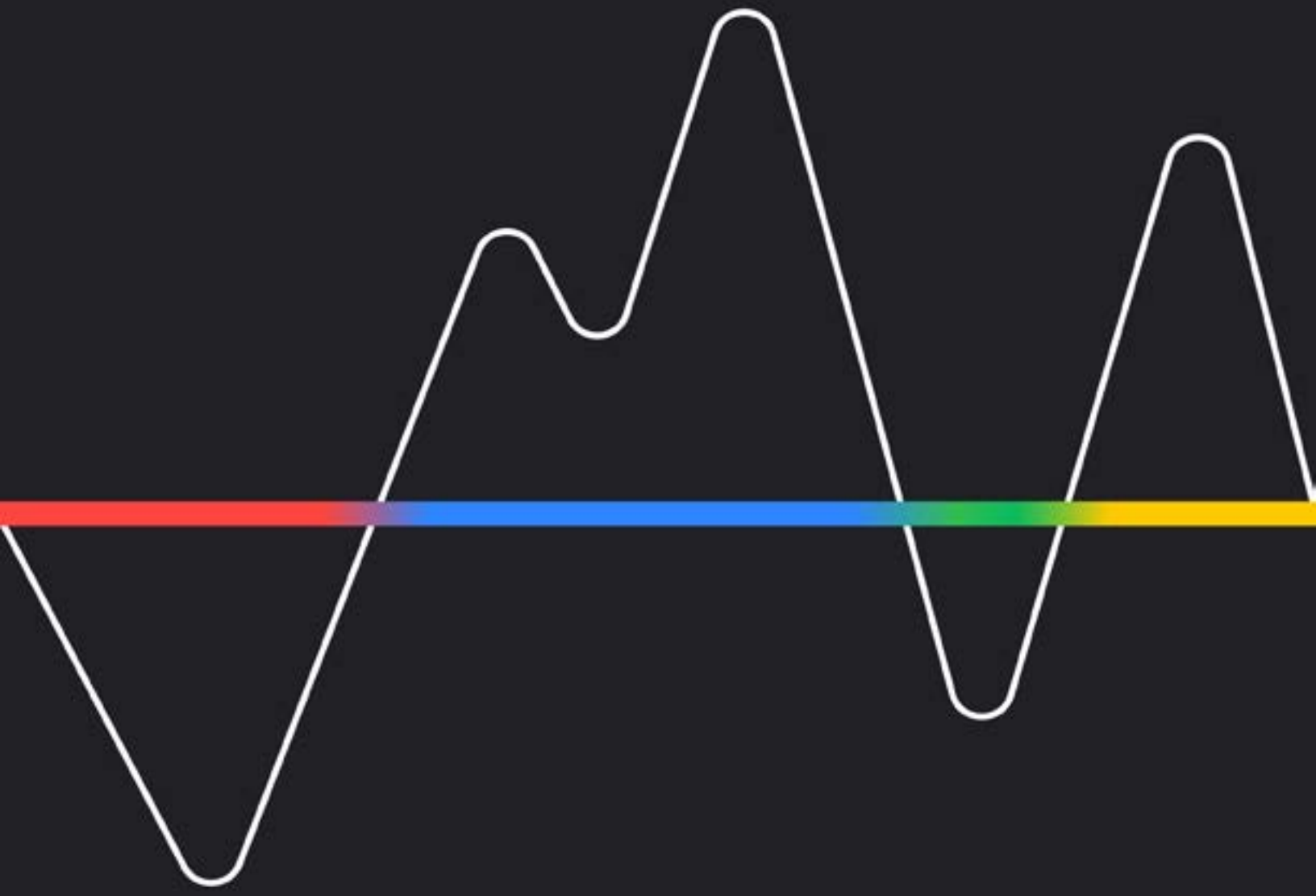
Mandiant

Google Cloud Security

M-Trends

Informe del 2026

Edición para ejecutivos



Edición para ejecutivos de M-Trends 2026

Prólogo

M-Trends es el análisis definitivo de las amenazas y tácticas que se usan en las vulneraciones, con base en más de 500,000 horas de investigaciones de incidentes de primera línea realizadas por Mandiant en el 2025. Junto con Google Threat Intelligence Group (GTIG), tenemos una visión completa del panorama actual de las amenazas, así como de las amenazas emergentes que impulsarán los ataques en el futuro.

En los informes recientes de GTIG, se confirma que los adversarios están adoptando la IA. Los agentes de amenazas están utilizando modelos de lenguaje grandes (LLM) para aplicar ingeniería social altamente personalizada, malware que consulta LLM durante la ejecución para evadir la detección y "ataques de síntesis" dirigidos a la lógica de aprendizaje automático de propiedad de la empresa. Los equipos rojos de Mandiant están incorporando técnicas basadas en IA en sus intervenciones para preparar a las organizaciones ante estas amenazas emergentes; sin embargo, los resultados de nuestro informe M-Trends 2026 demuestran que es fundamental mitigar las fallas humanas y sistémicas que permiten las violaciones de la seguridad.

