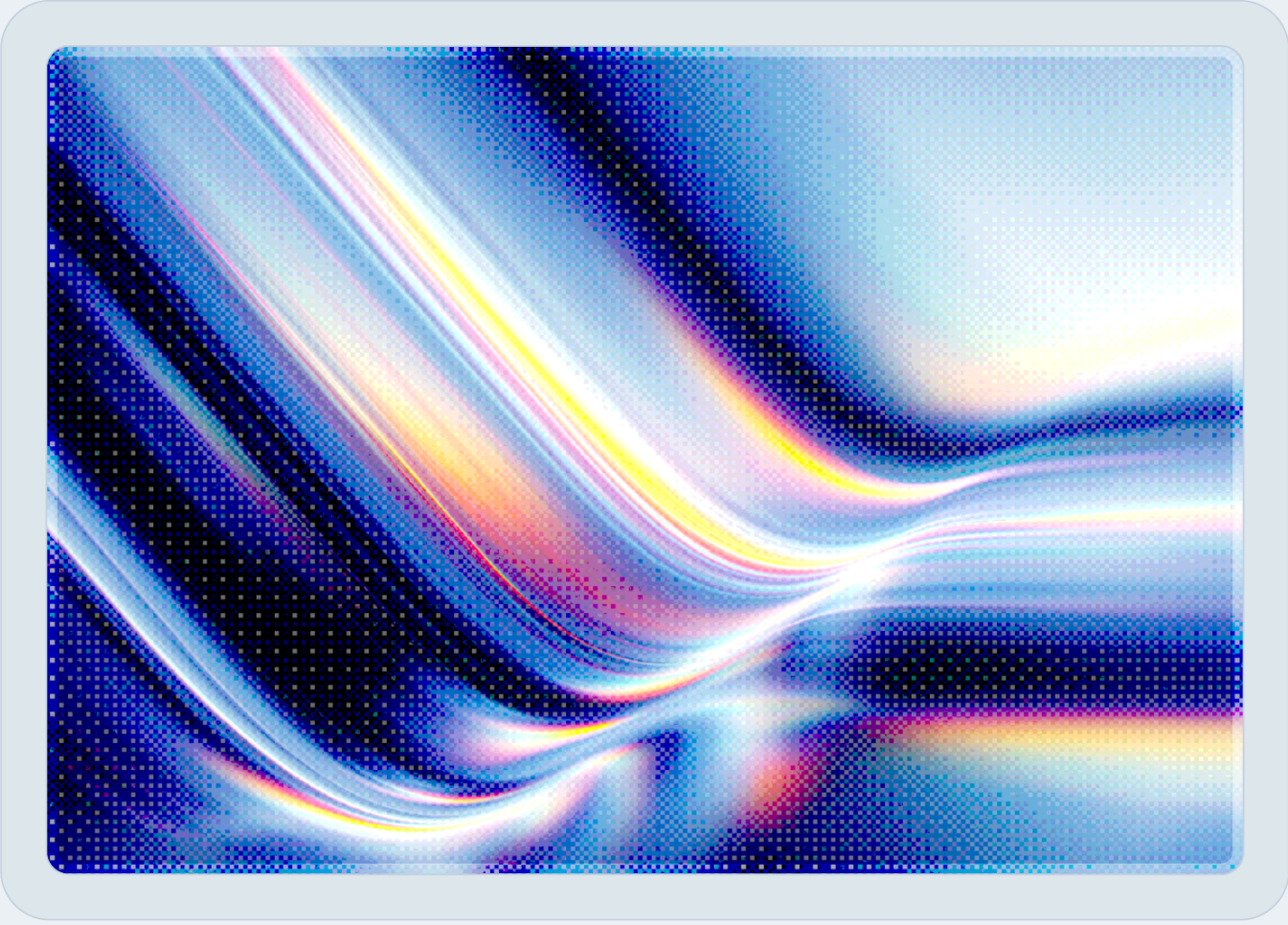
CSAI Foundation | Cloud Security Alliance

