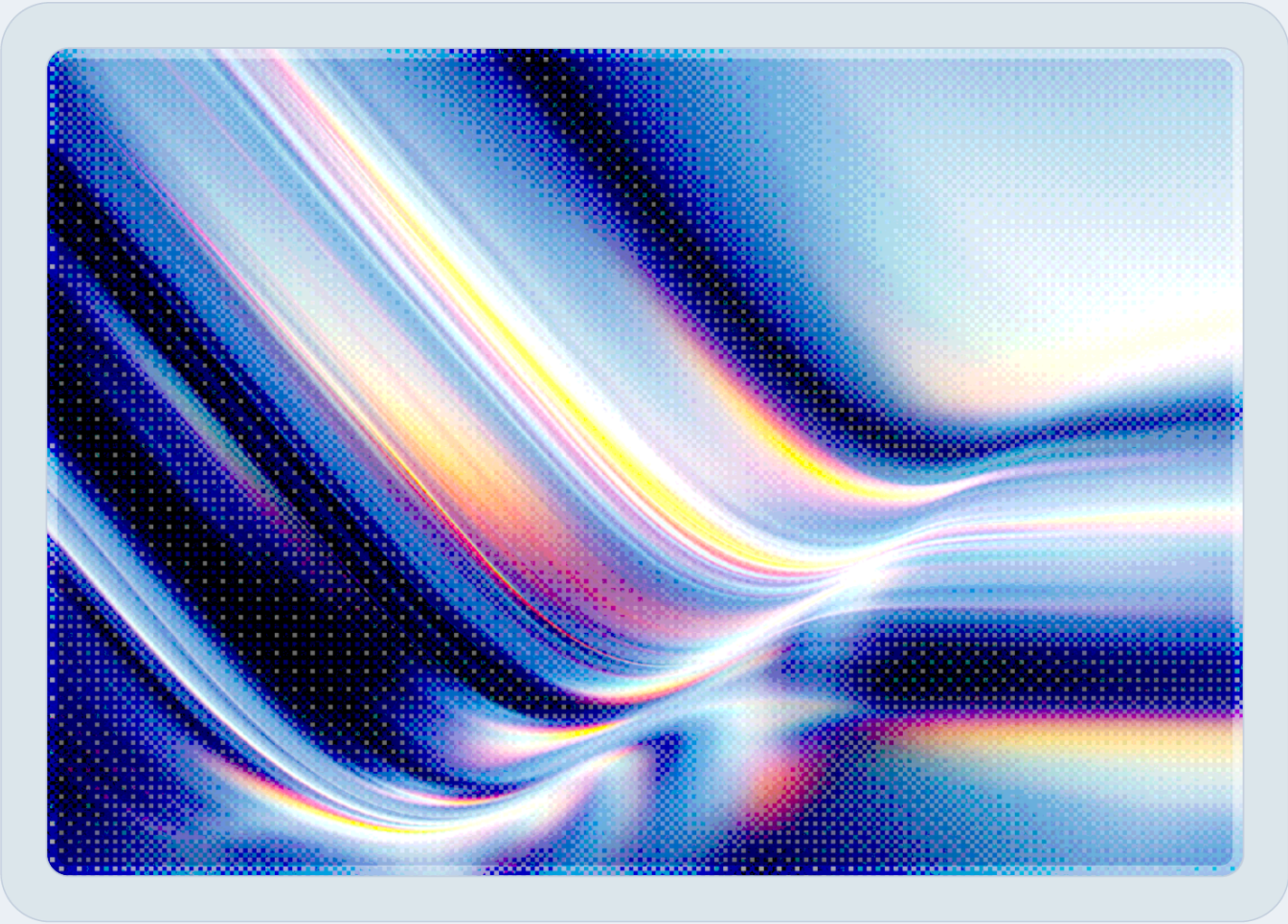
CSAI Foundation | Cloud Security Alliance