Una de las principales conclusiones extraídas de nuestros análisis de incidentes del 2025 es que un subconjunto de adversarios permanece sin ser detectado en las redes durante períodos más prolongados, a menudo por medio de la persistencia en dispositivos perimetrales que normalmente carecen de telemetría estándar. Mandiant respondió a tantos incidentes de este tipo en el 2025 que la mediana mundial del tiempo de permanencia aumentó de los 11 días del período anterior a 14 días, en gran parte debido al espionaje a largo plazo y a las operaciones de trabajadores de TI de la RPDC.

Otros grupos de amenazas priorizan la velocidad. Estamos observando una tendencia creciente en la que los socios de acceso inicial trabajan directamente con grupos secundarios en lugar de vender el acceso en mercados clandestinos. Esto da lugar a una "transferencia" que a veces se produce en menos de 30 segundos, lo que crea una situación en la que las alertas "menores" pueden convertirse rápidamente en vulneraciones mayores.

De forma simultánea, los adversarios atacan sistemáticamente la infraestructura, como las copias de seguridad, los servicios de identidad y las capas de virtualización, para impedir la recuperación, lo que ejerce una enorme presión sobre las organizaciones para que paguen los rescates exigidos o se arriesguen a perder la capacidad de recuperarse.

Para lograr una verdadera resiliencia operativa, las organizaciones deben moverse al ritmo del adversario. Una parte importante de ello consiste en comprender cómo los adversarios logran el éxito. Las empresas pueden cerrar las brechas importantes de visibilidad y adoptar las medidas de defensa que se detallan en M-Trends 2026 para pasar de la recuperación reactiva a la contención proactiva antes de que una alerta menor se convierta en una vulneración catastrófica.

Cifras

Las métricas que se informan en M-Trends 2026 se basan en las investigaciones que los consultores de Mandiant llevaron a cabo entre el 1 de junio y el 31 de diciembre de 2025 sobre la actividad de ataques dirigidos.

Conclusión

Los atacantes están cambiando sus estrategias de acceso inicial; los exploits siguen siendo las más comunes, pero la ingeniería social altamente interactiva basada en la voz está escalando al segundo lugar. Al mismo tiempo, la mediana del tiempo de permanencia global ha aumentado, ya que los grupos de espionaje sofisticados y las amenazas internas priorizan el acceso sigiloso y a largo plazo. A medida que los adversarios buscan oportunidades para aprovechar y utilizar la IA como arma, explotar vulnerabilidades de día cero en dispositivos perimetrales y coordinar el traspaso de información entre los socios de acceso inicial y los grupos de ciberdelitos, las organizaciones deben ir más allá de las defensas estáticas para supervisar de forma continua el comportamiento de las identidades y la infraestructura, como la virtualización, que tradicionalmente ha estado fuera del alcance de las soluciones de EDR y otras herramientas de seguridad similares.

¿Qué debo saber?

- Los grupos con motivaciones financieras representaron el 41% de los conjuntos de amenazas observados en el 2025 (un descenso en comparación con el 55% del 2024), y los grupos de ciberespionaje aumentaron al 16% (frente al 8%).
- Por sexto año consecutivo, los vectores de infección más comunes fueron los exploits (32%). El phishing telefónico experimentó un aumento repentino hasta alcanzar el 11%, lo que lo convirtió en el segundo vector más frecuente, mientras que el phishing por correo electrónico registró un descenso sostenido del 14% en el 2024 al 6% en el 2025. Si bien tanto el phishing por correo electrónico como el phishing telefónico se engloban dentro de un concepto más amplio de ingeniería social, la distinción entre ambos es fundamental para los defensores: los ataques interactivos son significativamente más difíciles de detectar con los controles técnicos automatizados y requieren estrategias de identificación diferentes.
- En los incidentes relacionados con el ransomware, las vulneraciones previas fueron el vector de infección inicial más frecuente, con un 30%. Algunos conjuntos de amenazas se enfocan en establecer un punto de apoyo inicial en muchas organizaciones a través de vectores de infección oportunistas de gran volumen para, luego, vender o transferir este acceso a otros conjuntos de amenazas para la explotación posterior a la vulneración.
- Los sectores más afectados fueron la alta tecnología, las finanzas, los servicios empresariales y profesionales, y la salud, aunque el alcance total de los incidentes afectó a más de 16 verticales de la industria. Cabe destacar que las investigaciones en el sector de alta tecnología superaron a las de las organizaciones del sector financiero, que tuvieron la mayor participación en el 2024 y el 2023.
- En todas las investigaciones realizadas en el 2025, en el 52% de los casos las organizaciones detectaron por primera vez indicios de actividad maliciosa de manera interna (un aumento con respecto al 43% del 2024), mientras que en el 34% de los casos fueron notificadas por una entidad externa (una disminución con respecto al 43%). Los adversarios informaron a las organizaciones sobre una vulneración de seguridad en el 14% de los casos.
- En lo que respecta al ransomware, los adversarios avisaron al objetivo el 44% de las veces, la detección se produjo internamente en el 41% de los casos y las entidades externas notificaron el 15% de las veces. Un porcentaje mucho mayor de notificaciones por parte de los adversarios es coherente con el modelo de negocio del ransomware.
- En el 2024, la mediana del tiempo de permanencia a nivel mundial aumentó de 11 a 14 días. A través de la comparación de la distribución entre el 2024 y el 2025, se revela una leve disminución en los incidentes detectados en una semana o menos, y un cambio hacia tiempos de permanencia más prolongados (entre una semana y seis meses). Es probable que este cambio

