LECCIONES APRENDIDAS

No necesitas entrar en el sistema operativo: ARMORY CASTLE, dos zero-days y el compromiso del plano de control de VMware

Análisis técnico de una operación de ciberespionaje atribuida a un actor presuntamente vinculado a China. Investigación DFIR, inteligencia de amenazas, MITRE ATT&CK, reglas Sigma y estrategias de defensa.

Fecha: 20 de noviembre de 2026
Ámbito: DFIR · Threat Intelligence · Seguridad de infraestructuras · Virtualización
Caso analizado: ARMORY CASTLE / UNC5221
Fuentes principales: Google Threat Intelligence Group, Mandiant, Broadcom y BeyondTrust.

Fuente: Google/Mandiant — ARMORY CASTLE Exploits vSphere (19/11/2026)

- **19 de noviembre de 2026:** publicación de la noticia original por Google Threat Intelligence Group y Mandiant.
- **Finales de agosto de 2026:** primeros ataques documentados contra BeyondTrust, utilizados como vía de entrada.
- **Principios de septiembre de 2026:** primeras evidencias de explotación del zero-day de VMware vCenter (**CVE-2026-123890**) por ARMORY CASTLE.

Lo más relevante: la presunta APT china habría estado explotando la vulnerabilidad de VMware aproximadamente **dos meses antes de su divulgación pública**.

1. Introducción: por qué debería importarte aunque no utilices VMware

Imagina que una organización dispone de servidores Windows y Linux correctamente parcheados, soluciones EDR desplegadas, autenticación multifactor, segmentación de red y un SOC que supervisa continuamente sus sistemas.

Aun así, un atacante consigue extraer información sensible de un controlador de dominio sin necesitar ejecutar malware dentro de su sistema operativo.

¿Cómo es posible?

Porque ese servidor no es una máquina física aislada. Es una máquina virtual cuyo disco, memoria, configuración y ciclo de vida dependen de una infraestructura administrada desde una capa superior.

El atacante no necesita vencer todas las defensas del sistema operativo invitado si consigue controlar la plataforma que lo administra.

Este es el problema que ilustra la investigación publicada el 19 de noviembre de 2026 por Google Threat Intelligence Group (GTIG) y Mandiant.

Según los investigadores, ARMORY CASTLE, un actor de ciberespionaje presuntamente vinculado a China y anteriormente identificado como UNC5221, encadenó dos vulnerabilidades de día cero para pasar de una solución de soporte remoto comprometida al entorno interno de una organización y, finalmente, al plano de control de VMware vSphere.

Una vez allí, utilizó herramientas especializadas para comprometer la autenticación de vCenter, obtener credenciales, acceder a hipervisores ESXi y extraer archivos directamente de discos virtuales.

La relevancia de este caso no se limita a VMware.

En prácticamente cualquier arquitectura moderna existe alguna forma de **plano de control**: la capa desde la que se crean, administran, conectan, modifican o eliminan los recursos informáticos.

Puede ser VMware vCenter, una plataforma de administración de hipervisores, una consola de infraestructura cloud, Kubernetes, una solución de copias de seguridad o un sistema de orquestación.

El principio es el mismo:

> Cuando un atacante controla la infraestructura que gobierna los sistemas, las defensas instaladas dentro de esos sistemas pueden dejar de ser suficientes.

Por eso este caso interesa a responsables de seguridad, equipos de respuesta a incidentes, analistas de inteligencia de amenazas, administradores de sistemas, arquitectos cloud y organizaciones que ni siquiera utilizan VMware.

No estamos ante otro incidente de ransomware que cifra máquinas virtuales. Estamos ante una operación de espionaje en la que el objetivo es acceder a información sensible aprovechando una capa de administración que muchas organizaciones no monitorizan con la misma intensidad que sus servidores.

Y, especialmente, estamos ante un caso que permite formular una pregunta incómoda:

¿Estamos protegiendo mejor nuestros servidores que los sistemas capaces de controlarlos?

---

2. Resumen ejecutivo

El 19 de noviembre de 2026, Mandiant y GTIG documentaron la explotación activa de **CVE-2026-123890**, una vulnerabilidad crítica de autenticación en VMware vCenter Server.

Broadcom la calificó con **CVSS v3.1 de 9,8 sobre 10** y publicó actualizaciones de seguridad mediante el boletín **VMSA-2026-0009.1**.

La investigación vincula la explotación con ARMORY CASTLE, un actor presuntamente asociado al ecosistema de ciberespionaje de la República Popular China.

Los principales hallazgos son:

Elemento	Hallazgo
Actor	ARMORY CASTLE, anteriormente UNC5221
Atribución	Presunto vínculo con China, según GTIG
Objetivo	Espionaje y acceso a información de alto valor
Primer zero-day	CVE-2026-102764, BeyondTrust Remote Support

Segundo zero-day	CVE-2026-123890, VMware vCenter Server
Severidad del segundo	CVSS 9,8, crítica
Herramienta de explotación	BONESAW
Herramienta de falsificación de identidad	SAWSNAKE
Persistencia y acceso	Webshells, puertas traseras, cuentas SSO y claves SSH
Impacto observado	Compromiso de vCenter, acceso a ESXi y extracción de archivos desde discos virtuales

La conclusión central es que la explotación de vCenter no fue el punto de entrada inicial, sino una fase posterior de una operación de movimiento lateral.

Este matiz es importante: una organización puede mantener su vCenter sin exposición directa a Internet y seguir estando en riesgo si otro sistema interno comprometido puede alcanzar sus interfaces de administración.

---

3. Entender el objetivo: qué es realmente el plano de control de VMware

Antes de analizar la vulnerabilidad, conviene distinguir tres componentes.

3.1. ESXi: el hipervisor

VMware ESXi es el hipervisor que permite ejecutar múltiples máquinas virtuales sobre un mismo servidor físico.

Administra recursos como CPU, memoria, almacenamiento y conectividad virtual.

Su posición en la arquitectura le proporciona capacidades sobre los sistemas invitados que no dependen necesariamente de las credenciales de Windows o Linux utilizadas dentro de cada máquina.

3.2. vCenter Server: la consola central de administración

vCenter es el sistema desde el que se gestionan múltiples hipervisores ESXi y sus recursos.

Permite administrar inventarios, permisos, configuraciones, máquinas virtuales, operaciones de almacenamiento y otros componentes de la infraestructura.

En muchas organizaciones se despliega como **vCenter Server Appliance (VCSA)**, una máquina virtual basada en Photon OS.

3.3. Las máquinas virtuales: los sistemas invitados

Son los servidores Windows y Linux donde normalmente instalamos antivirus, EDR, agentes de monitorización y herramientas de seguridad.

Pero existe una diferencia fundamental entre controlar un sistema invitado y controlar la infraestructura desde la que se administra.

Un administrador con suficientes privilegios sobre vSphere puede realizar operaciones sobre máquinas virtuales sin autenticarse necesariamente en sus sistemas operativos.

Esto no significa que cualquier compromiso de vCenter proporcione automáticamente acceso ilimitado a todos los datos de todas las máquinas. Las posibilidades dependen de los privilegios obtenidos, la configuración, el almacenamiento y las protecciones existentes.