Sandworm's Fake Job Interviews Deliver Trojanized WireGuard VPN

2026-08-13

AI-assisted Rapid Research



CSAI

CSA cloud security alliance®

Key Takeaways

CERT-UA disclosed on August 9, 2026, that UAC-0145 — a subcluster of Sandworm (also tracked as UAC-0002, APT44, and Seashell Blizzard, and assessed as affiliated with Russia's GRU military intelligence directorate) — has run a recruitment-fraud campaign against Ukrainian system administrators and IT professionals since at least May 2026 [1][2]. The operation impersonates recruiters from real IT and consulting firms, including ATLAS Business Group and Sopra Steria's Bulgarian office, and moves victims through a multi-stage funnel of job-site chat, Telegram screening, and an English-language Zoom interview before delivering a "technical assignment" that requires installing a VPN client [1][2]. That client, distributed from SourceForge under names such as "SopraVPN," is a modified build of the legitimate open-source WireGuard software that has been altered to accept a non-standard configuration field, decrypt an embedded PowerShell payload, and execute attacker commands on the victim's machine without their knowledge [1][2][3].

The trojanization technique is notable for where it hides malicious logic: not in a suspicious binary that endpoint detection tools might flag, but inside the WireGuard configuration file itself, which is customarily treated as inert connection data rather than executable content. The modified client adds a "SymmetricKey" field containing Base64-encoded AES-256-GCM ciphertext and derives the decryption key from the configuration's own "PrivateKey" value [1][2]. Independent technical write-ups of the CERT-UA findings further describe the process as obfuscated with a custom Base64 alphabet generated through a Fisher-Yates shuffle seeded by a CRC32 checksum [2][3]. Once decrypted, the embedded PowerShell runs on Windows via a scheduled task that fetches a secondary payload, while a Linux build of the same trojanized client uses cURL over the established VPN tunnel to retrieve and execute an attacker-supplied binary [1][2][3].

This campaign extends a pattern documented industry-wide: nation-state and financially motivated actors increasingly use fabricated recruitment processes — recognizable from North Korean IT-worker fraud and Lazarus Group's "Contagious Interview" campaigns, and now a Sandworm subcluster — because job interviews appear to create a socially sanctioned reason for a target to run unfamiliar software with elevated trust and reduced scrutiny. Organizations that employ or contract Ukrainian IT staff, and any enterprise whose administrators actively interview externally, should treat unsolicited "technical assignment" VPN installs as a credible initial-access vector and apply the mitigations below.

Background

Sandworm is widely assessed by Western governments and security researchers, including Google's Mandiant, as one of the most disruptive Russian state-sponsored threat actors of the past decade, with a track record that includes the 2015 and 2016 attacks on the Ukrainian power grid, the 2017 NotPetya wiper — which a White House assessment attributed to Russia's military and estimated caused more than $10 billion in global damage — and repeated destructive operations against Ukrainian government and critical infrastructure networks following Russia's 2022 invasion [9][10]. UAC-0145 is the identifier CERT-UA — Ukraine's Computer Emergency Response Team — uses to track a specific activity cluster it assesses as a Sandworm subcluster operating under the broader UAC-0002 designation; the group overlaps in public reporting with Microsoft's "Seashell Blizzard" and Mandiant's "APT44" naming conventions [1][2][9]. CERT-UA's disclosure describes an operation that began no later than May 2026 and, in a departure from Sandworm's historically destructive playbook, opens with a patient, weeks-long social engineering process rather than a technical exploit.

The campaign begins with reconnaissance against job-search platforms, where the operators review the résumés of Ukrainian system administrators and IT specialists before initiating contact under the guise of a hiring manager at a legitimate firm [1][2]. Reported impersonations include ATLAS Business Group and the Bulgarian office of Sopra Steria, a real multinational IT consulting company; in at least one instance, the group registered a look-alike domain resembling Sopra Steria's Bulgarian web presence to lend correspondence added legitimacy [4]. After an initial exchange through the job platform's own messaging system, the conversation is deliberately migrated to Telegram for a preliminary screening conducted by a purported HR representative, and from there to a Zoom videoconference with an English-speaking interviewer whose age and manner The Hacker News's reporting characterizes as consistent with either a real recruiter or a synthetic, AI-generated persona [1]. CERT-UA has not confirmed which is the case in every instance, and CSA treats the AI-generated-persona possibility as a plausible but unverified aggravating detail rather than a settled fact.