refleje la cantidad de incidentes de ciberespionaje e incidentes relacionados con trabajadores de TI norcoreanos, en los que los conjuntos de amenazas priorizan el mantenimiento de un acceso sigiloso y a largo plazo; ambas categorías tuvieron una mediana de tiempo de permanencia de 122 días.

- De las familias de malware observadas en las investigaciones del 2025, el 36% fueron ataques de puerta trasera, el 11% de downloaders, el 10% de ransomware, el 10% de droppers y el 9% de ladrones de credenciales.
- Luego de cinco años consecutivos de ser la familia de malware observada con más frecuencia en las investigaciones de Mandiant, Cobalt Strike BEACON descendió al cuarto puesto. La familia de malware observada con mayor frecuencia fue el downloader GOLDVEIN.JAVA, seguido del ransomware REDBIKE (Akira).
- Los conjuntos de amenazas están utilizando tácticas sigilosas y malware ligero (como la puerta trasera BRICKSTORM) en dispositivos que no son compatibles con EDR. Tanto los grupos con motivaciones financieras como los de ciberespionaje abusan de las funcionalidades nativas de los entornos locales y en la nube, así como de herramientas legítimas, para reducir las oportunidades de detección.
- Los conjuntos de amenazas adoptan cada vez más herramientas de IA para aumentar su productividad en materia de reconocimiento, ingeniería social y desarrollo de malware. Además, los atacantes utilizan la IA como arma en entornos hackeados. Por ejemplo, se observó que el ladrón de credenciales QUIETVAULT revisaba las máquinas objetivo en busca de herramientas de CLI basadas en IA para ejecutar instrucciones predefinidas y buscar archivos de configuración.

¿Qué debemos hacer?

- **Ampliar la visibilidad más allá de los extremos tradicionales:** Implementa la detección avanzada de amenazas en todo el ecosistema. En concreto, se debe incorporar el análisis del tráfico de red para los dispositivos perimetrales sin EDR y aplicar una telemetría estricta para la infraestructura de virtualización.
- **Administrar de forma enérgica las superficies de ataque expuestas a Internet:** Dado que los exploits siguen siendo el principal vector de infección inicial por sexto año consecutivo, las organizaciones deben priorizar la aplicación rápida de parches, el análisis de vulnerabilidades y el aislamiento estricto de los servidores de aplicaciones web externos, que son frecuentemente objetivo de campañas de día cero y día n.
- **Adaptar las capacitaciones de concientización sobre seguridad para ir más allá de la bandeja de entrada del correo electrónico:** Si bien el phishing por correo electrónico sigue siendo una táctica habitual de los agentes de amenazas, Mandiant observa cada vez más que estos agentes utilizan otros vectores de infección iniciales, como el phishing telefónico interactivo, el robo de credenciales y tácticas como ClickFix. Capacita específicamente a los empleados y al personal del departamento de ayuda de TI para que sepan reconocer los intentos de ingeniería social en tiempo real basada en voz, los señuelos a través de apps de mensajería y las solicitudes no autorizadas de restablecimiento de MFA.
- **Cambiar a la verificación continua de identidad:** Dado que la ingeniería social interactiva suele eludir la MFA tradicional y que diversos agentes de amenazas, incluidos trabajadores de TI norcoreanos, buscan mantener el acceso a largo plazo, las organizaciones deben aplicar un estricto principio de privilegio mínimo, auditar periódicamente las integraciones de SaaS/nube y buscar de forma proactiva comportamientos de identidad anómalos entre empleados y contratistas remotos.
- **Actualizar las guías de IR para el modelo moderno de extorsión:** Asegúrate de que los planes de respuesta y recuperación ante incidentes aborden tanto el ransomware basado en encriptación como la extorsión a través del robo de datos. Realiza ejercicios de debate enfocados en la detección temprana de infecciones oportunistas para reducir los tiempos de respuesta y evitar que los socios de acceso inicial transfieran información a grupos secundarios.
- **Proteger los entornos de desarrolladores y las cadenas de herramientas de IA:** Se observa que los conjuntos de amenazas utilizan la IA como arma en entornos hackeados. Debido a que notamos que el malware abusa activamente de herramientas legítimas de línea de comandos de IA locales para identificar y robar tokens de GitHub y NPM, las organizaciones deberían adoptar principios del Secure AI Framework (SAIF) de Google. En concreto, con la "ampliación de la detección y la respuesta" para incluir herramientas de IA en el modelo de amenazas, los equipos de seguridad pueden establecer estándares de comportamiento y supervisar si se producen instrucciones anómalas o robos de datos no autorizados procedentes de estas utilidades.
- **Involucrar equipos rojos para la emulación de amenazas modernas:** Prueba periódicamente tus defensas a través de una emulación realista de las tácticas más recientes de los adversarios. Mide el tiempo que tardan los equipos de seguridad en detectar las técnicas de “living off the land”, la explotación de dispositivos perimetrales y la ingeniería social interactiva, y en responder a ellas.