GentleKiller: Inside the Gentlemen RaaS EDR-Killer Suite

2026-06-22

AI-assisted Rapid Research



CSAI

CSA cloud security alliance®

Key Takeaways

- The Gentlemen ranked as the second most active ransomware-as-a-service operation in the first five months of 2026, with approximately 332 publicly listed victims and internal telemetry suggesting a true victim pool exceeding 1,570 corporate organizations [1][3].
- GentleKiller, the group's in-house defense-evasion framework, deploys at least eight variants that exploit vulnerable or malicious kernel-mode drivers (BYOVD) to terminate more than 400 security processes across 48 vendors [1][2].
- Each GentleKiller variant impersonates a legitimate security product through fabricated version metadata, copied digital signatures, and commercial code-protection packing, significantly undermining the reliability of signature-based detection [1][4].
- The Gentlemen supplement GentleKiller with at least three externally sourced EDR-killing tools—HexKiller, ThrottleBlood, and HavocKiller—which creates defense-evasion redundancy and, since these tools appear across multiple threat-actor clusters, may complicate attribution [2][4].
- Defenders should prioritize Hypervisor-Protected Code Integrity (HVCI) and Windows Defender Application Control (WDAC) policies over sole reliance on EDR process-monitoring, since BYOVD attacks operate at a layer below which user-space agents cannot observe [6] [7].

Background

The Gentlemen ransomware-as-a-service operation emerged in September 2025 and ranked as the second most active RaaS operation in the first five months of 2026 [3]. Led by a Russian-speaking administrator operating under the aliases zeta88 and hastalamuerte, the group operates a documented core of approximately nine named operators alongside a broader affiliate network [3]. The Gentlemen distinguished itself from competing RaaS programs through an unusually affiliate-friendly revenue split: affiliates retain 90% of ransom proceeds, above the typical affiliate split across competing RaaS programs, drawing experienced operators seeking higher margins [3][5].

By May 2026, the group had publicly claimed approximately 332 victims on its data-leak site, ranking it as the second most productive RaaS operation for that period among groups that publicly list victims [3] [5]. That figure likely substantially understates actual reach. Check Point Research gained access to a live command-and-control server linked to a Gentlemen affiliate, which revealed a botnet of more than 1,570 likely corporate victims—a gap that indicates many victims either paid quietly or remained in the pre-publication negotiation window when the data was captured [3]. On May 4, 2026, the group's internal backend database, called Rocket, was leaked, exposing 9 operator accounts, at least 8 distinct affiliate identifiers, and approximately 16.22 gigabytes of operational communications, infrastructure details, and victim records [3][5]. That breach gave researchers a rare and detailed picture of the group's tool development and affiliate provisioning practices, confirming what incident responders had previously inferred from forensic evidence.

Gentlemen intrusions consistently begin at internet-exposed edge infrastructure. The group targets Fortinet FortiGate and Cisco devices through credential brute-forcing, exploitation of known vulnerabilities, and access purchased from initial-access brokers [3]. The group actively tracks newly disclosed vulnerabilities—including CVE-2024-55591, CVE-2025-32433, and CVE-2025-33073—and demonstrates a consistent pattern of rapid weaponization [3][4]. Once inside, affiliates conduct Active Directory reconnaissance and privilege escalation before deploying ransomware. The Gentlemen practice double extortion, encrypting data and threatening public release to maximize pressure on victims. A documented negotiation from the leaked data resulted in payment of approximately $190,000 USD following an initial $250,000 demand [3]. The leaked data also revealed opportunistic data reuse: material exfiltrated from a UK software consultancy was subsequently weaponized in an attack against a Turkish company, illustrating a secondary threat multiplier inherent in double-extortion operations [3].

Victim concentrations span Southeast Asia, South America, and Western Europe [4][9].

Security Analysis

GentleKiller Architecture and BYOVD Exploitation

The operational core of the Gentlemen's defense-evasion capability is GentleKiller, an in-house EDR-killing framework that the leadership maintains and distributes to verified affiliates. The framework was discovered staged in a directory named GentlemenCollection during incident investigation, and the internal Rocket leak confirmed that zeta88 personally builds and provisions it as part of the RaaS offering [1][2]. ESET researchers hypothesized in February 2026 that GentleKiller was an internal development; that assessment was subsequently corroborated by Group-IB and Check Point, and then definitively confirmed by the leaked chats [1].

