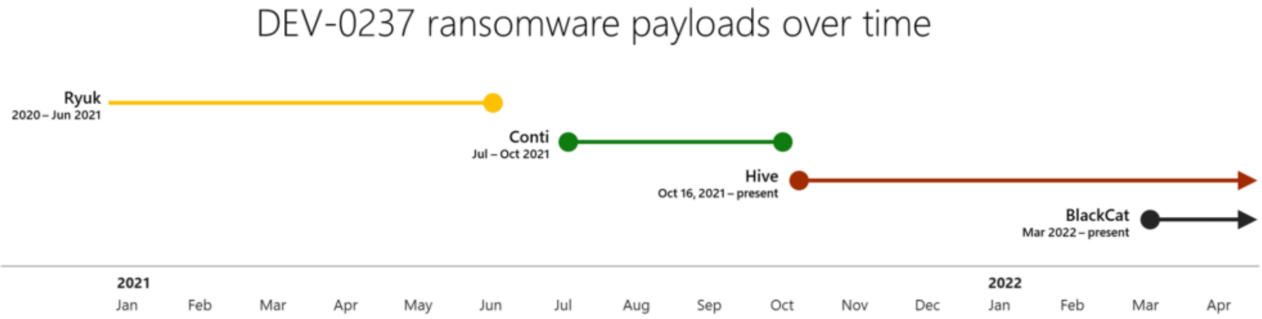


Figure 4. Ransomware payloads distributed by DEV-0237 between 2020 and April 2022

Beyond RaaS payloads, DEV-0237 uses the cybercriminal gig economy to also gain initial access to networks. DEV-0237's proliferation and success rate come in part from their willingness to leverage the network intrusion work and malware implants of other groups versus performing their own initial compromise and malware development.

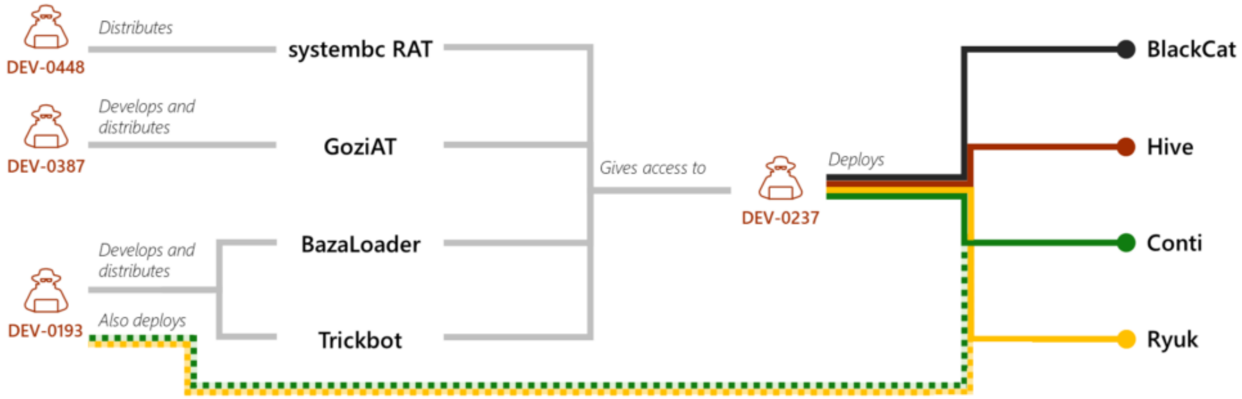


Figure 5. Examples of DEV-0237's relationships with other cybercriminal activity groups

Like all RaaS operators, DEV-0237 relies on compromised, highly privileged account credentials and security weaknesses once inside a network. DEV-0237 often leverages Cobalt Strike Beacon dropped by the malware they have purchased, as well as tools like SharpHound to conduct reconnaissance. The group often utilizes BITSadmin /transfer to stage their payloads. An often-documented trademark of Ryuk and Conti deployments is naming the ransomware payload *xxx.exe*, a tradition that DEV-0237 continues to use no matter what RaaS they are deploying, as most recently observed with BlackCat. In late March of 2022, DEV-0237 was observed to be using a new version of Hive again.

DEV-0206 and DEV-0243: An "evil" partnership

Malvertising, which refers to taking out a search engine ad to lead to a malware payload, has been used in many campaigns, but the access broker that Microsoft tracks as DEV-0206 uses this as their primary technique to gain access to and profile networks. Targets are lured by an ad purporting to be a browser update, or a software package, to download a ZIP file and double-click it. The ZIP package contains a JavaScript file (.js), which in most environments runs when double-clicked. Organizations that have changed the settings such that script files open with a text editor by default instead of a script handler are largely immune from this threat, even if a user double clicks the script.