Following the video interview, the candidate receives an email instructing them to complete a "technical assignment" that requires connecting to a corporate VPN, with configuration files attached for either Windows or Linux [1][2]. A second Zoom session is then scheduled, ostensibly to observe the candidate working through the assignment [1]. CSA assesses that this second session likely also gives the operators a window to monitor whether access has succeeded and to sustain social pressure that discourages the target from pausing to scrutinize the software they were told to install, though the source reporting does not state this motive explicitly. In CSA's assessment, every stage of this funnel functions to normalize an action — installing an externally provided VPN client on a work or work-capable machine — that a security-aware employee would ordinarily question, by embedding it inside a process (a job interview) that carries its own implicit deadlines and incentives to comply quickly.

Security Analysis

The VPN client at the center of this campaign is compiled from WireGuard's open-source codebase, which gives it functional parity with the legitimate application and helps it evade the reflexive suspicion that a wholly unfamiliar tool might invite [1][2]. CERT-UA's technical analysis, corroborated by independent reporting, identifies the core modification as support for a "SymmetricKey" configuration option that WireGuard's standard implementation does not recognize [1][2]. That field carries Base64-encoded data structured as an AES-256-GCM nonce, ciphertext, and authentication tag; the malicious client decodes it and derives the decryption key from the same configuration file's "PrivateKey" field — a value that, in a legitimate WireGuard deployment, exists solely to authenticate the tunnel, not to unlock a second layer of hidden instructions [1][2]. Technical write-ups of the sample further describe a custom Base64 alphabet, generated via a Fisher-Yates shuffle seeded by a CRC32 checksum, layered on top of the AES decryption to frustrate straightforward static analysis of the configuration file's contents [3].

The two observed payload-delivery paths diverge by operating system but converge on the same objective: establishing a channel for arbitrary command execution once the VPN tunnel is active. On Windows, the decrypted PowerShell creates a scheduled task — reported to masquerade under a path resembling a Microsoft Windows system directory — that reaches out to attacker infrastructure to retrieve an additional payload [1][2][4]. On Linux, the client instead shells out to cURL to pull an executable over the same VPN connection, with at least one analysis identifying a drop location disguised as a system time-synchronization utility [2][3][4]. In both cases, the victim's own VPN tunnel — the very connection the "technical assignment" instructed them to establish — becomes the transport for the intrusion, meaning that any network monitoring calibrated to trust traffic inside an established VPN session would need to be re-examined for this scenario. CERT-UA's public reporting does not identify a named malware family for the secondary payload, and CSA has not independently verified specific file hashes or command-and-control infrastructure at the time of this writing; organizations should rely on CERT-UA's forthcoming or updated indicator releases rather than secondary-source IOC lists that have not been cross-confirmed.

WireGuard has become a preferred protocol for both legitimate enterprises and privacy-conscious individuals because of its compact, auditable codebase and strong default cryptography, and open-source distribution through platforms like SourceForge is a common way for smaller or unofficial builds to reach users. That same openness is what the operators exploited: a security-conscious IT professional evaluating a "SopraVPN" download would find a functioning, standards-compliant WireGuard client with no obviously malicious binary behavior at rest, because the malicious logic activates only after a legitimate-looking configuration file supplies the hidden key material. CSA assesses that this makes the choice of WireGuard as the trojanization target a meaningful signal, and that the technique is

conceptually distinct from — and potentially more durable than — trojanizing a binary outright, since the payload's activation is gated behind attacker-controlled configuration data that a static or dynamic malware scan of the installer alone would not necessarily surface.

Table 1 summarizes how this campaign's technique compares with the two other primary vectors by which VPN infrastructure has been weaponized against enterprises in 2026.