Sin embargo, el caso ARMORY CASTLE demuestra que, en determinadas condiciones, el atacante puede llegar a **extraer datos directamente del almacenamiento virtual**, evitando las rutas tradicionales de acceso al sistema operativo invitado.

Ese es el cambio de perspectiva que debería orientar la estrategia defensiva.

---

4. Qué ocurrió: reconstrucción cronológica del incidente

La investigación de Mandiant permite reconstruir una cadena de ataque que comienza fuera de VMware y termina en la extracción de información sensible desde un hipervisor.

Cronología documentada

Fecha	Acontecimiento

13 de julio de 2026	Marca temporal de compilación recuperada en BONESAW. Es un indicio, no una fecha de desarrollo demostrada.
29 de julio de 2026	Divulgación de CVE-2026-59309, otra vulnerabilidad distinta de vCenter.
Finales de agosto de 2026	Explotación de CVE-2026-102764 en BeyondTrust Remote Support, según Mandiant.
Principios de septiembre de 2026	Evidencias de ejecución de BONESAW contra un vCenter mediante CVE-2026-123890.
15 de octubre de 2026	GTIG y Mandiant publican la investigación sobre BeyondTrust.
19 de noviembre de 2026	Divulgación pública de CVE-2026-123890 y publicación del análisis de ARMORY CASTLE contra vSphere.
20 de noviembre de 2026	El boletín de Broadcom figura actualizado a la versión VMSA-2026-0009.1.

Una precisión importante: la marca temporal de compilación de BONESAW puede manipularse. Por tanto, no demuestra por sí sola cuándo se desarrolló la herramienta.

La evidencia de explotación procede principalmente de la reconstrucción forense de los incidentes.

La cadena de ataque, de principio a fin

Fase 1 — Acceso inicial: explotación de BeyondTrust Remote Support.

Fase 2 — Persistencia: instalación de herramientas maliciosas en el dispositivo comprometido.

Fase 3 — Movimiento lateral: acceso a un appliance Linux interno con menor cobertura de seguridad.

Fase 4 — Escalada de privilegios: obtención de capacidades de root.

Fase 5 — Reconocimiento: identificación de servicios e interfaces de VMware.

Fase 6 — Explotación de vCenter: utilización de BONESAW y CVE-2026-123890.

Fase 7 — Compromiso de identidad: extracción de secretos, creación de cuentas SSO y falsificación de tokens SAML.

Fase 8 — Acceso a ESXi: uso de credenciales administrativas y establecimiento de persistencia SSH.

Fase 9 — Acceso a datos: exploración de volúmenes VMFS y extracción de archivos desde discos virtuales.

Fase 10 — Exfiltración y persistencia adicional: preparación de archivos, transferencia entre sistemas comprometidos y despliegue de puertas traseras.

El aspecto distintivo es la combinación de dos zero-days con herramientas diseñadas específicamente para diferentes plataformas.

No se trató de un único exploit lanzado contra un servidor vulnerable.

Fue una operación en varias etapas.

---

5. El primer zero-day: BeyondTrust como puerta de entrada

La primera vulnerabilidad documentada fue **CVE-2026-102764**, que afectaba a BeyondTrust Remote Support.

Este producto se utiliza para prestar soporte técnico y administrar remotamente equipos.

Precisamente por su función, puede disponer de conectividad y permisos privilegiados dentro de una organización.

Mandiant observó que ARMORY CASTLE explotó la vulnerabilidad antes de su divulgación pública.

El fallo permitía realizar operaciones de carga de archivos no autorizadas y podía facilitar el control del dispositivo afectado.

La investigación anterior de GTIG identificó herramientas como:

- **SAWFLY:** webshell utilizada para mantener acceso.
- **BARECLAW:** herramienta para facilitar comunicaciones y movimiento lateral.

También se observaron técnicas de persistencia y eliminación de rastros.

La vulnerabilidad se corrigió mediante **BT26-05**, según el fabricante.

Primera lección: el riesgo de los sistemas de acceso privilegiado

Las herramientas de soporte remoto, VPN, gestión de identidades, administración de copias de seguridad y otros appliances de infraestructura tienen algo en común: suelen disfrutar de conectividad excepcional.

Por eso resultan tan atractivos para un atacante.

El problema no es únicamente que puedan contener una vulnerabilidad.

El problema es que, una vez comprometidos, pueden convertirse en un punto de partida excelente para alcanzar sistemas que nunca estuvieron expuestos directamente a Internet.

Un sistema de administración comprometido puede heredar oportunidades de acceso que un equipo de usuario convencional no tendría.

---

6. El segundo zero-day: CVE-2026-123890 y la ruptura de la autenticación de vCenter

Este es el núcleo técnico de la investigación.

6.1. Qué vulnerabilidad se explotó

CVE-2026-123890 es una vulnerabilidad crítica de **omisión de autenticación** en VMware Directory Service.

Según Mandiant, el problema afecta al mecanismo de autenticación **Secure Remote Password (SRP-6a)** implementado sobre interfaces **DCE/RPC**.

La explotación se relaciona con el servicio `rpc_srp_verifier`, accesible mediante TCP/2012.

En condiciones normales, el protocolo SRP está diseñado para permitir una autenticación segura sin transmitir directamente la contraseña.

Sin embargo, la implementación vulnerable presentaba problemas criptográficos y de autorización que permitían obtener información de la clave de sesión.

Mandiant describe, además, el aprovechamiento de características predecibles relacionadas con la generación de identificadores UUIDv1.

El resultado era especialmente grave:

Un atacante con acceso de red al servicio vulnerable podía establecer una sesión autenticada como `administrator@vsphere.local` sin conocer su contraseña.

6.2. Por qué el fallo era crítico

El problema no consistía simplemente en revelar información de configuración.

La sesión obtenida podía utilizarse para realizar operaciones privilegiadas contra los servicios de directorio de vCenter.

La investigación documenta el abuso de la interfaz RPC `vmdirdbcp` para obtener archivos, incluida la base de datos:

`/storage/db/vmware-vmdir/data.mdb`

Esta base de datos es un objetivo de gran valor porque contiene información del directorio de vCenter y material relacionado con sus mecanismos de identidad.

Una vez obtenida, el atacante podía extraer secretos y preparar nuevas operaciones de autenticación fraudulenta.

6.3. BONESAW: la herramienta utilizada

Mandiant identificó una utilidad desarrollada en Go, denominada **BONESAW**, que automatizaba varias fases del compromiso.

Entre sus capacidades documentadas destacan:

1. Explotar CVE-2026-123890.
2. Obtener una copia de la base de datos LMDB de vCenter.
3. Extraer material criptográfico asociado a los servicios de identidad.
4. Crear cuentas administrativas SSO.
5. Facilitar el compromiso de los mecanismos de autenticación y autorización de vCenter.

El análisis del binario recuperado reveló el nombre interno `vcenter-db-teamcheck-go` y artefactos de compilación compatibles con un entorno de desarrollo Windows.

Estos datos ayudan a caracterizar la herramienta, pero no permiten identificar por sí solos a sus desarrolladores.

6.4. Diferencia con CVE-2026-59309

Es importante no confundir **CVE-2026-123890** con **CVE-2026-59309**.