Once successfully executed, the JavaScript framework, also referred to SocGholish, acts as a loader for other malware campaigns that use access purchased from DEV-0206, most commonly Cobalt Strike payloads. These payloads have, in numerous instances, led to custom Cobalt Strike loaders attributed to DEV-0243. DEV-0243 falls under activities tracked by the cyber intelligence industry as "EvilCorp," The custom Cobalt Strike loaders are similar to those seen in publicly documented Blister malware's inner payloads. In DEV-0243's initial partnerships with DEV-0206, the group deployed a custom ransomware payload known as WastedLocker, and then expanded to additional DEV-0243 ransomware payloads developed in-house, such as PhoenixLocker and Macaw.

Around November 2021, DEV-0243 started to deploy the LockBit 2.0 RaaS payload in their intrusions. The use of a RaaS payload by the "EvilCorp" activity group is likely an attempt by DEV-0243 to avoid attribution to their group, which could discourage payment due to their sanctioned status.

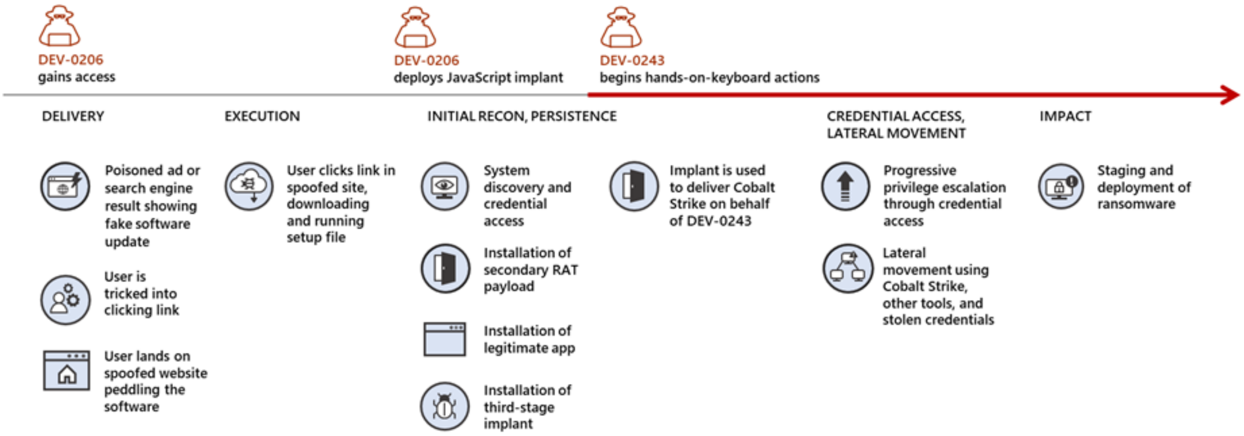

DEV-0206
gains access
DEV-0206
deploys JavaScript implant
DEV-0243
begins hands-on-keyboard actions
DELIVERY
EXECUTION
INITIAL RECON, PERSISTENCE
CREDENTIAL ACCESS, LATERAL MOVEMENT
IMPACT
Poisoned ad or search engine result showing fake software update
User is tricked into clicking link
User lands on spoofed website peddling the software
User clicks link in spoofed site, downloading and running setup file
System discovery and credential access
Installation of secondary RAT payload
Installation of legitimate app
Installation of third-stage implant
Implant is used to deliver Cobalt Strike on behalf of DEV-0243
Progressive privilege escalation through credential access
Lateral movement using Cobalt Strike, other tools, and stolen credentials
Staging and deployment of ransomware

Figure 6. The handover from DEV-0206 to DEV-0243

DEV-0401: China-based lone wolf turned LockBit 2.0 affiliate

Differing from the other RaaS developers, affiliates, and access brokers profiled here, DEV-0401 appears to be an activity group involved in all stages of their attack lifecycle, from initial access to ransomware development. Despite this, they seem to take some inspiration from successful RaaS operations with the frequent rebranding of their ransomware payloads. Unique among human-operated ransomware threat actors tracked by Microsoft, DEV-0401 is confirmed to be a China-based activity group.