Vector	Entry Point	Mechanism	Primary Target
UAC-0145 trojanized WireGuard (this campaign)	Social engineering (fake job interview)	Malicious config field decrypts and executes embedded PowerShell/shell commands	Individual IT professionals, primarily in Ukraine
FortiBleed credential exposure [5]	Legacy password hashing and an authentication bypass (CVE-2026-24858)	Offline cracking and credential stuffing against exposed SSL VPN accounts	Internet-facing Fortinet FortiGate appliances at scale
Nation-state AI-weaponized phishing [6]	AI-generated spear phishing and deepfake vishing	Credential harvesting or fraudulent authorization via impersonation	Enterprise employees and financial authorization workflows

The comparison underscores that VPN and remote-access infrastructure is under sustained pressure from structurally different directions at once — a scaled technical exploitation of perimeter appliances on one hand, and a patient, individually targeted social engineering operation on the other — and that defenses tuned for one do not necessarily cover the other.

Recommendations

Immediate Actions

Organizations with Ukrainian-based IT staff, contractors, or job applicants — and any organization whose system administrators are actively engaged in the external job market — should issue guidance now warning against installing any VPN client, remote-access tool, or "technical assignment" software supplied by a prospective employer outside of an established, verifiable corporate onboarding process. Security teams should search endpoint and download logs for evidence of SourceForge-hosted VPN clients bearing names associated with impersonated firms, including variations on "SopraVPN," and treat any WireGuard configuration file containing a non-standard "SymmetricKey" field as a strong indicator of compromise warranting immediate isolation of the host [1][2]. Any machine that connected through a suspect client should be treated as a confirmed incident rather than a suspected one, given that the technique is specifically designed to achieve silent command execution.

Short-Term Mitigations

CERT-UA's own guidance recommends restricting access to corporate resources to managed, continuously monitored devices protected by endpoint detection and response, including in bring-your-own-device scenarios where employees might reasonably be asked to use personal equipment for interviews or technical assessments [2]. Recruitment and hiring teams should establish a verified channel — a callback to a published company phone number, or confirmation through an existing corporate identity system — before any candidate is asked to install software, and should never require installation of a VPN client, remote-desktop tool, or code-execution environment as part of a screening or take-home technical assessment; where a live coding or systems exercise is genuinely necessary, it should run in an employer-provisioned, sandboxed, or browser-based environment that the candidate does not need to install locally. Security awareness training for IT and administrative staff should explicitly incorporate this fake-interview pattern alongside the North Korean IT-worker and "Contagious Interview" scenarios it resembles, since all three exploit the same underlying trust dynamic between job candidate and prospective employer.

Strategic Considerations

The durability of this technique — hiding malicious logic in configuration data rather than in an executable — argues for treating configuration files, not just binaries, as a category of content requiring inspection before deployment on any managed endpoint, a principle that extends well beyond WireGuard to any

tool whose configuration format supports encoded or scripted content. More broadly, this campaign is a reminder that the recruitment pipeline is, in CSA's assessment, an under-defended piece of enterprise attack surface: HR, recruiting, and hiring-manager workflows commonly sit outside the scope of security operations, yet they now regularly involve external parties sending files, links, and software to prospective employees before any employment relationship or corporate device-management policy applies. Organizations should work with HR and talent-acquisition functions to define minimum security requirements for the technical interview process itself, and should factor recruitment-stage social engineering into enterprise threat models alongside more conventional email- and collaboration-platform-based vectors.

CSA Resource Alignment

This campaign connects directly to CSA AI Safety Initiative and Zero Trust publications that examine adjacent dimensions of the same threat convergence: AI-augmented social engineering, VPN and perimeter trust exploitation, identity spoofing, and nation-state adoption of high-fidelity phishing tradecraft.

AI Superpersuasion: Influence Operations and Enterprise Security Risk [7] (June 28, 2026) is the most directly applicable prior CSA work, given the UAC-0145 campaign's use of a sustained, multi-session video interview process and the unverified but plausible use of an AI-generated interviewer persona. That research note documents how frontier AI models now outperform expert human persuaders and enable attackers to sustain believable, personalized deception across extended interactions — precisely the capability that would let a small threat actor cluster run a convincing, weeks-long fake hiring process against dozens of targets in parallel. Its recommendation to shift from single-channel identity trust and content-based detection toward procedural, multi-party verification and Zero Trust architecture applies directly to the hiring and technical-assessment workflows this campaign exploits.