Ambas afectan a mecanismos de autenticación de vCenter, pero no son la misma vulnerabilidad.

CVE-2026-59309 afecta a una ruta LDAP del servicio `vmdir`, relacionada con los puertos TCP/389 y TCP/636.

CVE-2026-123890 afecta a una interfaz DCE/RPC relacionada con SRP y TCP/2012.

Mandiant documentó que BONESAW se ejecutó contra un servidor que ya estaba parcheado frente a CVE-2026-59309.

Esto es relevante porque demuestra que el segundo fallo no podía considerarse resuelto por la actualización que corregía el primero.

6.5. Versiones afectadas y actualizaciones

El boletín oficial de Broadcom establece las siguientes versiones corregidas:

Producto o rama	Actualización indicada
vCenter 9.1.x	9.1.1.0200
vCenter 9.0.x	9.0.2.0200
vCenter 8.0	8.0 U3m
vCenter 7.0	Consultar con Broadcom si existe contrato de soporte extendido
VMware Cloud Foundation 5.x	5.2.5

También existen indicaciones específicas para otras soluciones que integran vCenter.

Broadcom no publica un workaround oficial para esta CVE. La corrección recomendada es instalar la actualización correspondiente.

La restricción de conectividad hacia los servicios vulnerables puede reducir la exposición, pero no sustituye al parche.

El boletín VMSA-2026-0009.1 incluye otras vulnerabilidades —CVE-2026-123891, CVE-2026-123892, CVE-2026-123893 y CVE-2026-123894— en componentes como SDDC Manager y VMware Tools.

No debe interpretarse que ARMORY CASTLE explotó todas ellas: **la explotación documentada en este incidente corresponde a CVE-2026-123890 y, previamente, a CVE-2026-102764**.

---

7. Del control de vCenter a la falsificación de identidad: Golden SAML

Una de las fases más interesantes del caso es la utilización de **SAWSNAKE**.

Esta herramienta, también desarrollada en Go, permitía trabajar con material criptográfico de los servicios de identidad de vCenter.

Mandiant la relaciona con técnicas de **Golden SAML**.

¿Qué significa Golden SAML?

SAML es un estándar utilizado para intercambiar información de autenticación entre sistemas.

En un funcionamiento legítimo, un proveedor de identidad emite una aserción firmada que otro servicio acepta como prueba de autenticación.

Si un atacante obtiene la clave privada utilizada para firmar esas aserciones, puede llegar a crear tokens que aparenten proceder de una autoridad legítima.

El problema deja de ser únicamente el robo de una contraseña.

El atacante puede intentar **fabricar una prueba de identidad aparentemente válida**.

¿Qué hizo SAWSNAKE?

Según la investigación, SAWSNAKE permitía:

- Analizar la base de datos `data.mdb`.
- Extraer claves privadas y certificados de confianza.
- Generar aserciones SAML firmadas.
- Obtener cookies de sesión administrativas.
- Suplantar identidades frente a vCenter.

Esto tiene una consecuencia importante para las estrategias basadas en MFA.

La autenticación multifactor no protege frente a todos los escenarios si el atacante consigue comprometer la autoridad que emite o valida los tokens de autenticación.

No significa que MFA sea inútil.

Significa que la seguridad de las claves de firma, los servicios de identidad y las relaciones de confianza es tan importante como la autenticación del usuario.

Lección para el SOC

Una autenticación técnicamente válida no garantiza que sea legítima.

Cuando se sospecha que una clave de firma ha sido comprometida, no basta con revisar intentos fallidos de inicio de sesión.

Hay que investigar las sesiones aceptadas, las identidades creadas, los cambios de privilegios y la integridad de la infraestructura de confianza.

---

8. Las TTPs observadas: cómo avanzó ARMORY CASTLE

8.1. Reconocimiento de la infraestructura

Tras comprometer el appliance Linux interno, el actor realizó exploraciones de red utilizando `nmap`.

Mandiant observó búsquedas dirigidas a puertos como:

- TCP/443: interfaces web y API.
- TCP/2012: servicio RPC relevante para la explotación.
- TCP/389: LDAP.
- TCP/5480: administración de VCSA.
- TCP/9080: otro servicio incluido en el reconocimiento observado.

También se investigaron rutas como `/sdk` y `/ui`.

Este comportamiento indica que el atacante estaba identificando específicamente sistemas VMware.

Oportunidad defensiva: las conexiones de un appliance de soporte o de una máquina Linux no administrativa hacia TCP/2012 de vCenter deberían considerarse anómalas salvo justificación operativa.

8.2. Webshells en el appliance Linux

Mandiant identificó dos herramientas:

SLIMSAVAGE, desplegada como `ActiveSer.jsp`, proporcionaba ejecución remota de comandos.

CAGEMATCH, desplegada como `u.jsp`, permitía cargar archivos y preparar nuevas herramientas.

La combinación es significativa: una webshell ejecutaba órdenes y la otra facilitaba el traslado de binarios.

8.3. Escalada de privilegios

El actor utilizó `gdb` para modificar permisos relacionados con `/bin/env`, habilitando una vía de ejecución privilegiada mediante el bit SUID.

Esta técnica merece atención porque no depende de una vulnerabilidad de kernel convencional.

Aprovecha herramientas legítimas del sistema para modificar su comportamiento de seguridad.

Por tanto, el *hunting* debe contemplar tanto la ejecución anómala de depuradores como los cambios de permisos en binarios sensibles.

8.4. Compromiso de vCenter

Una vez preparado el entorno, el atacante utilizó BONESAW contra vCenter.

Mandiant observó la creación de una cuenta administrativa SSO y la extracción de la base de datos LMDB.

Posteriormente se desplegó SAWSNAKE para facilitar la falsificación de tokens SAML.

8.5. Puertas traseras y persistencia

La investigación identifica también **BEARBONES**, una puerta trasera basada en una versión modificada de Dropbear SSH.

Entre sus características se encuentran la alteración del mecanismo de autenticación y la capacidad de registrar credenciales.

Un indicador especialmente distintivo es su banner SSH:

`SSH-2.0-pulsesecure_23345`

Mandiant observó intentos de escucha en los puertos TCP/53 y TCP/33992.

También identificó **LAZYBONES**, una puerta trasera escrita en Go que utilizaba Slack como canal de mando y control.

Su comunicación se realizaba mediante HTTPS, con mensajes cifrados y codificados.

La utilización de un servicio legítimo como Slack complica las estrategias basadas exclusivamente en bloquear dominios sospechosos.

8.6. Habilitación temporal de SSH en vCenter

ARMORY CASTLE utilizó la API de administración de VCSA para habilitar SSH.

La investigación menciona una petición HTTP `PUT` hacia:

`/rest/appliance/access/ssh`

Tras acceder y obtener información sensible, el actor volvió a deshabilitar el servicio.

Esto proporciona una lección de detección especialmente útil:

No basta con comprobar si SSH está habilitado ahora. Hay que registrar cuándo se habilitó, desde dónde, por quién y cuánto tiempo permaneció activo.

8.7. Acceso a ESXi y persistencia mediante claves SSH

Mandiant observó autenticaciones hacia ESXi utilizando `vpxuser` y `root`.