DEV-0401 differs from many of the attackers who rely on purchasing access to existing malware implants or exposed RDP to enter a network. Instead, the group heavily utilizes unpatched vulnerabilities to access networks, including vulnerabilities in Exchange, Manage Engine AdSelfService Plus, Confluence, and Log4j 2. Due to the nature of the vulnerabilities they preferred, DEV-0401 gains elevated credentials at the initial access stage of their attack.

Once inside a network, DEV-0401 relies on standard techniques such as using Cobalt Strike and WMI for lateral movement, but they have some unique preferences for implementing these behaviors. Their Cobalt Strike Beacons are frequently launched via DLL search order hijacking. While they use the common Impacket tool for WMI lateral movement, they use a customized version of the *wmiexec.py* module of the tool that creates renamed output files, most likely to evade static detections. Ransomware deployment is ultimately performed from a batch file in a share and Group Policy, usually written to the NETLOGON share on a Domain Controller, which requires the attackers to have obtained highly privileged credentials like Domain Administrator to perform this action.

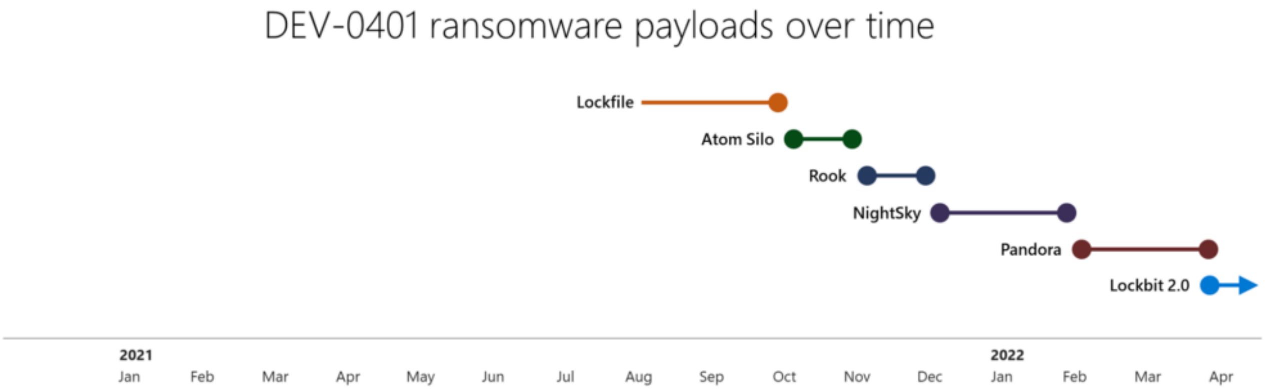


Figure 7. Ransomware payloads distributed by DEV-0401 between 2021 and April 2022
Because DEV-0401 maintains and frequently rebrands their own ransomware payloads, they can appear as different groups in payload-driven reporting and evade detections and actions against them. Their payloads are sometimes rebuilt from existing for-purchase ransomware tools like Rook, which shares code similarity with the Babuk ransomware family. In February of 2022, DEV-0401 was observed deploying the Pandora ransomware family, primarily via unpatched VMware Horizon systems vulnerable to the Log4j 2 CVE-2021-44228 vulnerability.

Like many RaaS operators, DEV-0401 maintained a leak site to post exfiltrated data and motivate victims to pay, however their frequent rebranding caused these systems to sometimes be unready for their victims, with their leak site sometimes leading to default web server landing pages when victims attempt to pay. In a notable shift—possibly related to victim payment issues—DEV-0401 started deploying LockBit 2.0 ransomware payloads in April 2022.

DEV-0537: From extortion to destruction

An example of a threat actor who has moved to a pure extortion and destruction model without deploying ransomware payloads is an activity group that Microsoft tracks as DEV-0537, also known as LAPSUS$. Microsoft has detailed DEV-0537 actions taken in early 2022 in this blog. DEV-0537 started targeting organizations mainly in Latin America but expanded to global targeting, including government entities, technology, telecom, retailers, and healthcare. Unlike more opportunistic attackers, DEV-0537 targets specific companies with an intent. Their initial access techniques include exploiting unpatched vulnerabilities in internet-facing systems, searching public code repositories for credentials, and taking advantage of weak passwords. In addition, there is evidence that DEV-0537 leverages credentials stolen by the Redline password stealer, a piece of malware available for purchase in the cybercriminal economy. The group also buys credentials from underground forums which were gathered by other password-stealing malware.

