MANDIANT
NOW PART OF Google Cloud

Informe de la Mirada a la Ventaja del Defensor, edición 4

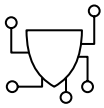
El informe de la Mirada a la Ventaja del Defensor brinda información sobre los temas de defensa cibernética de creciente importancia basada en las observaciones del frente de batalla de Mandiant y las experiencias del mundo real. Esta publicación cubre una amplia gama de temas, incluida la incorporación de seguridad en los sistemas de IA, las mejores prácticas para comunicaciones eficaces ante las crisis durante un incidente y la mitigación de los riesgos más recientes para IoT y la infraestructura de red perimetral.

Asegurar los sistemas de inteligencia artificial

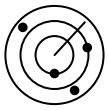
La inteligencia artificial (IA) avanza rápidamente y es importante que con ella evolucionen estrategias eficaces de gestión de riesgos. Mandiant cree que es importante incorporar seguridad en los sistemas de IA desde el principio para evitar las soluciones de seguridad integradas que hemos visto afectar a las redes y DevOps. Para ayudar a lograr esto, Google introdujo recientemente Secure AI Framework (SAIF), un marco conceptual para sistemas de IA seguros.

SAIF ofrece un enfoque práctico para abordar las preocupaciones más importantes, incluida la seguridad (por ejemplo, gestión de acceso, seguridad del endpoint y las redes, ataques a la cadena de suministro, etc.), gestión de riesgos del modelo IA/ML (por ejemplo, transparencia y responsabilidad del modelo, envenenamiento de datos, linaje de datos, etc.), privacidad y cumplimiento (por ejemplo, privacidad de datos y uso de datos confidenciales) y personas y organizaciones (por ejemplo, deficiencias en talento, gobernanza e informes para la junta directiva).

Hay seis elementos centrales de SAIF que en conjunto guían a las organizaciones para crear e implementar sistemas de IA de manera segura y responsable.



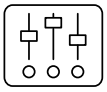
Ampliar bases de seguridad sólidas al entorno de la inteligencia artificial



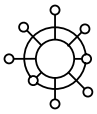
Ampliar la detección y la respuesta para llevar la inteligencia artificial al universo de amenazas de una organización

Automatizar las defensas para seguir el ritmo de las amenazas nuevas y existentes

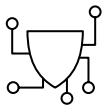
Armonizar los controles a nivel de plataforma a fin de garantizar una seguridad consistente en toda la organización



Adaptar los controles para ajustar las mitigaciones y crear bucles de retroalimentación más rápidos para la implementación de inteligencia artificial



Contextualizar los riesgos del sistema de IA en los procesos comerciales circundantes



Ampliar bases de seguridad sólidas al entorno de la inteligencia artificial

- **Revisar qué controles de seguridad existentes en todos los dominios de seguridad se aplican a los sistemas de IA**

  Los controles de seguridad existentes en todos los dominios de seguridad se aplican a los sistemas de IA de varias maneras. Por ejemplo, se pueden utilizar controles de seguridad de los datos con el fin de proteger los datos que utilizan los sistemas de IA para capacitar y operar; los controles de seguridad de las aplicaciones se pueden utilizar para proteger el software en el que se implementan los sistemas de IA; los controles de seguridad de la infraestructura se pueden utilizar para proteger la infraestructura subyacente de la que dependen los sistemas de inteligencia artificial; y se pueden utilizar controles de seguridad operativos para garantizar que los sistemas de inteligencia artificial funcionen de manera segura.

  Los controles específicos que se necesitan variarán según el uso de la inteligencia artificial, así como los sistemas y entornos de inteligencia artificial específicos.

- **Evaluar la relevancia de los controles tradicionales para las amenazas y riesgos de la inteligencia artificial utilizando los marcos disponibles**

  Los controles de seguridad tradicionales pueden ser relevantes para las amenazas y riesgos de la inteligencia artificial, pero es posible que sea necesario adaptarlos para que sean eficaces o agregar capas adicionales a la postura de defensa para ayudar a cubrir los riesgos específicos de la inteligencia artificial. Por ejemplo, el cifrado de datos puede ayudar a proteger los sistemas de IA del acceso no autorizado al limitar el acceso de las claves a determinadas funciones, pero también puede ser necesario utilizarlo para proteger los modelos de inteligencia artificial y los datos que los sustentan contra el robo o la manipulación.

- **Realizar un análisis para determinar qué controles de seguridad deben agregarse debido a amenazas, regulaciones, etc. específicas de la inteligencia artificial.**

  Junto con el equipo reunido, revise cómo sus controles actuales se corresponden con su caso de uso de inteligencia artificial, realice una evaluación adecuada de estos controles y luego cree un plan para abordar las áreas de brechas. Una vez realizado todo esto, mida también la eficacia de estos controles en función de si reducen el riesgo y qué tan bien abordan el uso previsto de la inteligencia artificial.