FortiBleed: Default Credential Exploitation and Mass Fortinet Compromise [5] (June 20, 2026) provides the complementary technical-exploitation half of the current VPN threat picture. Where FortiBleed documents credential harvesting, default-account abuse, and authentication-bypass exploitation against legitimate VPN appliances at internet scale, the UAC-0145 campaign shows a threat actor achieving an equivalent outcome — arbitrary command execution through a VPN connection — by trojanizing the client software itself rather than attacking the server side. Read together, the two research notes make the case that CSA's Software Defined Perimeter and Zero Trust Network Access guidance, including its **Executive View on how Zero Trust protects organizations by securely**

connecting users to resources from anywhere [11], addresses VPN risk from both directions: it reduces the blast radius of a compromised legitimate VPN appliance and removes the incentive for a target to install an unvetted third-party VPN client in the first place.

The impersonation at the heart of this campaign — fabricated ATLAS Business Group and Sopra Steria personas, a look-alike Sopra Steria Bulgaria domain, and a possibly synthetic interviewer — is itself the identity-spoofing pattern addressed in CSA's **Using Zero Trust to Counter Identity Spoofing & Abuse** [12], which argues for verifying counterparties through continuous, contextual signals rather than a single point-in-time check such as a video call or a plausible-looking company domain. Applied to recruitment workflows, that guidance supports the verified-channel practice recommended above: confirming a recruiter's identity through an independent, previously established corporate channel before any file or software exchange takes place.

AI-Weaponized Phishing: Nation-State Quality at Commodity Scale [6] (June 14, 2026) situates this campaign within the broader pattern the note identifies of nation-state actors — it names Iranian, Russian, Chinese, and North Korean clusters specifically — adopting increasingly high-fidelity social engineering tradecraft that has already diffused into commodity criminal tooling. UAC-0145's use of Telegram screening, a full video interview cycle, and a plausible corporate impersonation reflects the same operational patience and production quality that note attributes to state-directed campaigns, reinforcing its recommendation that security awareness programs be recalibrated to assume any unsolicited professional contact, however credible it appears, may be adversary-controlled.

Beyond these artifacts, this campaign's exploitation of implicit trust in a familiar, Microsoft- or industry-recognized tool category maps to the identity and access management and threat and vulnerability management domains of CSA's **AI Controls Matrix (AICM) v1.1** [8], the organization's current control framework superseding the Cloud Controls Matrix, particularly its provisions for verifying the integrity and provenance of software introduced to managed endpoints.

References

[1] The Hacker News. "Sandworm-Linked UAC-0145 Uses Fake Job Interviews to Push VPN That Can Run Commands." The Hacker News, August 11, 2026.

[2] BleepingComputer. "Sandworm hackers target IT pros with trojanized WireGuard VPN client." BleepingComputer, August 2026.

[3] Cyber Press. "Sandworm Fake Job Interviews Push Trojanized WireGuard VPN to Infect IT Professionals." Cyber Press, August 12, 2026.

[4] Cyber Security News. "Sandworm Fake Job Interviews Push Trojanized WireGuard VPN to Infect IT Professionals." Cyber Security News, August 12, 2026.

[5] Cloud Security Alliance. "FortiBleed: Default Credential Exploitation and Mass Fortinet Compromise." Cloud Security Alliance, June 20, 2026.

[6] Cloud Security Alliance. "AI-Weaponized Phishing: Nation-State Quality at Commodity Scale." Cloud Security Alliance, June 14, 2026.

[7] Cloud Security Alliance. "AI Superpersuasion: Influence Operations and Enterprise Security Risk." Cloud Security Alliance, June 28, 2026.

[8] Cloud Security Alliance. "AI Controls Matrix (AICM) v1.1." Cloud Security Alliance, 2026.

[9] Google Cloud (Mandiant). "APT44: Unearthing Sandworm." Google Cloud Blog, April 17, 2024.

[10] CNN. "White House blasts Russia for NotPetya cyberattack." CNN Politics, February 15, 2018.

[11] Cloud Security Alliance. "An Executive View on How Zero Trust Protects Organizations by Securely Connecting Users to Resources from Anywhere." Cloud Security Alliance.

[12] Cloud Security Alliance. "Using Zero Trust to Counter Identity Spoofing & Abuse." Cloud Security Alliance.