Once initial access to a network is gained, DEV-0537 takes advantage of security misconfigurations to elevate privileges and move laterally to meet their objectives of data exfiltration and extortion. While DEV-0537 doesn’t possess any unique technical capabilities, the group is especially cloud-aware. They target cloud administrator accounts to set up forwarding rules for email exfiltration and tamper with administrative settings on cloud environments. As part of their goals to force payment of ransom, DEV-0537 attempts to delete all server infrastructure and data to cause business disruption. To further facilitate the achievement of their goals, they remove legitimate admins and delete cloud resources and server infrastructure, resulting in destructive attacks.

DEV-0537 also takes advantage of cloud admin privileges to monitor email, chats, and VOIP communications to track incident response efforts to their intrusions. DEV-0537 has been observed on multiple occasions to join incident response calls, not just observing the response to inform their attack but unmuting to demand ransom and sharing their screens while they delete their victim’s data and resources.

Defending against ransomware: Moving beyond protection by detection

A durable security strategy against determined human adversaries must include the goal of mitigating classes of attacks and detecting them. Ransomware attacks generate multiple, disparate security product alerts, but they could easily get lost or not responded to in time. Alert fatigue is real, and SOCs can make their lives easier by looking at trends in their alerts or grouping alerts into incidents so they can see the bigger picture. SOCs can then mitigate alerts using hardening capabilities like attack surface reduction rules. Hardening against common threats can reduce alert volume and stop many attackers before they get access to networks.

Attackers tweak their techniques and have tools to evade and disable security products. They are also well-versed in system administration and try to blend in as much as possible. However, while attacks have continued steadily and with increased impact, the attack techniques attackers use haven’t changed much over the years. Therefore, a renewed focus on prevention is needed to curb the tide.

Ransomware attackers are motivated by easy profits, so adding to their cost via security hardening is key in disrupting the cybercriminal economy.

Building credential hygiene

More than malware, attackers need credentials to succeed in their attacks. In almost all attacks where ransomware deployment was successful, the attackers had access to a domain admin-level account or local administrator passwords that were consistent throughout the environment. Deployment then can be done through Group Policy or tools like PsExec (or clones like PAExec, CSExec, and WinExeSvc). Without the credentials to provide administrative access in a network, spreading ransomware to multiple systems is a bigger challenge for attackers. Compromised credentials are so important to these attacks that when cybercriminals sell ill-gotten access to a network, in many instances, the price includes a guaranteed administrator account to start with.

Credential theft is a common attack pattern. Many administrators know tools like Mimikatz and LaZagne, and their capabilities to steal passwords from interactive logons in the LSASS process. Detections exist for these tools accessing the LSASS process in most security products. However, the risk of credential exposure isn’t just limited to a domain administrator logging in interactively to a workstation. Because attackers have accessed and explored many networks during their attacks, they have a deep knowledge of common network configurations and use it to their advantage. One common misconfiguration they exploit is running services and scheduled tasks as highly privileged service accounts.

Too often, a legacy configuration ensures that a mission-critical application works by giving the utmost permissions possible. Many organizations struggle to fix this issue even if they know about it, because they fear they might break applications. This configuration is especially dangerous as it leaves highly privileged credentials exposed in the LSA Secrets portion of the registry, which users with administrative access can access. In organizations where the local administrator rights haven’t been removed from end users, attackers can be one hop away from domain admin just from an initial attack like a banking trojan. Building credential hygiene is developing a logical segmentation of the network, based on privileges, that can be implemented alongside network segmentation to limit lateral movement.

Here are some steps organizations can take to build credential hygiene:

- Aim to run services as Local System when administrative privileges are needed, as this allows applications to have high privileges locally but can’t be used to move laterally. Run services as Network Service when accessing other resources.
- Use tools like LUA Buglight to determine the privileges that applications really need.
- Look for events with EventID 4624 where the logon type is 2, 4, 5, or 10 *and* the account is highly privileged like a domain admin. This helps admins understand which credentials are vulnerable to theft via LSASS or LSA Secrets. Ideally, any highly privileged account like a Domain Admin shouldn’t be exposed on member servers or workstations.
- Monitor for EventID 4625 (Logon Failed events) in Windows Event Forwarding when removing accounts from privileged groups. Adding them to the local administrator group on a limited set of machines to keep an application running still reduces the scope of an attack as against running them as Domain Admin.
- Randomize Local Administrator passwords with a tool like Local Administrator Password Solution (LAPS) to prevent lateral movement using local accounts with shared passwords.
- Use a cloud-based identity security solution that leverages on-premises Active Directory signals get visibility into identity configurations and to identify and detect threats or compromised identities

Auditing credential exposure

Auditing credential exposure is critical in preventing ransomware attacks and cybercrime in general. BloodHound is a tool that was originally designed to provide network defenders with insight into the number of administrators in their environment. It can also be a powerful tool in reducing privileges tied to administrative account and understanding your credential exposure. IT security teams and SOCs can work together with the authorized use of this tool to enable the reduction of exposed credentials. Any teams deploying BloodHound should monitor it carefully for malicious use. They can also use this detection guidance to watch for malicious use.