- **Prepararse para almacenar y realizar un seguimiento de los activos, el código y los datos de capacitación de la cadena de suministro**

  Las organizaciones que utilizan sistemas de IA deben prepararse para almacenar y rastrear los activos, el código y los datos de capacitación de la cadena de suministro. Esto incluye identificar, categorizar y proteger todos los activos, así como supervisar el acceso o uso no autorizado. Al tomar estas medidas, las organizaciones pueden ayudar a proteger sus sistemas de IA frente a los ataques.

- **Garantizar que su gobernanza de datos y gestión del ciclo de vida sean escalables y estén adaptados a la inteligencia artificial**

  Dependiendo de la definición de gobernanza de datos a la que adhiera, existen hasta seis dominios de decisión para la gobernanza de datos:

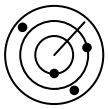
  - Calidad de los datos
  - Seguridad de los datos
  - Arquitectura de los datos
  - Metadatos
  - Ciclo de vida de los datos
  - Almacenamiento de datos

  La gobernanza de datos de inteligencia artificial será más importante que nunca. Por ejemplo, un fundamento clave de la eficacia de los modelos de inteligencia artificial son los conjuntos de datos de capacitación. Asegúrese de tener un sistema de gestión del ciclo de vida adecuado cuando se trata de conjuntos de datos, con un fuerte énfasis en la seguridad como parte del ciclo de vida (es decir, tener medidas de seguridad desde la creación de datos hasta la destrucción final de los datos integradas durante todo el ciclo de vida). El linaje de datos también desempeñará un papel clave y ayudará a responder preguntas con respecto a la privacidad y la propiedad intelectual. Si sabe quién creó los datos, de dónde provienen y qué constituye el conjunto de datos, será mucho más fácil responder preguntas sobre los temas antes mencionados.

  A medida que crece la adopción de la inteligencia artificial, el éxito de su organización probablemente dependerá de escalar estos dominios de decisión de manera ágil. Para ayudar a respaldar este esfuerzo, es fundamental revisar su estrategia de gobernanza de datos con un equipo multifuncional y ajustarla potencialmente a fin de garantizar que refleje los avances en inteligencia artificial.

- **Retener y volver a capacitar**

  No estamos hablando de inteligencia artificial, sino de personas. Para muchas organizaciones, encontrar el talento adecuado en seguridad, privacidad y cumplimiento puede implicar un recorrido de varios años. Tomar medidas para retener este talento puede contribuir a su éxito, ya que las personas pueden volver a capacitarse con habilidades relevantes para la inteligencia artificial, y eso resulta ser más rápido que contratar talento externo que tal vez tenga el conocimiento específico requerido de inteligencia artificial, pero que no tiene el conocimiento institucional, el cual puede llevar más tiempo adquirir.



Ampliar la detección y la respuesta para llevar la inteligencia artificial al universo de amenazas de una organización

- **Desarrollar la comprensión de las amenazas más importantes para los escenarios de uso de la inteligencia artificial, los tipos de inteligencia artificial utilizados, etc.**

  Las organizaciones que utilizan sistemas de IA deben comprender las amenazas relevantes para sus escenarios específicos de uso de inteligencia artificial. Esto incluye comprender los tipos de inteligencia artificial que utilizan, los datos que utilizan para entrenar sistemas de IA y las posibles consecuencias de una vulneración de seguridad. Al tomar estas medidas, las organizaciones pueden ayudar a proteger sus sistemas de IA frente a los ataques.

- **Prepararse para responder a los ataques contra la IA y también a los problemas planteados por los resultados de la IA**

  Las organizaciones que utilizan sistemas de IA deben tener un plan para detectar y responder ante incidentes de seguridad, y mitigar los riesgos de que los sistemas de IA tomen decisiones dañinas o sesgadas. Al tomar estas medidas, las organizaciones pueden ayudar a proteger sus sistemas de IA y a sus usuarios de los daños.

- **Específicamente para la IA generativa, centrarse en los resultados de la IA: prepararse para hacer cumplir las políticas de seguridad del contenido**

  La IA generativa es una herramienta poderosa para crear una variedad de contenido, desde texto hasta imágenes y videos. Sin embargo, este poder también conlleva el potencial de uso indebido. Por ejemplo, la IA generativa podría usarse para crear contenido dañino, como discursos de odio o imágenes violentas. Para mitigar estos riesgos, es importante prepararse para exigir el cumplimiento de las políticas de seguridad del contenido.