El actor también añadió una clave pública al archivo de claves autorizadas del usuario root.

Esta técnica permite mantener acceso SSH sin depender de la contraseña original.

8.8. La fase decisiva: extracción de datos desde discos virtuales

La investigación documenta la utilización de una herramienta denominada `vmfs_browser.py`.

El atacante la empleó para explorar volúmenes VMFS y archivos de discos virtuales VMDK.

Entre los objetivos observados se encontraban archivos asociados a un controlador de dominio Windows:

- `NTDS`
- `SYSTEM`
- `SECURITY`

Estos archivos son especialmente sensibles porque pueden contener información fundamental para la seguridad de Active Directory.

El actor preparó archivos comprimidos para su posterior transferencia.

Aquí se materializa la idea central de este artículo: el atacante accedió a información del sistema invitado desde la capa de almacenamiento virtual, sin necesitar que su herramienta de extracción se ejecutara dentro de ese Windows.

Un EDR instalado en el controlador de dominio puede no observar una lectura de disco realizada desde el hipervisor como si fuese un proceso ejecutado dentro del invitado.

Esto no implica que la operación sea indetectable.

Implica que los sensores deben situarse también en la infraestructura de virtualización y en sus rutas de administración.

---

9. ¿Cómo se descubrió la intrusión?

Conviene distinguir lo que sabemos de lo que no sabemos.

Lo documentado: Mandiant reconstruyó las operaciones mediante una investigación de respuesta a incidentes y evidencias forenses, incluyendo registros, artefactos y herramientas recuperadas.

Lo que no queda establecido públicamente: cuál fue exactamente la primera alerta que llevó a la organización afectada a solicitar ayuda.

No sería riguroso afirmar que la intrusión fue descubierta por un EDR, un SIEM, una alerta concreta de vCenter o una investigación proactiva si la fuente no lo especifica.

Evidencias que permitieron reconstruir el caso

La investigación describe diferentes tipos de artefactos:

En BeyondTrust:

Registros de acceso, syslog y otros archivos internos de la solución, además de evidencias de modificaciones y herramientas desplegadas.

En el appliance Linux:

Webshells JSP, binarios maliciosos, operaciones de descarga, comandos ejecutados y artefactos asociados a la escalada de privilegios.

En vCenter:

Creación de cuentas SSO, utilización de interfaces administrativas, habilitación de SSH y acceso a información sensible.

En ESXi:

Autenticaciones, claves SSH, actividad de exploración de volúmenes y operaciones sobre archivos de máquinas virtuales.

En la infraestructura de red:

Reconocimiento, conexiones entre sistemas y comunicaciones asociadas a herramientas de persistencia.

La dificultad forense

En una investigación de este tipo, los registros del sistema operativo invitado son insuficientes.

La línea temporal debe construirse correlacionando:

1. El appliance inicialmente comprometido.
2. Los sistemas intermedios utilizados para movimiento lateral.
3. El plano de administración de VMware.
4. Los hipervisores ESXi.
5. Los registros de red.
6. Los sistemas de identidad.
7. Las evidencias de transferencia y exfiltración.

Si uno de estos niveles no conserva registros, la reconstrucción puede quedar incompleta.

---

10. Mapeo MITRE ATT&CK: técnicas demostradas y alcance de la evidencia

El siguiente mapeo es una interpretación defensiva de las conductas descritas por Mandiant.

No constituye una matriz oficial publicada por GTIG para este incidente.

Técnica ATT&CK	ID	Evidencia en el caso
Exploit Public-Facing Application	T1190	Explotación inicial de BeyondTrust Remote Support
Network Service Discovery	T1046	Exploraciones con `nmap` de servicios VMware
Server Software Component: Web Shell	T1505.003	SLIMSAVAGE y CAGEMATCH
Unix Shell	T1059.004	Ejecución de comandos mediante webshell
Setuid and Setgid	T1548.001	Manipulación de permisos SUID
Exploitation of Remote Services	T1210	Explotación de vCenter desde un sistema interno
Create Account	T1136	Creación de una cuenta administrativa SSO
Forge Web Credentials: SAML Tokens	T1606.002	Capacidades Golden SAML de SAWSNAKE
Remote Services: SSH	T1021.004	Acceso lateral mediante SSH
Account Manipulation: SSH Authorized Keys	T1098.004	Incorporación de una clave pública a root
OS Credential Dumping: NTDS	T1003.003	Extracción de archivos NTDS desde un disco virtual
Archive Collected Data	T1560	Preparación de archivos comprimidos

Web Service	T1102	Uso de Slack como infraestructura C2 por LAZYBONES
Create or Modify System Process: Systemd Service	T1543.002	Persistencia mediante un servicio systemd
Indicator Removal: File Deletion	T1070.004	Eliminación de herramientas y artefactos

Hay que distinguir tres niveles de certeza.

Observado: las operaciones que Mandiant describe expresamente en su investigación.

Inferido mediante ATT&CK: la correspondencia entre esas operaciones y una técnica del marco.

No demostrado: cualquier comportamiento que podría haber sido posible gracias a los privilegios obtenidos, pero cuya ejecución no se ha documentado.

Por ejemplo, disponer de control administrativo sobre vCenter puede facilitar acciones destructivas. Sin embargo, eso no demuestra que ARMORY CASTLE ejecutara ransomware o destruyera máquinas virtuales en este incidente.

La inteligencia de amenazas pierde valor cuando convierte posibilidades técnicas en hechos.

---

11. Threat hunting: qué deberíamos buscar

Una búsqueda eficaz no debería limitarse a comprobar los indicadores publicados.

Hay que investigar las transiciones entre fases del ataque.

Hipótesis de hunting H1: acceso inesperado a servicios de administración

Hipótesis: un sistema que no pertenece a la red de administración intenta comunicarse con interfaces sensibles de vCenter.

Telemetría: firewall, NetFlow, Zeek, IDS/NDR.

Búsqueda: conexiones hacia vCenter en TCP/2012, 5480, 22, 389 y 636 procedentes de sistemas no autorizados.

Priorización: tráfico desde appliances de soporte, VPN, servidores de aplicaciones o equipos de usuario.

Hipótesis H2: creación de identidades SSO no autorizadas

Hipótesis: el atacante crea una cuenta administrativa para mantener acceso.

Telemetría: eventos de identidad y auditoría SSO de vCenter.

Búsqueda: creación de usuarios, asignación de privilegios globales y autenticaciones desde orígenes anómalos.

Priorización: cuentas creadas fuera de ventanas de cambio o utilizadas inmediatamente después de su creación.

Hipótesis H3: activación temporal de SSH

Hipótesis: un atacante habilita SSH en VCSA para extraer secretos y lo deshabilita después.

Telemetría: registros de VAMI, eventos de configuración, registros de autenticación y tráfico de red.

Búsqueda: peticiones a `/rest/appliance/access/ssh`, cambios de estado del servicio y sesiones SSH posteriores.

Priorización: habilitación y deshabilitación en un intervalo corto sin cambio autorizado.

Hipótesis H4: acceso al almacenamiento virtual desde el hipervisor

Hipótesis: un atacante explora archivos VMDK y extrae datos sensibles desde ESXi.