Microsoft has observed ransomware attackers also using BloodHound in attacks. When used maliciously, BloodHound allows attackers to see the path of least resistance from the systems they have access, to highly privileged accounts like domain admin accounts and global administrator accounts in Azure.

Prioritizing deployment of Active Directory updates

Security patches for Active Directory should be applied as soon as possible after they are released. Microsoft has witnessed ransomware attackers adopting authentication vulnerabilities within one hour of being made public and as soon as those vulnerabilities are included in tools like Mimikatz. Ransomware activity groups also rapidly adopt vulnerabilities related to authentication, such as ZeroLogon and PetitPotam, especially when they are included in toolkits like Mimikatz. When unpatched, these vulnerabilities could allow attackers to rapidly escalate from an entrance vector like email to Domain Admin level privileges.

Cloud hardening

As attackers move towards cloud resources, it’s important to secure cloud resources and identities as well as on-premises accounts. Here are ways organizations can harden cloud environments:

Cloud identity hardening

- Implement the Azure Security Benchmark and general best practices for securing identity infrastructure, including:

    Prevent on-premises service accounts from having direct rights to the cloud resources to prevent lateral movement to the cloud.

    Ensure that “break glass” account passwords are stored offline and configure honey-token activity for account usage.

    Implement Conditional Access policies enforcing Microsoft’s Zero Trust principles.
- Enable risk-based user sign-in protection and automate threat response to block high-risk sign-ins from all locations and enable MFA for medium-risk ones.
- Ensure that VPN access is protected via modern authentication methods.

Multifactor authentication (MFA)

- Enforce MFA on all accounts, remove users excluded from MFA, and strictly require MFA from all devices, in all locations, at all times.
- Enable passwordless authentication methods (for example, Windows Hello, FIDO keys, or Microsoft Authenticator) for accounts that support passwordless. For accounts that still require passwords, use authenticator apps like Microsoft Authenticator for MFA. Refer to this article for the different authentication methods and features.
- Identify and secure workload identities to secure accounts where traditional MFA enforcement does not apply.
- Ensure that users are properly educated on not accepting unexpected two-factor authentication (2FA).
- For MFA that uses authenticator apps, ensure that the app requires a code to be typed in where possible, as many intrusions where MFA was enabled (including those by DEV-0537) still succeeded due to users clicking “Yes” on the prompt on their phones even when they were not at their computers. Refer to this article for an example.
- Disable legacy authentication.

Cloud admins

- Ensure cloud admins/tenant admins are treated with the same level of security and credential hygiene as Domain Admins.
- Address gaps in authentication coverage.

Addressing security blind spots

In almost every observed ransomware incident, at least one system involved in the attack had a misconfigured security product that allowed the attacker to disable protections or evade detection. In many instances, the initial access for access brokers is a legacy system that isn't protected by antivirus or EDR solutions. It's important to understand that the lack security controls on these systems that have access to highly privileged credentials act as blind spots that allow attackers to perform the entire ransomware and exfiltration attack chain from a single system without being detected. In some instances, this is specifically advertised as a feature that access brokers sell.

Organizations should review and verify that security tools are running in their most secure configuration and perform regular network scans to ensure appropriate security products are monitoring and protecting all systems, including servers. If this isn't possible, make sure that your legacy systems are either physically isolated through a firewall or logically isolated by ensuring they have no credential overlap with other systems.

For Microsoft 365 Defender customers, the following checklist eliminates security blind spots:

- Turn on cloud-delivered protection in Microsoft Defender Antivirus to cover rapidly evolving attacker tools and techniques, block new and unknown malware variants, and enhance attack surface reduction rules and tamper protection.
- Turn on tamper protection features to prevent attackers from stopping security services.
- Run EDR in block mode so that Microsoft Defender for Endpoint can block malicious artifacts, even when a non-Microsoft antivirus doesn't detect the threat or when Microsoft Defender Antivirus is running in passive mode. EDR in block mode also blocks indicators identified proactively by Microsoft Threat Intelligence teams.
- Enable network protection to prevent applications or users from accessing malicious domains and other malicious content on the internet.
- Enable investigation and remediation in full automated mode to allow Microsoft Defender for Endpoint to take immediate action on alerts to resolve breaches.
- Use device discovery to increase visibility into the network by finding unmanaged devices and onboarding them to Microsoft Defender for Endpoint.
- Protect user identities and credentials using Microsoft Defender for Identity, a cloud-based security solution that leverages on-premises Active Directory signals to monitor and analyze user behavior to identify suspicious user activities, configuration issues, and active attacks.