Una infección menor hoy puede ser un ataque de ransomware mañana

Conclusión

La mayor colaboración entre los ciberdelincuentes redujo drásticamente el margen de maniobra para la defensa, lo que disminuye la mediana del tiempo entre el acceso inicial oportunista de un grupo y el momento en que un segundo grupo de amenazas tiene acceso a tan solo 22 segundos; en años anteriores, este tiempo rondaba las 8 horas. Este cambio exige que las organizaciones traten las alertas de bajo impacto como indicadores críticos, lo que requiere una corrección inmediata antes de que los actores de alto impacto puedan sacar provecho del acceso.

¿Qué debo saber?

- **El "tiempo de traspaso" se redujo drásticamente:** La mediana del tiempo transcurrido entre un evento de acceso inicial y el traspaso de información a un grupo de amenazas secundario se redujo de más de 8 horas en el 2022 a tan solo 22 segundos en el 2025. Esta disminución sugiere un cambio hacia la coordinación directa, en la que los socios de acceso inicial establecen el acceso específicamente para que los grupos secundarios lo utilicen de inmediato, en lugar de venderlo más adelante a través de canales clandestinos.
- **Las intrusiones de bajo impacto pueden ser precursoras de ataques de alto impacto:** Los agentes de amenazas utilizan un modelo de división del trabajo en el que grupos especializados emplean técnicas de bajo impacto, como anuncios maliciosos (malvertising) o actualizaciones falsas del navegador, para abrirse paso. Debido a que estos vectores iniciales parecen malware de bajo impacto, las organizaciones que buscan únicamente tácticas de alto impacto a menudo los pasan por alto hasta que es demasiado tarde.
- **Preparación previa de los socios:** En lugar de limitarse a vender una puerta de acceso a los mercados clandestinos, los socios de acceso inicial ahora preparan de antemano el malware, los túneles o las puertas traseras preferidos del grupo secundario durante la infección inicial. Eludir el mercado clandestino y preconfigurar el entorno permite a los actores posteriores aprovechar estos puntos de apoyo ya establecidos a su propio ritmo, lo que significa que están equipados por completo para iniciar operaciones de alto impacto desde el momento en que interactúan por primera vez con la red.

¿Qué debemos hacer?

- **Tratar las alertas de bajo impacto como indicadores críticos:** La disparidad entre la baja gravedad percibida de las infecciones iniciales y sus posibles consecuencias de alto impacto exige una reestructuración de las guías de respuesta. Los equipos de seguridad deben tratar las alertas rutinarias de malware como indicadores de alta prioridad de una transferencia inminente a un grupo secundario.
- **Establecer un conjunto básico de herramientas preaprobadas:** Los equipos de TI y de seguridad deben definir e implementar un conjunto de herramientas preaprobadas y almacenadas centralmente para los usuarios. Esto reduce el ruido de las instalaciones automáticas oportunistas y permite que los defensores identifiquen las desviaciones respecto al conjunto básico y respondan rápidamente a ellas.
- **Optimizar para la correlación de eventos y el contexto:** Los flujos de trabajo defensivos deberían enfocarse en la correlación de eventos, en lugar de en la revisión de alertas individuales. Enriquecer las alertas con datos contextuales permite a los analistas detectar patrones de comportamiento específicos que indican que se está produciendo una posible intrusión de alto impacto.
- **Corregir antes de que comience la actividad interactiva:** El objetivo de los defensores es corregir una intrusión durante la fase de acceso inicial, que a menudo es un evento no interactivo, antes de que un segundo grupo comience a realizar operaciones directas a través del teclado. Detener el ataque en esta fase de sistema único es mucho más eficaz que intentar recuperarse de la actividad de alto impacto posterior.

El ransomware ahora es un problema de resiliencia

Conclusión

Los grupos que usan ransomware ya no se limitan a encriptar los datos; destruyen activamente la capacidad de recuperación. Se centran en los sistemas más importantes, lo que, en la práctica, obliga a elegir entre pagar o reconstruir. Ahora, la verdadera resiliencia implica diseñar redes, de manera que las herramientas de recuperación estén segmentadas y protegidas. El bloqueo de los sistemas centrales y la dificultad para que los atacantes se muevan dentro de la red brindan a los equipos de seguridad una ventaja frente a las amenazas de ransomware actuales.