Telemetría: registros de ESXi, sesiones SSH, transferencias SCP, operaciones de almacenamiento y registros de red.

Búsqueda: accesos administrativos no habituales, uso de herramientas de exploración VMFS, archivos comprimidos en datastores y transferencias desde el hipervisor.

Priorización: operaciones sobre discos de controladores de dominio, servidores de identidad y otros activos críticos.

Hipótesis H5: persistencia en sistemas Linux intermedios

Hipótesis: el atacante utiliza webshells, herramientas de depuración y binarios en directorios temporales.

Telemetría: auditd, EDR Linux, integridad de archivos, syslog y registros del servidor web.

Búsqueda: archivos JSP inesperados, procesos shell iniciados por Java, uso de `gdb` sobre procesos privilegiados y modificaciones SUID.

---

12. Reglas Sigma propuestas para el SOC

Las siguientes reglas son **propuestas originales de detección basadas en las TTPs documentadas**. No son reglas oficiales de Mandiant ni han sido validadas en un laboratorio de producción.

Sigma es un formato de detección, no una fuente de telemetría.

Por tanto, cada regla necesita un origen de datos real y, en algunos casos, un mapeo de campos antes de convertirse a la sintaxis del SIEM correspondiente.

Regla 1. Detección de webshells asociadas a la campaña

Objetivo: identificar la creación de archivos con nombres utilizados por SLIMSAVAGE y CAGEMATCH.

Fuente necesaria: eventos de creación de archivos procedentes de FIM, auditd enriquecido o EDR Linux.

```
None
title: ARMORY CASTLE - Known JSP Webshell File Creation

id: 7d26d2f1-9725-40e9-98f0-7a93416b2031

status: experimental

description: Detects filenames associated with SLIMSAVAGE and
CAGEMATCH.

references:

  - https://cloud.google.com/blog/topics/threat-intelligence/
armory-castle-exploits-vsphere

logsource:

  category: file_event
```

```
  product: linux
detection:
  selection:
    TargetFilename|endswith:
      - '/ActiveSer.jsp'
      - '/u.jsp'
  condition: selection
falsepositives:
  - Legitimate application files with matching names
level: high
tags:
  - attack.persistence
  - attack.t1505.003
```

Limitación: el nombre del archivo puede cambiar fácilmente. Es una detección útil para buscar artefactos conocidos, no una defensa suficiente frente a nuevas variantes.

Regla 2. Ejecución sospechosa de GDB con archivos de comandos temporales

Objetivo: detectar una técnica semejante a la utilizada para modificar permisos privilegiados.

Fuente necesaria: telemetría de creación de procesos Linux.

```
None
title: ARMORY CASTLE - Suspicious GDB Command File Execution
id: 541bb99c-8304-4e35-bfe2-c4cb5fb65190
status: experimental
```

```
description: Detects GDB execution using command files staged
in temporary directories.
references:
  - https://cloud.google.com/blog/topics/threat-intelligence/
armory-castle-exploits-vsphere
logsource:
  category: process_creation
  product: linux
detection:
  selection_image:
    Image|endswith:
      - '/gdb'
  selection_command:
    CommandLine|contains|all:
      - '--command='
      - '/tmp/'
  condition: selection_image and selection_command
falsepositives:
  - Authorized debugging or maintenance activities
level: high
tags:
  - attack.privilege-escalation
  - attack.t1548.001
```

Limitación: detecta una implementación concreta de la técnica. No detectará todas las modificaciones de permisos SUID ni todos los usos maliciosos de GDB.

Debe complementarse con auditoría de cambios de permisos en binarios sensibles.

Regla 3. Habilitación de SSH mediante la API de VCSA

Objetivo: alertar sobre el uso de la interfaz administrativa empleada durante la intrusión.

Fuente necesaria: registros HTTP de VAMI que permitan identificar método y ruta.

Los campos `cs-method` y `cs-uri-stem` representan un esquema normalizado propuesto para los registros HTTP. Deben mapearse desde el formato real de VAMI.

```
None
title: VMware VCSA - SSH Enablement Through Management API
id: e8ac50d5-287f-45be-a1ca-8fb09c78ac34
status: experimental
description: Detects PUT requests to the VCSA SSH management
API.
references:
  - https://cloud.google.com/blog/topics/threat-intelligence/
armory-castle-exploits-vsphere
logsource:
  category: webserver
detection:
  selection:
    cs-method: 'PUT'
    cs-uri-stem: '/rest/appliance/access/ssh'
  condition: selection
falsepositives:
  - Authorized maintenance enabling or disabling SSH
level: high
tags:
  - attack.lateral-movement
```

```
  - attack.t1021.004
```

Recomendación: correlacionar la petición con el usuario autenticado, la IP de origen, el estado anterior y posterior de SSH y las conexiones al puerto 22.

La petición `PUT` identifica una modificación de configuración, pero no demuestra por sí sola que SSH haya quedado habilitado. Para afirmarlo es necesario conocer el contenido de la petición o el cambio de estado resultante.

Regla 4. Acceso a TCP/2012, interfaz sensible de vCenter

Objetivo: localizar conexiones hacia el servicio relacionado con CVE-2026-123890.

Fuente necesaria: eventos de conexión normalizados procedentes de firewall o NDR.

```
None
title: VMware vCenter - Connection to Sensitive RPC Service

id: 04f5df90-a3e4-4d67-93bc-56ccfc0474b8

status: experimental

description: Identifies connections to TCP port 2012 for
further investigation.

references:

  - https://cloud.google.com/blog/topics/threat-intelligence/
armory-castle-exploits-vsphere

logsource:

  category: network_connection

detection:

  selection:

    DestinationPort: 2012

  condition: selection

falsepositives:

  - Legitimate VMware infrastructure communication
```

```
level: low
tags:
  - attack.discovery
  - attack.t1046
```

Importante: esta regla es deliberadamente una búsqueda de exposición, no una alerta concluyente de explotación.

En producción debe restringirse a los destinos vCenter inventariados y excluir los orígenes autorizados.

Una conexión a TCP/2012 no demuestra el uso de BONESAW ni de CVE-2026-123890.

Regla 5. Banner SSH asociado a BEARBONES

Objetivo: identificar un artefacto de red de mayor especificidad.

Fuente necesaria: registros SSH de Zeek con el banner del servidor disponible en el campo `server`.

```
None
title: ARMORY CASTLE - BEARBONES Suspicious SSH Server Banner
id: 41a92b1f-6712-4b78-b56a-9f7d96bc9d54
status: experimental
description: Detects the SSH banner associated with BEARBONES.
references:
  - https://cloud.google.com/blog/topics/threat-intelligence/
armory-castle-exploits-vsphere
logsource:
  product: zeek
  service: ssh
detection:
```

```
  selection:
    server|contains: 'pulsesecure_23345'
  condition: selection
falsepositives:
  - Deliberate banner emulation
level: high
tags:
  - attack.command-and-control
```

Limitación: el campo exacto depende de la versión de Zeek y de cómo se normalicen sus registros. Además, el banner es modificable por el atacante.