- **Ajustar la política de uso indebido y los procesos de respuesta a incidentes conforme a los incidentes específicos de la IA, como la creación de contenido malicioso o violaciones de la privacidad de la IA**

  A medida que los sistemas de IA se vuelven más complejos y generalizados, es importante ajustar su política para abordar los casos de uso indebido y también ajustar sus procesos de respuesta a incidentes a fin de tener en cuenta los tipos de incidentes específicos de la IA. Este tipo de incidentes pueden incluir creación de contenido malicioso, violaciones de la privacidad de la IA, sesgo de la IA y uso indebido general del sistema.

Automatizar las defensas para seguir el ritmo de las amenazas nuevas y existentes

- **Identificar la lista de capacidades de seguridad de IA centradas en proteger los sistemas de IA, entrenar las canalizaciones de datos, etc.**

  Las tecnologías de seguridad de IA pueden proteger los sistemas de IA de una variedad de amenazas, incluidas las vulneraciones de datos, la creación de contenido malicioso y el sesgo de la IA. Algunas de estas tecnologías incluyen cifrado de datos tradicional, control de acceso, auditoría que puede ampliarse con IA y tecnologías más nuevas que pueden realizar capacitación, protección de datos y protección de modelos.

- **Utilizar defensas de IA para contrarrestar las amenazas de la IA, pero mantener a las personas informadas para tomar decisiones cuando sea necesario**

  La IA se puede utilizar para detectar y responder ante amenazas contra la IA, como vulneraciones de datos, creación de contenido malicioso y sesgos de la IA. Sin embargo, las personas deben permanecer informadas para tomar decisiones importantes, como determinar qué constituye una amenaza y cómo responder a ella. Esto se debe a que los sistemas de IA pueden estar sesgados o cometer errores, y es necesaria la supervisión humana para garantizar que los sistemas de IA se utilicen de manera ética y responsable.

- **Utilizar la IA para automatizar tareas que consumen mucho tiempo, reducir el trabajo y acelerar los mecanismos defensivos**

  Aunque parezca un punto más bien simplista a la luz de los usos de la IA, utilizarla para acelerar tareas que consumen mucho tiempo conducirá, en última instancia, a resultados más rápidos. Por ejemplo, realizar ingeniería inversa a un binario de malware puede llevar mucho tiempo. Sin embargo, la IA puede revisar rápidamente el código relevante y brindar información procesable al analista. Con esta información, el analista podría pedirle al sistema que genere una regla YARA en busca de estas acciones. En este ejemplo, hay una reducción inmediata del esfuerzo y un rendimiento más rápido para la postura defensiva.

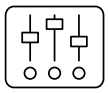
Armonizar los controles a nivel de plataforma a fin de garantizar una seguridad consistente en toda la organización

- **Revisar el uso de la IA y el ciclo de vida de las aplicaciones basadas en IA**

  Como se mencionó en el Paso 1, comprender el uso de la IA es un componente clave. Una vez que la IA se utilice más ampliamente en su organización, debe implementar un proceso de revisión periódica del uso para identificar y mitigar los riesgos de seguridad. Esto incluye revisar los tipos de modelos y aplicaciones de IA que se utilizan, los datos empleados para entrenar y ejecutar modelos de IA, las medidas de seguridad implementadas para proteger los modelos y aplicaciones de IA, los procedimientos para supervisar y responder ante incidentes de seguridad de IA, y conocimiento y capacitación sobre los riesgos de seguridad de la IA para todos los empleados.

- **Evitar la fragmentación de los controles intentando estandarizar las herramientas y los marcos**

  Después de implementar el proceso mencionado anteriormente, podrá comprender mejor las herramientas, los controles de seguridad y los marcos existentes actualmente. Al mismo tiempo, es importante examinar si su organización cuenta con marcos diferentes o superpuestos para controles de seguridad y cumplimiento con el fin de ayudar a reducir la fragmentación. La fragmentación aumentará la complejidad y creará un solapamiento considerable, lo que aumentará los costos y las ineficiencias. Al armonizar sus marcos y controles, y comprender su aplicabilidad en su contexto de uso de IA, limitará la fragmentación y proporcionará un enfoque "adecuado" para los controles para mitigar el riesgo. Esta guía se refiere principalmente a los marcos y estándares de control existentes, pero el mismo principio (por ejemplo, tratar de mantener el número total lo más pequeño posible) se aplicaría a los marcos y estándares nuevos y emergentes para la IA.



Adaptar los controles para ajustar las mitigaciones y crear bucles de retroalimentación más rápidos para la implementación de inteligencia artificial

- **Llevar a cabo ejercicios de Red Team para mejorar la seguridad de los productos y capacidades impulsados por IA**

  Los ejercicios de Red Team son un método de prueba de seguridad en el que un equipo de hackers éticos intenta explotar las vulnerabilidades en los sistemas y aplicaciones de una organización. Esto puede ayudar a las organizaciones a identificar y mitigar los riesgos de seguridad en sus sistemas de IA antes de que puedan ser explotados por actores maliciosos.