GentleKiller is built entirely around the Bring Your Own Vulnerable Driver (BYOVD) technique. BYOVD involves loading a legitimately signed but exploitable kernel-mode driver onto the target system and then using that driver's vulnerabilities to execute code at ring-0 privilege [1][6]. Because security software typically runs in user space, kernel-level execution effectively blinds endpoint detection: once a vulnerable driver grants ring-0 access, GentleKiller can terminate any process, including EDR agents, without triggering the user-mode hooks those agents rely upon. The drivers exploited carry valid vendor signatures, which convey only that the driver was once signed by a trusted party—not that it is safe to load in the current operational context.

ESET researchers identified at least eight distinct GentleKiller variants, each built around a different vulnerable or malicious driver. The following table summarizes the known variants and their associated drivers:

Variant Name	Driver File	Exploited Component
Kaspersky	eb.sys	Custom rootkit
FACEIT Anti-Cheat	nseckrnl.sys	NSecsoft NSecKrnl
Valorant	GameDriverX64.sys	Anti-cheat component
Javelin (Safetica)	stpm_old.sys / stpm_new.sys	Safetica TPM driver
Zemana WatchDog	dmx.sys	Zemana AntiMalware
Network Blocker	360netmon_wfp.sys	Qihoo 360
Cleaner	IMFForceDelete	IObit
G11	PoisonX	Rootkit

Each variant impersonates a different legitimate security product through fabricated version metadata, copied icons, and invalid digital signatures transplanted from legitimate executables [1][2]. The binaries are protected with commercial code-protection tools—either Enigma or Themida—which add obfuscation resistance against static analysis [2]. Despite this surface variation, ESET identified numerous structural and behavioral commonalities across all eight variants, strongly indicating derivation from a shared internal development template [1][4]. The shared architecture enables the group to rapidly generate new variants by substituting the underlying driver while preserving the evasion layer, process-termination logic, and target process list. The process-termination logic itself runs on a timer, scanning for and killing targeted processes approximately every two seconds [6].

GentleKiller's target set spans more than 400 process names linked to 48 security products, including CrowdStrike Falcon, Microsoft Defender, SentinelOne, Sophos, Palo Alto Cortex XDR, Bitdefender, Trend Micro, ESET, McAfee/Trellix, and Kaspersky [1][2][4]. The breadth of coverage suggests a deliberate design goal [1][3]: rather than targeting a narrow set of common defenders, GentleKiller appears engineered to function across the wide range of enterprise security stacks an affiliate might encounter.

Supplementary EDR-Killing Tools

Beyond GentleKiller, the Gentlemen provision affiliates with at least three externally sourced EDR-killing utilities. HexKiller exploits the BdApi driver associated with Baidu Antivirus, while ThrottleBlood abuses the ThrottleStop.sys driver from the TechPowerUp performance utility. HavocKiller, first observed in March 2026, abuses a Huawei Audio driver (havoc.sys) [2][4]. Prior to their incorporation into the Gentlemen toolkit, HexKiller had been associated with the Warlock ransomware gang, and ThrottleBlood had appeared in MedusaLocker and DragonForce operations [2]. The Gentlemen operators standardized these imported tools by applying the same impersonation techniques and commercial packing conventions used across GentleKiller variants, presenting them as a cohesive defensive capability to affiliates [4].

This multi-tool approach serves two distinct operational purposes. The immediate operational benefit is redundancy: if one utility is blocked or flagged by a particular target's defenses, alternatives remain available within the same session. The secondary effect is that some of these tools appear across multiple distinct threat-actor clusters, which complicates incident-attribution decisions and may delay response escalation [2].

Credential Harvesting and Full Intrusion Chain

GentleKiller operates as one stage in a broader intrusion chain. Gentlemen affiliates deploy OxideHarvest, a Rust-based credential stealer, for post-exploitation credential collection [2]. The broader toolkit includes ZeroPulse for reconnaissance, Velociraptor (repurposed from its legitimate DFIR context), NetExec for remote execution, RelayKing for NTLM relay attacks, TaskHound for scheduled-task manipulation, PrivHound for privilege escalation, and CertiHound for Active Directory Certificate Services abuse [3]. Lateral movement frequently employs Cloudflare Zero Trust tunnels and custom VPN deployments, blending exfiltration and command-and-control traffic with legitimate cloud service activity [3].