Qué reglas priorizaría

Para un SOC que quiera aplicar las lecciones de este caso, la prioridad no debería ser desplegar únicamente reglas basadas en nombres de malware.

Las detecciones más valiosas son las que alertan sobre:

- Acceso no autorizado a servicios de administración.
- Creación de cuentas privilegiadas en vCenter.
- Habilitación de SSH y cambios de configuración.
- Manipulación de claves SSH.
- Actividad anómala sobre discos virtuales.
- Extracción de datos desde sistemas de infraestructura.
- Modificación de los mecanismos de identidad.

Las reglas Sigma anteriores son un punto de partida. Las detecciones de comportamiento deben completarse con correlaciones específicas del entorno.

---

13. Indicadores de compromiso publicados

GTIG incluyó los siguientes indicadores de red asociados a una presunta infraestructura de ocultación utilizada por el actor:

```
None
103.147.251.5
103.190.38.3
118.174.128.34
118.174.128.36
118.174.128.38
176.9.89.66
185.71.15.109
220.121.128.41
222.108.77.44
61.216.56.127
81.45.89.43
```

También publicó los siguientes nombres y rutas de archivos:

Indicador	Herramienta asociada
`ActiveSer.jsp`	SLIMSAVAGE
`u.jsp`	CAGEMATCH
`/tmp/123`	BONESAW
`/tmp/sss`	SAWSNAKE
`/tmp/dbr`	BEARBONES
`/usr/bin/vmtoolscs`	LAZYBONES

Otros indicadores relevantes incluyen:

- Banner SSH `SSH-2.0-pulsesecure_23345`.
- Creación de cuentas SSO administrativas no autorizadas.
- Uso anómalo de `/rest/appliance/access/ssh`.
- Claves SSH no autorizadas en cuentas privilegiadas.
- Directorios temporales y ocultos utilizados para preparar herramientas y archivos.

Advertencia CTI: las IP publicadas no deben interpretarse como direcciones permanentemente maliciosas. Pueden corresponder a infraestructura comprometida de terceros y cambiar de propietario o función.

Deben utilizarse para búsquedas históricas contextualizadas, no como prueba aislada de atribución.

GTIG también publicó reglas YARA específicas para SLIMSAVAGE, CAGEMATCH, BONESAW, SAWSNAKE, BEARBONES y LAZYBONES.

Para conservar sus condiciones originales y evitar errores de transcripción, recomiendo obtenerlas directamente del informe de GTIG enlazado en la bibliografía.

---

14. La pirámide del dolor aplicada a ARMORY CASTLE

La *Pyramid of Pain*, propuesta por David Bianco, describe el esfuerzo que debe realizar un adversario cuando el defensor detecta y bloquea distintos tipos de indicadores.

Cuanto más alto se sitúa una detección en la pirámide, más difícil suele resultar para el atacante adaptarse sin modificar su forma de operar.

Nivel 1. Hashes

Ejemplo: hashes de muestras concretas de BONESAW o SAWSNAKE.

Dolor para el atacante: bajo.

Recompilar una herramienta puede generar un hash diferente.

Utilidad: identificación forense y búsqueda de muestras conocidas.

Nivel 2. Direcciones IP

Ejemplo: las once direcciones publicadas por GTIG.

Dolor: relativamente bajo.

Un actor puede utilizar nuevas direcciones, proxies o dispositivos comprometidos.

Utilidad: correlación histórica, enriquecimiento y bloqueo temporal.

Nivel 3. Dominios y servicios de comunicación

Ejemplo: infraestructura utilizada para C2 y servicios legítimos abusados para transportar comandos.

Dolor: variable.

El caso LAZYBONES demuestra por qué bloquear un dominio completo como Slack puede resultar impracticable.

Utilidad: análisis de destinos, patrones de comunicación y cuentas o aplicaciones no autorizadas.

Nivel 4. Artefactos de red y host

Ejemplos:

- Banner SSH `pulsesecure_23345`.
- Archivos JSP inesperados.
- Uso de directorios temporales para herramientas.
- Claves SSH incorporadas a cuentas privilegiadas.

Dolor: medio.

El atacante puede modificar nombres y cadenas, pero hacerlo exige adaptar su infraestructura y procedimientos.

Nivel 5. Herramientas

Ejemplos:

- BONESAW.
- SAWSNAKE.
- BEARBONES.
- LAZYBONES.
- Utilidades de exploración de VMFS.

Dolor: alto.

Una detección robusta de las capacidades y el comportamiento de estas herramientas puede obligar al actor a rediseñar partes de su operación.

Nivel 6. Tácticas, técnicas y procedimientos

Ejemplos:

- Pasar de un appliance comprometido al plano de control.
- Explotar servicios de autenticación de infraestructura.
- Crear cuentas SSO privilegiadas.
- Obtener claves de firma.
- Habilitar SSH temporalmente.
- Acceder a discos virtuales para extraer información sensible.
- Mantener persistencia mediante claves SSH y servicios legítimos.

Dolor: muy alto.

Si una organización detecta estas conductas con independencia de la herramienta utilizada, el atacante debe modificar su estrategia operativa.

Conclusión de la pirámide

La prioridad defensiva debería evolucionar de:

«¿Conocemos las IP y los hashes de ARMORY CASTLE?»

a:

«¿Detectaríamos a cualquier actor que intentase comprometer nuestro plano de control y leer discos virtuales desde él?»

La segunda pregunta es mucho más útil.

---

15. ¿Cómo podría haberse evitado el ataque?

No existe una medida que garantice impedir todas las intrusiones, especialmente cuando intervienen vulnerabilidades de día cero.

Sin embargo, este incidente muestra varios puntos donde los defensores podrían haber interrumpido la cadena de ataque.

15.1. Tratar vCenter como un activo de máxima criticidad

En organizaciones con Active Directory, vCenter puede tener un impacto comparable al de otros activos Tier 0.

Su compromiso puede facilitar operaciones sobre servidores que contienen información crítica.

Por tanto, su administración debería estar sometida a controles especialmente estrictos.

15.2. Segmentar el plano de administración

No basta con impedir que vCenter sea accesible desde Internet.

También es necesario restringir qué sistemas internos pueden comunicarse con sus interfaces.

Un appliance de soporte remoto comprometido no debería disponer automáticamente de conectividad hacia los servicios de identidad y administración de vCenter.

Las políticas deberían limitar, según necesidad real, el acceso a puertos como TCP/2012, 443, 5480 y 22.

La segmentación debe aplicarse mediante controles independientes y verificables, idealmente fuera del sistema cuya seguridad se intenta proteger.

15.3. No confiar exclusivamente en el firewall de VCSA

Mandiant advierte de limitaciones del firewall gestionado desde VAMI.

En particular, señala problemas de granularidad, visibilidad y dependencia de la propia interfaz administrativa.

La investigación recomienda reforzar los controles de red y registrar intentos de conexión rechazados.

GTIG también enlaza un proyecto de Mandiant denominado **VCSA Hardening & Logging Tool**, que permite configurar políticas restrictivas y auditoría adicional.

No obstante, modificar directamente reglas `iptables` de VCSA puede afectar a la conectividad y a la compatibilidad con el soporte del fabricante.