¿Qué debo saber?

- **El ransomware evolucionó hacia la negación de la recuperación:** Los operadores de ransomware cambiaron su objetivo principal: pasaron del simple robo de datos a la denegación de la recuperación. Los atacantes ahora se centran en las vías administrativas y del sistema, específicamente en los servicios de identidad, los planos de administración de la virtualización y la infraestructura de las copias de seguridad, para reducir la capacidad de recuperación de una organización y maximizar la presión para que esta pague.
- **La identidad es el nuevo perímetro:** Grupos de amenazas sofisticados manipulan el plano de control de identidades para obtener el control total de los entornos objetivo. Observamos que los atacantes aprovechan los parámetros de configuración erróneos para emitir certificados y crear cuentas de administrador que eluden la autenticación de varios factores (MFA) y la rotación de contraseñas. En algunos casos, los atacantes roban bases de datos completas o comprometen las cuentas de los usuarios en la nube para destruir la infraestructura de backend, lo que impide que los defensores accedan a sus propias cuentas de emergencia durante una crisis.
- **Se están destruyendo las redes de seguridad tradicionales:** La estrategia de depender de las copias de seguridad está fracasando, ya que los atacantes buscan activamente las arquitecturas de respaldo y las destruyen. Los agentes de amenazas realizan operaciones de reconocimiento para localizar las ubicaciones de almacenamiento y recuperar los parámetros de configuración de encriptación antes de borrar sistemáticamente los objetos de copia de seguridad del almacenamiento en la nube y los sistemas locales. En entornos locales, se observa que los atacantes desvinculan los entornos de virtualización de las plataformas de copia de seguridad o encriptan los puntos de recuperación locales, lo que hace que las guías estándares de respuesta ante incidentes resulten ineficaces, ya que no se dispone de las herramientas necesarias para ejecutarlas.

¿Qué debemos hacer?

- **Aumentar la seguridad mínima viable:** Las organizaciones deben considerar la identidad como su perímetro principal. Esto requiere endurecer las vías de alto impacto implementando la administración de identidades privilegiadas y el acceso condicional para garantizar que los privilegios administrativos tengan una duración limitada y se respeten. Los equipos de seguridad deben aplicar la MFA a todos los accesos, utilizar estaciones de trabajo con acceso privilegiado endurecidas y auditar los nombres principales de servicio (SPN). La telemetría debe ajustarse para detectar comportamientos de "living off the land" e identificar las apropiaciones administrativas antes de que tengan un impacto generalizado en el entorno.
- **Aislar los planos de control críticos de nivel 0:** Las plataformas de virtualización y administración deben tratarse como activos de "nivel 0", con las restricciones de acceso más severas. Las organizaciones deben implementar la segmentación de confianza cero para aislar estos sistemas y cortar formalmente la integración con Active Directory (AD) para evitar que una única vulneración de identidad se traduzca en una encriptación masiva. Usar una administración dedicada fuera de banda con cuentas locales protegidas con la MFA garantiza la separación estructural y evita situaciones en las que se pierdan simultáneamente las capacidades de producción y recuperación.
- **Mejorar la confiabilidad de la ruta de recuperación:** La capacidad de recuperación debe sobrevivir a la vulneración del entorno de producción. Las organizaciones necesitan establecer un entorno de recuperación aislado y dedicado en el que los equipos puedan limpiar, validar y almacenar en etapa intermedia los restablecimientos del sistema, lo que ayuda a mitigar un bucle de reinfección. Esto requiere mantener versiones sin conexión o inmutables de los activos de nivel 0, específicamente los catálogos de identidades y de copias de seguridad, para que no corran la misma suerte que la red de producción. Los planes de recuperación ante desastres deben considerar explícitamente la pérdida total del tejido de identidad principal.

Las intrusiones de varios años ponen de relieve una persistencia extrema

Conclusión

Lo ideal es prevenir, pero la preparación es imprescindible. Los agentes de amenazas sofisticados mantienen el acceso durante varios años explotando los puntos débiles y la confianza administrativa. Si no puedes demostrar el alcance de una intrusión debido a lagunas en el registro de datos, corres el riesgo de perder la confianza del cliente, ya que, en el peor de los casos, te verás en la obligación de suponer y divulgar un posible robo de datos. Considera la visibilidad como una auditoría continua para garantizar que puedas detectar y corregir estas amenazas antes de que se conviertan en crisis incontrolables.

¿Qué debo saber?