The sequence—initial access through edge devices, credential harvesting, domain-wide lateral movement, EDR elimination via GentleKiller, data exfiltration, and finally ransomware deployment—illustrates a full-lifecycle intrusion methodology, from initial access through post-exfiltration encryption. The group's systematic targeting of NAS devices and backup systems prior to encryption indicates a deliberate effort to remove recovery options [3].

Rapid PoC Operationalization

A notable operational characteristic of the Gentlemen is their speed in weaponizing newly disclosed BYOVD proofs-of-concept. ESET documented that the group incorporates new BYOVD PoCs within days of public disclosure, a pace that has repeatedly outstripped the update cycles for Microsoft's Vulnerable Driver Blocklist [1][4][8]. That blocklist is updated infrequently—typically one to two times per year alongside major Windows releases—and does not cover the full driver attack surface, creating a window between disclosure and blocklist inclusion that the Gentlemen have systematically leveraged [7].

Recommendations

Immediate Actions

Security teams should immediately audit Windows endpoints for unauthorized kernel driver loading. Behavioral alerts that correlate driver installation events with process-termination activity targeting security software provide a high-confidence near-term detection signal for GentleKiller [6]. Monitoring for the GentlemenCollection staging directory name and the specific driver filenames identified in ESET's analysis provides additional host-based indicators; the full IOC set from that research should be ingested into SIEM and threat-intelligence platforms without delay [1]. Incident responders should also verify that all enrolled EDR agents are actively reporting telemetry—silence from an enrolled endpoint that remains network-active is a strong indicator that an EDR-killing tool has already been deployed [4][6].

Edge infrastructure deserves immediate hardening review. The Gentlemen consistently gain initial access through exposed FortiGate and Cisco management interfaces. Management-plane interfaces should not be publicly reachable, all known vulnerabilities on edge devices should be patched, and authentication anomalies on VPN concentrators and firewalls should trigger real-time alerting [3][6].

Short-Term Mitigations

Enabling Hypervisor-Protected Code Integrity (HVCI) and the full Virtualization-Based Security (VBS) stack on managed endpoints is widely considered the strongest architectural defense against BYOVD attacks [7][8]. HVCI enforces that only drivers verified by a trusted code-integrity policy can load into the kernel, blocking GentleKiller and similar tools from achieving ring-0 execution even when an attacker already holds local administrator privileges. Where HVCI is not immediately deployable—often due to legacy hardware compatibility constraints—Windows Defender Application Control policies should be configured to restrict which drivers are permitted to load, with the specific drivers observed in GentleKiller analysis explicitly blocked [6][7].

Microsoft's Vulnerable Driver Blocklist provides a baseline defense that should be enabled but should not be treated as comprehensive, given its infrequent update cadence and limited coverage of the overall driver ecosystem [7][8]. Organizations managing their own WDAC policies can supplement the blocklist with the specific drivers identified in GentleKiller analysis and update those policies independently of Windows release cycles, closing the window between public disclosure and blocklist inclusion.

Network-layer monitoring provides a critical detection layer that BYOVD attacks cannot easily suppress. Because GentleKiller operates by killing endpoint agent processes, an attacker who succeeds eliminates host-based telemetry while typically leaving network telemetry intact. SIEM rules that flag endpoints going silent in EDR consoles while remaining active on the network can surface compromises even when local agents have been blinded [6][8]. Combining EDR telemetry health monitoring with network flow analysis significantly raises the cost for attackers who must then suppress both simultaneously.

Strategic Considerations

The Gentlemen's rapid operationalization of BYOVD PoCs reveals a structural tension in how kernel driver trust is managed across the Windows ecosystem. Organizations that rely primarily on user-space EDR controls face a fundamental constraint: any tool operating at the kernel layer can circumvent user-space detection by design. A defense architecture that treats kernel-level driver integrity as a first-class security control—enforcing HVCI and WDAC from initial provisioning rather than retrofitting them onto existing deployments—materially raises the cost of BYOVD exploitation. This architecture represents a resilience investment that benefits defense against GentleKiller specifically and against the broader class of kernel-level evasion tools generally.