No debería desplegarse en producción sin laboratorio, revisión de dependencias, control de cambios y procedimiento de recuperación por consola.

15.4. Utilizar estaciones de administración privilegiada

Las operaciones sobre vCenter y ESXi deberían realizarse desde equipos de administración controlados, con identidades diferenciadas y acceso restringido.

Una cuenta administrativa utilizada desde cualquier ordenador de la red aumenta la superficie de ataque.

15.5. Centralizar los registros antes del incidente

Es esencial enviar registros de VCSA y ESXi a un repositorio externo protegido.

También debe verificarse que la auditoría del sistema operativo del appliance se está reenviando realmente.

Mandiant advierte de que la configuración habitual de syslog de VCSA no garantiza, por sí sola, el envío remoto de todos los eventos de `auditd`.

Esto importa porque un atacante con acceso privilegiado puede borrar registros locales.

15.6. Monitorizar la integridad

Se recomienda vigilar:

- Archivos de configuración críticos.
- Claves SSH.
- Servicios systemd.
- Cambios de permisos SUID.
- Binarios inesperados.
- Modificaciones de mecanismos de autenticación.
- Nuevas cuentas y asignaciones de privilegios.

15.7. Restringir la salida de red

Los appliances de infraestructura no deberían disponer de acceso irrestricto a Internet.

Una política de salida basada en destinos y servicios autorizados puede dificultar la utilización de C2 y la exfiltración.

15.8. Parchar, pero no depender únicamente del parcheo

Instalar las actualizaciones oficiales es imprescindible.

Sin embargo, durante la ventana en la que una vulnerabilidad es desconocida, no existe un parche disponible.

La estrategia defensiva debe asumir que puede aparecer otro zero-day.

La segmentación, el mínimo privilegio, la monitorización y la protección de los servicios de identidad son precisamente las medidas que reducen el impacto de ese escenario.

---

16. Respuesta a incidentes: qué hacer si sospechamos que vCenter está comprometido

Si existe una sospecha fundada de compromiso, la respuesta no debería limitarse a actualizar vCenter y cambiar la contraseña de administrador.

El caso demuestra que el atacante puede haber obtenido claves, tokens, cuentas y acceso a sistemas dependientes.

Fase 1. Contención

Restringir las conexiones hacia el plano de administración y detener los accesos no autorizados.

Antes de apagar sistemas o modificar indiscriminadamente servicios, valorar la preservación de evidencias volátiles.

La estrategia concreta debe adaptarse a la gravedad del incidente y a los requisitos de continuidad.

Fase 2. Preservación forense

Recopilar, con procedimientos documentados:

- Registros de VCSA.
- Registros de ESXi.
- Eventos SSO.
- Historial de operaciones administrativas.
- Logs de VAMI.
- Registros SSH.
- Telemetría de red.
- Evidencias de modificaciones de cuentas y permisos.
- Artefactos sospechosos en sistemas Linux intermedios.

Siempre que sea posible, preservar copias verificables y mantener una cadena de custodia.

Fase 3. Análisis de identidad

Investigar:

- Cuentas SSO creadas recientemente.
- Privilegios globales asignados.

- Sesiones administrativas anómalas.
- Integridad de las claves de firma STS.
- Credenciales y secretos accesibles desde vCenter.

Si existen indicios de extracción de claves de firma, la recuperación debe incluir su sustitución mediante procedimientos compatibles con el fabricante y la invalidación de las relaciones de confianza afectadas.

Cambiar una contraseña no neutraliza necesariamente una clave de firma robada.

Fase 4. Análisis de ESXi

Revisar:

- Accesos de `root` y `vpxuser`.
- Claves SSH autorizadas.
- Sesiones SCP.
- Cambios en servicios de administración.
- Archivos inesperados en datastores.
- Operaciones de exploración y copia de discos virtuales.

Fase 5. Evaluación del alcance

Si se confirma que se extrajeron archivos de un controlador de dominio, la investigación debe ampliarse a Active Directory.

En ese caso, puede ser necesario planificar una recuperación de identidad más amplia, incluida la rotación controlada de secretos de dominio.

Estas operaciones deben realizarse después de asegurar la contención y siguiendo procedimientos específicos para evitar nuevas oportunidades de compromiso o interrupciones innecesarias.

Fase 6. Recuperación

Restablecer los componentes comprometidos desde un estado de confianza.

Aplicar las actualizaciones de Broadcom.

Eliminar persistencia, cuentas y claves no autorizadas.

Revisar las credenciales que pudieron quedar expuestas.

Validar que los sistemas recuperados no siguen confiando en material criptográfico comprometido.

Fase 7. Seguimiento

Mantener una vigilancia reforzada de autenticaciones, accesos a administración y conexiones entre segmentos.

Una operación de espionaje puede mantener múltiples vías de persistencia. Recuperar un único appliance no garantiza haber eliminado todas ellas.

---

17. Lecciones para AWS, Azure, Google Cloud y otras plataformas

Este incidente no demuestra que CVE-2026-123890 afecte a AWS, Azure o Google Cloud.

La vulnerabilidad es específica de VMware vCenter.

Lo que sí permite es extraer lecciones arquitectónicas aplicables a cualquier entorno donde exista una separación entre el plano de control y los recursos administrados.

AWS

En AWS, las capacidades de administración sobre recursos EC2 dependen de servicios y permisos del plano de control, principalmente mediante IAM y las API correspondientes.

Un compromiso de credenciales con permisos suficientes puede permitir operaciones sensibles sobre instancias, volúmenes y snapshots sin utilizar las credenciales del sistema operativo invitado.

Esto no implica que AWS Nitro presente la vulnerabilidad de VMware ni que ambos modelos de seguridad sean equivalentes.

La lección es proteger especialmente las identidades, las API y los permisos que gobiernan la infraestructura.

Microsoft Azure

En Azure, las identidades y roles con privilegios sobre máquinas virtuales, discos y recursos de infraestructura pueden realizar operaciones administrativas desde el plano de control.

Las estrategias defensivas deben incluir auditoría de operaciones, control de privilegios y supervisión de accesos a recursos sensibles.

Google Cloud

En Google Cloud, las identidades con permisos de administración sobre Compute Engine y otros servicios pueden modificar recursos mediante API.

La protección de IAM, los registros de auditoría y las restricciones sobre operaciones de almacenamiento y administración son fundamentales.

Kubernetes

El paralelismo también es útil en Kubernetes.

Un compromiso del plano de control o de identidades con privilegios elevados puede permitir modificar cargas de trabajo y acceder a recursos sin comprometer individualmente cada contenedor.

La lección común

La pregunta correcta no es únicamente:

«¿Está protegido el sistema operativo?»

También debemos preguntarnos:

«¿Quién puede administrar ese sistema desde fuera y qué ocurriría si esa autoridad fuese comprometida?»

---

18. Lecciones aprendidas: siete cambios que deberían salir de este incidente

Lección 1. El plano de control es un objetivo prioritario

La infraestructura de administración no es un componente auxiliar.

Puede ser una de las partes más críticas de toda la arquitectura.

Lección 2. Una red interna no es necesariamente una red de confianza

El acceso inicial se produjo en una solución diferente de VMware.