Reducing the attack surface

Microsoft 365 Defender customers can turn on attack surface reduction rules to prevent common attack techniques used in ransomware attacks. These rules, which can be configured by all Microsoft Defender Antivirus customers and not just those using the EDR solution, offer significant hardening against attacks. In observed attacks from several ransomware-associated activity groups, Microsoft customers who had the following rules enabled were able to mitigate the attack in the initial stages and prevented hands-on-keyboard activity:

In addition, Microsoft has changed the default behavior of Office applications to block macros in files from the internet, further reduce the attack surface for many human-operated ransomware attacks and other threats.

Hardening internet-facing assets and understanding your perimeter

Organizations must identify and secure perimeter systems that attackers might use to access the network. Public scanning interfaces, such as RiskIQ, can be used to augment data. Some systems that should be considered of interest to attackers and therefore need to be hardened include:

- Secure Remote Desktop Protocol (RDP) or Windows Virtual Desktop endpoints with MFA to harden against password spray or brute force attacks.
- Block Remote IT management tools such as Teamviewer, Splashtop, Remote Manipulator System, Anydesk, Atera Remote Management, and ngrok.io via network blocking such as perimeter firewall rules if not in use in your environment. If these systems are used in your environment, enforce security settings where possible to implement MFA.

Ransomware attackers and access brokers also use unpatched vulnerabilities, whether already disclosed or zero-day, especially in the initial access stage. Even older vulnerabilities were implicated in ransomware incidents in 2022 because some systems remained unpatched, partially patched, or because access brokers had established persistence on a previously compromised systems despite it later being patched.

Some observed vulnerabilities used in campaigns between 2020 and 2022 that defenders can check for and mitigate include:

- Citrix ADC systems affected by CVE-2019-19781
- Pulse Secure VPN systems affected by CVE-2019-11510, CVE-2020-8260, CVE-2020-8243, CVE-2021-22893, CVE-2021-22894, CVE-2021-22899, and CVE-2021-22900
- SonicWall SSLVPN affected by CVE-2021-20016
- Microsoft SharePoint servers affected by CVE-2019-0604
- Unpatched Microsoft Exchange servers
- Zoho ManageEngine systems affected by CVE-2020-10189
- FortiGate VPN servers affected by CVE-2018-13379

- Apache log4j CVE-2021-44228

Ransomware attackers also rapidly adopt new vulnerabilities. To further reduce organizational exposure, Microsoft Defender for Endpoint customers can use the threat and vulnerability management capability to discover, prioritize, and remediate vulnerabilities and misconfigurations.

Microsoft 365 Defender: Deep cross-domain visibility and unified investigation capabilities to defend against ransomware attacks

The multi-faceted threat of ransomware requires a comprehensive approach to security. The steps we outlined above defend against common attack patterns and will go a long way in preventing ransomware attacks. Microsoft 365 Defender is designed to make it easy for organizations to apply many of these security controls.

Microsoft 365 Defender's industry-leading visibility and detection capabilities, demonstrated in the recent MITRE Engenuity ATT&CK® Evaluations, automatically stop most common threats and attacker techniques. To equip organizations with the tools to combat human-operated ransomware, which by nature takes a unique path for every organization, Microsoft 365 Defender provides rich investigation features that enable defenders to seamlessly inspect and remediate malicious behavior across domains.

Learn how you can stop attacks through automated, cross-domain security and built-in AI with Microsoft Defender 365.

In line with the recently announced expansion into a new service category called **Microsoft Security Experts**, we're introducing the availability of Microsoft Defender Experts for Hunting for public preview. Defender Experts for Hunting is for customers who have a robust security operations center but want Microsoft to help them proactively hunt for threats across Microsoft Defender data, including endpoints, Office 365, cloud applications, and identity.

Join our research team at the **Microsoft Security Summit** digital event on May 12 to learn what developments Microsoft is seeing in the threat landscape, as well as how we can help your business mitigate these types of attacks. Ask your most pressing questions during the live chat Q&A. Register today.