- **Sigilo y persistencia generalizados:** Algunos agentes de amenazas están logrando tiempos de permanencia superiores a un año, a menudo priorizando las credenciales legítimas sobre el malware personalizado. Su objetivo son las infraestructuras de virtualización y los dispositivos perimetrales no administrados, y utilizan herramientas integradas para camuflarse entre la actividad administrativa habitual y evitar que los detecten.
- **Brechas críticas de visibilidad:** Las estrategias de registro actuales no logran capturar el alcance total de estas intrusiones. En los casos que involucran BRICKSTORM (una puerta trasera sigilosa) con un tiempo de permanencia promedio de 393 días, la retención estándar de registros de 90 días impide que las organizaciones identifiquen el vector de acceso inicial. Además, la dependencia de EDR crea puntos ciegos, ya que los agentes se centran en los dispositivos perimetrales de la red y los protocolos de autenticación que están fuera del alcance de los agentes de extremo.
- **La oposición a los análisis forenses frena a los equipos de respuesta:** Se observa que hay agentes sofisticados que destruyen las pruebas necesarias para determinar el alcance de un incidente. Con el uso de técnicas para enmascarar las modificaciones de archivos y borrar los registros del sistema, los agentes de amenazas hacen que sea prácticamente imposible reconstruir la cronología del robo de datos. Con esta ausencia de pruebas forenses definitivas, las organizaciones se arriesgan a sufrir graves perjuicios relacionados con reglamentación y de reputación, ya que se ven obligadas a suponer y divulgar una situación de violación de la seguridad en el peor de los casos.

¿Qué debemos hacer?

- **Ampliar la retención y el alcance de los registros:** Dado el prolongado tiempo de permanencia de ciertos agentes de amenazas sofisticados, las políticas de retención estándar de 90 días son insuficientes para evaluar el alcance de las intrusiones. Las organizaciones deben priorizar el envío de registros a un almacenamiento centralizado a largo plazo, capturando específicamente datos de puntos ciegos como dispositivos perimetrales de red, hipervisores y protocolos de autenticación que las herramientas de EDR estándar suelen pasar por alto.
- **Garantizar la integridad de los registros:** Para contrarrestar las técnicas antiforenses, como la limpieza de registros locales y la alteración de marcas de tiempo, prioriza el envío inmediato de los registros a una ubicación centralizada. Esto garantiza que las pruebas se conserven incluso si se hackea el organismo receptor. Además, configura los repositorios para que generen alertas en caso de una extinción inesperada de los registros, lo que permite a los equipos de operaciones identificar cuándo un atacante está inhabilitando intencionalmente los controles de seguridad para ocultar sus huellas.
- **Cambiar a la búsqueda proactiva:** Los equipos de seguridad deberían ir más allá de las alertas reactivas e implementar una rutina de búsqueda proactiva de amenazas. Esto implica el uso de técnicas de análisis avanzadas, como la clasificación por pilas de datos, para identificar valores atípicos en el comportamiento autorizado. La integración de la inteligencia contra amenazas más reciente permite a los equipos distinguir la actividad administrativa benigna de la de los adversarios que utilizan herramientas nativas para ocultarse a plena vista.
- **Validar activos y supuestos:** Los líderes deben fomentar la colaboración entre los equipos de seguridad y de infraestructura para auditar los entornos. Realiza pruebas periódicas para comprobar las suposiciones y asegurarte de que los activos fundamentales están generando realmente la telemetría esperada. Este proceso también debería utilizarse con el objetivo de identificar y quitar tecnologías y sistemas en desuso para reducir la carga administrativa y la superficie de ataque.

Enfoque de los adversarios en la infraestructura de virtualización

Conclusión

Las plataformas de virtualización pasaron de ser infraestructura de backend a estar en la primera línea. Los atacantes explotan la naturaleza de "nivel 0" de los hipervisores para eludir las defensas a nivel de sistema invitado, incorporar una persistencia profunda que sobreviva a la corrección estándar e implementar ransomware a un nivel que imposibilita la recuperación tradicional. Para proteger esta pila, es necesario tratar el plano de administración como un activo fundamental aislado y eliminar los puntos ciegos graves en el registro de eventos para restablecer la visibilidad.

¿Qué debo saber?

- **Los hipervisores representan un punto ciego en materia de seguridad:** Los hipervisores suelen ejecutar sistemas operativos de propiedad que son incompatibles con las herramientas de EDR estándar, lo que crea una brecha de visibilidad que permite a los atacantes implementar malware o crear máquinas virtuales no administradas para llevar a cabo ataques sin que se activen las alertas de EDR o SIEM.
- **Los atacantes eluden las defensas de las máquinas huésped:** Los adversarios más sofisticados ya no necesitan acceder a los sistemas objetivo para robar datos. Atacan directamente la capa de almacenamiento de virtualización, por lo que pueden clonar discos virtuales y extraer bases de datos sensibles de Active Directory sin interactuar en ningún momento con el sistema operativo huésped ni con sus controles de seguridad.
- **El ransomware ahora ataca los almacenes de datos:** Las campañas modernas de ransomware están cambiando de estrategia, ya que encriptan los almacenes de datos de los hipervisores en lugar de las máquinas individuales. Esta táctica permite a un agente de amenazas apagar y bloquear simultáneamente todos los servidores de una máquina anfitrión, lo que deja inoperables e irrecuperables todas las máquinas virtuales asociadas.
- **Los atacantes están incorporando la persistencia profunda:** Los adversarios atacan cada vez más la capa subyacente del hipervisor para implementar puertas traseras ocultas y máquinas virtuales maliciosas. Debido a que estos mecanismos se ejecutan por debajo del entorno estándar del servidor, suelen sobrevivir a los esfuerzos habituales de corrección de incidentes y a los reinicios del sistema, lo que otorga a los atacantes acceso permanente a través de una puerta trasera.