El movimiento lateral permitió alcanzar servicios que podían no estar expuestos a Internet.

Lección 3. Los zero-days hacen insuficiente una estrategia basada solo en parches

Parchar es obligatorio, pero no sustituye a los controles que dificultan la explotación y el movimiento lateral.

Lección 4. El compromiso de identidad puede sobrevivir al cambio de contraseñas

Si se extraen claves de firma o se crean cuentas administrativas ocultas, la recuperación debe abordar los mecanismos de confianza afectados.

Lección 5. El EDR del invitado no ve necesariamente lo que sucede en el hipervisor

La cobertura de detección debe extenderse a las API, los servicios de administración, los hipervisores y el almacenamiento.

Lección 6. Los registros de infraestructura son evidencias de primer nivel

Sin logs de vCenter, ESXi, SSO y red, el equipo DFIR puede no ser capaz de reconstruir cómo se accedió a la información.

Lección 7. Las TTPs importan más que los indicadores efímeros

Los nombres de archivos, los hashes y las direcciones IP cambian.

Las conductas necesarias para comprometer y utilizar el plano de control son más difíciles de sustituir por completo.

---

19. Plan de acción recomendado para un equipo de seguridad

Prioridad inmediata

1. Inventariar las instancias de vCenter y verificar versiones.
2. Aplicar las correcciones oficiales para CVE-2026-123890.
3. Revisar la exposición de BeyondTrust Remote Support y el estado de CVE-2026-102764.
4. Comprobar qué sistemas pueden alcanzar TCP/2012 y otras interfaces administrativas.
5. Buscar cuentas SSO y claves SSH no autorizadas.

6. Revisar los indicadores y artefactos publicados por GTIG.

Prioridad a corto plazo

1. Centralizar y proteger registros de VCSA, ESXi y SSO.
2. Configurar alertas de habilitación de SSH y creación de cuentas privilegiadas.
3. Separar las estaciones de administración del resto de la red.
4. Establecer políticas de acceso restrictivas hacia el plano de control.
5. Validar la capacidad real del SOC para investigar operaciones sobre discos virtuales.

Prioridad estructural

1. Clasificar el plano de virtualización como infraestructura crítica.
2. Incorporar escenarios de compromiso de vCenter a los ejercicios de respuesta a incidentes.
3. Probar la recuperación de identidad y de infraestructura desde estados de confianza.
4. Revisar los permisos de las herramientas de soporte, backup y orquestación.
5. Medir la cobertura de detección por técnica ATT&CK, no únicamente por producto instalado.

---

20. Conclusión: defender el sistema operativo ya no es suficiente

ARMORY CASTLE no destaca únicamente por haber explotado dos vulnerabilidades de día cero.

Destaca por la forma en que utilizó el acceso inicial para avanzar hacia componentes cada vez más estratégicos de la infraestructura.

Primero comprometió una solución de soporte remoto.

Después alcanzó un appliance Linux interno.

A continuación explotó vCenter, obtuvo capacidades administrativas, comprometió mecanismos de identidad y accedió a ESXi.

Finalmente, utilizó herramientas especializadas para extraer información desde discos virtuales.

La operación demuestra que un atacante avanzado puede buscar deliberadamente los lugares donde las defensas tradicionales tienen menor visibilidad.

No necesita necesariamente instalar malware dentro del servidor que contiene los datos.

Puede intentar acceder a esos datos desde la infraestructura que lo gobierna.

La respuesta defensiva no consiste en abandonar el EDR, la autenticación multifactor o el parcheo. Todas esas medidas siguen siendo esenciales.

Consiste en reconocer que son solo una parte de la arquitectura de seguridad.

La verdadera lección es que no basta con proteger las máquinas: también hay que proteger, monitorizar y poder investigar todo aquello que tiene autoridad sobre ellas.

Y esta lección seguirá siendo válida mucho después de que se hayan corregido las CVE de este incidente.

Bibliografía y fuentes técnicas

[1] Google Threat Intelligence Group y Mandiant. *ARMORY CASTLE Exploits vSphere*. 19 de noviembre de 2026. Fuente principal del análisis técnico, herramientas, TTPs e indicadores.
https://cloud.google.com/blog/topics/threat-intelligence/armory-castle-exploits-vsphere

[2] Broadcom. *VMSA-2026-0009.1: VMware vCenter, Tools, and SDDC Manager updates address multiple vulnerabilities*. Publicado el 19 de noviembre de 2026; actualizado el 20 de noviembre. Fuente oficial para severidad, productos afectados y versiones corregidas.
https://support.broadcom.com/web/ecx/support-content-notification/-/external/content/SecurityAdvisories/0/39495

[3] Google Threat Intelligence Group y Mandiant. *ARMORY CASTLE Exploits BeyondTrust*. 15 de octubre de 2026. Investigación del acceso inicial y la explotación de CVE-2026-102764.
https://cloud.google.com/blog/topics/threat-intelligence/armory-castle-exploits-beyondtrust

[4] Mandiant. *vSphere and BRICKSTORM Malware: A Defender's Guide*. 2 de abril de 2026. Referencia para arquitectura defensiva, registros y endurecimiento de VMware.
https://cloud.google.com/blog/topics/threat-intelligence/vsphere-brickstorm-defender-guide

[5] Mandiant. *VCSA Hardening & Logging Tool*. Repositorio técnico de referencia. Su aplicación requiere pruebas y revisión de compatibilidad.
https://github.com/mandiant/vcsa-hardening-tool

[6] Mandiant. *VMware ESXi Zero-Day Used by Chinese Espionage Actor to Perform Privileged Guest Operations on Compromised Hypervisors*. 13 de junio de 2023. Antecedente relevante de ataques contra la capa de virtualización.
https://cloud.google.com/blog/topics/threat-intelligence/vmware-esxi-zero-day-bypass

[7] Mandiant. *Chinese Espionage Group UNC3886 Found Exploiting CVE-2023-34048 Since Late 2021*. 19 de enero de 2024. Contexto histórico de explotación de VMware por actores de ciberespionaje.
https://cloud.google.com/blog/topics/threat-intelligence/chinese-vmware-exploitation-since-2021

[8] MITRE ATT&CK. Base de conocimiento de tácticas y técnicas adversarias.
https://attack.mitre.org/

[9] David Bianco. *The Pyramid of Pain*. Marco conceptual para priorizar detecciones según el coste que imponen al atacante.
https://detect-respond.blogspot.com/2013/03/the-pyramid-of-pain.html

---

Nota metodológica

Este artículo se basa en información pública disponible hasta el **20 de noviembre de 2026**.

Los hechos atribuidos a ARMORY CASTLE proceden principalmente de las investigaciones de GTIG y Mandiant. La evaluación de atribución geopolítica corresponde a esos investigadores y no constituye una identificación judicial ni una prueba pública concluyente de patrocinio estatal.

El mapeo MITRE ATT&CK, las hipótesis de *threat hunting*, las reglas Sigma y las recomendaciones de priorización constituyen análisis defensivo elaborado a partir de las evidencias publicadas.

Las reglas Sigma requieren validación, adaptación de campos y pruebas en el entorno de destino antes de utilizarse como detecciones de producción.