The Gentlemen's double-extortion model and their demonstrated willingness to reuse exfiltrated data against third-party organizations broadens the blast radius of any single intrusion beyond the immediate victim. Organizations in regions where the group has concentrated activity—Southeast Asia, South America, and Western Europe—should treat data exfiltration as an expected intrusion component and

scope incident response plans accordingly [4][9]. Backup integrity and tested restoration procedures address the encryption component; the exfiltration component requires a parallel response track that accounts for regulatory notification obligations, downstream victim communication, and the possibility that exfiltrated data may be reused in subsequent attacks against the organization's customers or partners.

CSA Resource Alignment

GentleKiller and the Gentlemen RaaS operation are directly relevant to several active CSA frameworks and guidance documents.

CSA's Zero Trust Guidance addresses the core architectural premise that BYOVD exploitation subverts: implicit trust in signed kernel drivers. A zero trust posture that extends continuous verification to the kernel layer—through HVCI and WDAC enforcement—treats driver loading as a privileged action requiring policy authorization rather than signature verification alone. Organizations applying CSA Zero Trust principles to endpoint architecture should extend that model explicitly to cover kernel driver loading policy and not assume that OS-level driver signing provides sufficient assurance.

The Cloud Controls Matrix (CCM) is relevant across several control domains. The Threat and Vulnerability Management (TVM) domain applies to the group's rapid exploitation of newly disclosed BYOVD vulnerabilities and to the need to correlate public vulnerability disclosures with driver inventories on managed endpoints. The Security Incident Management, E-Discovery, and Cloud Forensics (SEF) domain covers the double-extortion and data-reuse dimensions, where exfiltrated data creates downstream legal and notification obligations that extend beyond the immediate intrusion. The Identity and Access Management (IAM) domain is implicated by the group's extensive credential-harvesting toolkit—OxideHarvest, CertiHound, and PrivHound—which collectively target privileged credentials and Active Directory Certificate Services to enable persistent access and lateral movement.

CSA's AI Safety Initiative research on AI-augmented threat actors is increasingly relevant to groups like the Gentlemen that demonstrate unusually rapid PoC operationalization. While public reporting does not currently attribute the group's pace directly to AI tooling, the pattern of incorporating new BYOVD PoCs within days of release is consistent with AI-assisted triage of vulnerability disclosures and automated tooling adaptation. Organizations applying CSA's MAESTRO framework to their defensive AI deployments should account for adversarial AI use in accelerating traditional exploit-development cycles as a near-term threat scenario, not merely a hypothetical one.

The CSA STAR program's transparency and continuous-monitoring tier is relevant to supply chain risk in this context. The Gentlemen's use of externally sourced tools—HexKiller, ThrottleBlood, and HavocKiller—previously associated with Warlock, MedusaLocker, and DragonForce, illustrates how EDR-killing capabilities diffuse across the ransomware ecosystem. Third-party security assessments informed by STAR disclosures should account for the possibility that a shared defense-evasion layer may appear in intrusions attributed to multiple distinct groups, since BYOVD utilities are increasingly treated as commodity components traded and shared across criminal operators [2][8].

References

[1] ESET Research. "Killing me gently: Inside Gentlemen's EDR killer framework." WeLiveSecurity, June 2026.

[2] BleepingComputer. "Gentlemen ransomware uses multiple EDR killers to disable defenses." BleepingComputer, June 2026.

[3] Check Point Research. "Thus Spoke...The Gentlemen." Check Point Research, June 2026.

[4] Help Net Security. "GentleKiller targets more than 400 security processes across 48 products." Help Net Security, June 18, 2026.

[5] The Hacker News. "The Gentlemen RaaS Uses GentleKiller EDR Framework Targeting 400 Security Processes." The Hacker News, June 2026.

[6] CybersecurityNews. "GentleKiller Ransomware Abuses Vulnerable Drivers to Disable 400+ EDR Security Processes." Cybersecurity News, June 2026.

[7] Halcyon AI. "Understanding BYOVD Attacks and Mitigation Strategies." Halcyon, May 7, 2025.

[8] Threat Intel Report. "BYOVD in 2026: the signed-driver loophole powering EDR bypass at scale." Threat Intel Report, February 21, 2026.

[9] Halcyon AI. "The Gentlemen Ransomware Group Is Scaling Faster Than Any Other Group on Record." Halcyon, 2026.