¿Qué debemos hacer?

- **Separar la identidad de la administración:** Quita los hipervisores y la infraestructura de copia de seguridad del dominio de Active Directory corporativo. Depender del mismo proveedor de identidad tanto para la producción como para la infraestructura crea un punto único de fallo. Implementa un proveedor de identidad (IdP) dedicado exclusivamente a la infraestructura o aplica el acceso justo a tiempo (JIT).
- **Aislar el plano de administración:** Trata las interfaces de virtualización como recursos de nivel 0. Restringe el tráfico de administración a un segmento de red dedicado y protegido por un firewall, accesible únicamente a través de estaciones de trabajo con acceso privilegiado y endurecidas, e implementa la MFA resistente al phishing.
- **Fomentar una resiliencia inmutable:** Para contrarrestar la destrucción de las capacidades de recuperación, los entornos de copia de seguridad deben estar aislados y utilizar almacenamiento inmutable. Se deben realizar pruebas de restauración periódicas a partir de estas copias aisladas para verificar que las copias de seguridad funcionen según lo previsto durante un evento de recuperación de alta presión.
- **Centralizar la telemetría de la infraestructura:** Exige el reenvío de los registros de administración de virtualización y de nivel de hipervisor a tu SIEM central. Dado que los atacantes eluden los sistemas operativos huésped, la telemetría del hipervisor es ahora la única forma de detectar el acceso no autorizado, la manipulación de instantáneas o la creación de máquinas virtuales maliciosas.

Explotación sistemática de los dispositivos de red perimetrales y centrales

Conclusión

Los atacantes utilizan los dispositivos de red perimetrales y centrales como armas para eludir las herramientas de seguridad modernas, explotando las vulnerabilidades más rápido de lo que se publican los parches y abusando de las funciones nativas de los dispositivos para robar datos silenciosamente. Debido a que estas puertas de enlace fundamentales a menudo no están catalogadas ni supervisadas, ofrecen a los adversarios un acceso invisible y a largo plazo. Las organizaciones deben priorizar con urgencia el descubrimiento exhaustivo de activos, la administración estricta de parches y el registro centralizado para recuperar el control del perímetro de su red.

¿Qué debo saber?

- **La explotación de vulnerabilidades de día cero se está acelerando:** El tiempo promedio para aprovechar las vulnerabilidades se redujo drásticamente, lo que significa que los agentes de amenazas están comprometiendo cada vez más los sistemas, incluidos los dispositivos de red perimetrales y centrales, incluso antes de que los proveedores publiquen un parche.
- **Los dispositivos de red son puntos ciegos de seguridad:** Dado que los dispositivos perimetrales no pueden ejecutar software de EDR tradicional, los atacantes los utilizan como refugios seguros. Implementan malware personalizado que se ejecuta en memoria, el cual es capaz de evadir la detección y sobrevivir en entornos con telemetría deficiente.
- **Los adversarios eluden los extremos:** Los atacantes avanzados llevan a cabo casi todo el ciclo de vida del ataque, desde el reconocimiento hasta el robo de datos, directamente desde la infraestructura de red, eludiendo por completo las estaciones de trabajo y los servidores que están bien supervisados.
- **La respuesta ante incidentes está muy limitada:** Los dispositivos perimetrales tienen requisitos mínimos de almacenamiento y de tiempo de actividad estrictos. Cuando se produce una vulneración de seguridad, suele ser imposible realizar un análisis forense estándar del sistema de archivos, lo que dificulta confirmar la presencia de un atacante antes de que se pierdan las pruebas.
- **Los atacantes utilizan las funciones nativas como armas:** Los adversarios utilizan cada vez más subshells administrativos integrados y la funcionalidad nativa de captura de paquetes para recopilar copias del tráfico en tiempo real entre los dispositivos. Esto les permite extraer contraseñas de protocolos de red en texto simple y operar sin llamar la atención, sin necesidad de implementar malware detectable.

¿Qué debemos hacer?

- **Centralizar y conservar los registros de red:** No confíes únicamente en los datos de extremos. Envía los registros fundamentales de los dispositivos de red, especialmente los registros de aplicaciones y de administración, a una SIEM centralizada, y conserva los registros administrativos durante al menos un año para garantizar la visibilidad durante una investigación.
- **Implementar una gestión estricta de vulnerabilidades:** Integra los dispositivos de red en programas integrales de análisis de vulnerabilidades y administración de activos. Establece responsabilidades claras y calendarios de aplicación de parches escalonados para corregir las vulnerabilidades antes de que los atacantes puedan explotarlas sin interrumpir las operaciones comerciales.
- **Desarrollar guías de actuación ante incidentes específicos para cada red:** Los equipos de seguridad deben desarrollar diagramas arquitectónicos detallados y guías de respuesta. Planifica con antelación los procedimientos de recolección forense para garantizar que las pruebas volátiles fundamentales no se destruyan durante un reinicio o un ciclo de encendido/apagado estándar.
- **Identificar infraestructura de red no catalogada:** Dado que los dispositivos perimetrales se implementan con frecuencia y, luego, se olvidan, los equipos de seguridad deben utilizar análisis de descubrimiento exhaustivos para identificar los sistemas no catalogados. Las organizaciones no pueden defender un activo cuya existencia desconocen, aplicarle parches ni supervisarlo.

El efecto dominó de las vulneraciones de seguridad de los servicios SaaS externos

Conclusión

El cambio hacia una infraestructura centrada en la nube transformó las aplicaciones de SaaS en vías de acceso para ataques masivos a la cadena de suministro. Los agentes de amenazas eluden las defensas estándar robando tokens de integración y aprovechando las apps de terceros no verificadas, lo que convierte una simple violación de seguridad de un proveedor en una crisis empresarial en cadena. Las organizaciones deberían adoptar un sistema de verificación de identidad continua, administrar estrictamente el consentimiento del usuario final para las aplicaciones y aplicar una gestión rigurosa de los riesgos de terceros antes de realizar cualquier adquisición.

¿Qué debo saber?

- **Se está eludiendo la MFA:** Los atacantes ya no se limitan a robar contraseñas; también recopilan tokens de OAuth de larga duración y cookies de sesión. Dado que estas credenciales por lo general siguen siendo válidas después de salir de la sesión, los atacantes pueden secuestrar sesiones sin que se activen las alertas de MFA.
- **Las integraciones son el nuevo perímetro:** Los agentes de amenazas comprometen a los proveedores de SaaS externos para robar claves codificadas y tokens de acceso personal, y utilizan esos secretos robados para infiltrarse sin problemas en los entornos de clientes downstream y llevar a cabo robos de datos a gran escala.
- **La ingeniería social está en aumento:** Grupos con motivaciones económicas atacan agresivamente a los departamentos de ayuda de TI a través del phishing por voz (vishing). Usan la identidad de los empleados para eludir los controles, obtener acceso inicial al SaaS y extraer claves con privilegios altos del entorno.

¿Qué debemos hacer?

- **Descubrir y administrar el ecosistema de SaaS:** La administración de activos tradicional a menudo no incorpora apps en la nube. Implementa herramientas de administración de la postura de seguridad de SaaS (SSPM) para inventariar activamente todas las aplicaciones, las integraciones y los secretos ocultos (como claves de API) y así eliminar los puntos ciegos.
- **Endurecer los controles de identidad:** Exige que todas las aplicaciones de SaaS se enruten a través de un proveedor de identidad central (IdP). Aplica un principio de privilegio mínimo estricto para las claves de API de terceros, elimina los permisos de comodín y utiliza el acceso justo a tiempo (JIT) para minimizar los privilegios existentes.
- **Automatizar la administración del ciclo de vida:** Implementa la rotación automatizada de todos los secretos y las credenciales de cuentas de servicio. Aplicar períodos de validez extremadamente cortos para los tokens de acceso y las sesiones del navegador devaluará de forma instantánea las cookies robadas que se venden en la Web oscura.
- **Restringir el consentimiento de los usuarios finales a aplicaciones:** Los administradores deben inhabilitar la capacidad de que los usuarios finales den su consentimiento a aplicaciones de terceros no verificadas. Esto garantiza que solo las aplicaciones verificadas puedan obtener tokens persistentes, lo que impide de forma eficaz que las aplicaciones de OAuth maliciosas accedan al entorno.
- **Implementar una administración estricta de los riesgos de proveedores:** Un riesgo aceptado en una aplicación de terceros supone un riesgo para el núcleo de producción. Las organizaciones deben implementar un programa sólido de administración de riesgos de terceros para examinar a los proveedores antes de su incorporación, exigiendo medidas como el inicio de sesión único, el registro de auditoría detallado y las prácticas de desarrollo seguras como parte del proceso de adquisición.

Descarga el informe completo.

Si tu organización sospecha que sufrió un incidente cibernético o experimentas una violación de la seguridad, comunícate con Mandiant para obtener asistencia para la respuesta ante incidentes.

Google Cloud

Para obtener más información, visita cloud.google.com.