- **Mantenerse al tanto de los nuevos ataques, incluida la inyección de prompt, el envenenamiento de datos y los ataques de evasión**

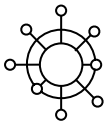
  Estos ataques pueden explotar las vulnerabilidades de los sistemas de IA para causar daños, como filtrar datos confidenciales, realizar predicciones incorrectas o interrumpir operaciones. Al mantenerse actualizadas sobre los últimos métodos de ataque, las organizaciones pueden tomar medidas para mitigar estos riesgos.

- **Aplicar técnicas de aprendizaje automático para mejorar la precisión y la velocidad de la detección**

  Si bien es fundamental centrarse en asegurar el uso de la IA, esta también puede ayudar a las organizaciones a lograr mejores resultados de seguridad a escala. Las capacidades de detección y respuesta asistidas por IA, por ejemplo, pueden ser un activo importante para cualquier organización. Al mismo tiempo, es esencial mantener a las personas informadas para supervisar sistemas, procesos y decisiones de IA relevantes. Con el tiempo, este esfuerzo puede impulsar el aprendizaje continuo para mejorar las protecciones básicas de la IA, actualizar la capacitación y ajustar los conjuntos de datos para los modelos básicos y los modelos de aprendizaje automático utilizados para crear protecciones. A su vez, esto permitirá a las organizaciones responder estratégicamente a los ataques a medida que evoluciona el entorno de amenazas. El aprendizaje continuo también es fundamental para mejorar la precisión, reducir la latencia y aumentar la eficiencia de las protecciones.

- **Crear un bucle de retroalimentación**

  Para maximizar el impacto de los tres elementos anteriores, es fundamental crear un bucle de retroalimentación. Por ejemplo, si su equipo de Red Team descubre una manera de hacer un mal uso de su sistema de IA, esa información debe enviarse a su organización para ayudar a mejorar las defensas, en lugar de enfocarse únicamente en la remediación. De manera similar, si su organización descubre un nuevo vector de ataque, debe incorporarlo a su conjunto de datos de capacitación como parte del aprendizaje continuo. Para garantizar que se aproveche la retroalimentación, es importante considerar varias vías de ingesta y comprender bien qué tan rápido se puede incorporar la retroalimentación a sus protecciones.



Contextualizar los riesgos del sistema de IA en los procesos comerciales circundantes

- **Establecer un marco modelo de gestión de riesgos y formar un equipo que comprenda los riesgos relacionados con la IA**

  Las organizaciones deben desarrollar un proceso para identificar, evaluar y mitigar los riesgos asociados con los modelos de IA. El equipo debe estar compuesto por expertos en inteligencia artificial, seguridad y gestión de riesgos.

- **Crear un inventario de modelos de IA y su perfil de riesgo en función de los casos de uso específicos y la responsabilidad compartida al aprovechar soluciones y servicios de terceros**

  Las organizaciones deben crear un inventario completo de modelos de IA y evaluar su perfil de riesgo en función de los casos de uso específicos, la sensibilidad de los datos y la responsabilidad compartida al aprovechar soluciones y servicios de terceros. Esto significa identificar todos los modelos de IA en uso, comprender los riesgos específicos asociados con cada modelo e implementar controles de seguridad para mitigar esos riesgos, además de contar con funciones y responsabilidades claras.

- **Implementar políticas, protocolos y controles de privacidad de datos, riesgo cibernético y riesgo de terceros durante todo el ciclo de vida del modelo de aprendizaje automático, a fin de guiar el desarrollo, la implementación, la supervisión y la validación del modelo**

  Las organizaciones deben implementar políticas, protocolos y controles de privacidad de datos, riesgo cibernético y riesgo de terceros durante todo el ciclo de vida del modelo de aprendizaje automático, a fin de guiar el desarrollo, la implementación, la supervisión y la validación del modelo. Esto significa desarrollar e implementar políticas, protocolos y controles que aborden los riesgos específicos asociados con cada etapa del ciclo de vida del modelo de aprendizaje automático. Tenga en cuenta el cuarto elemento del marco para asegurarse de no crear una fragmentación indebida.

- **Realizar una evaluación de riesgos que considere el uso organizacional de la IA**

  Las organizaciones deben identificar y evaluar los riesgos asociados con el uso de la IA e implementar controles de seguridad para mitigar esos riesgos. Las organizaciones también deben abarcar prácticas de seguridad para supervisar y validar la eficacia de los controles, incluida la explicabilidad de los resultados del modelo y la supervisión de desviaciones. Como se menciona en los dos primeros elementos, es importante crear un equipo multifuncional y desarrollar una comprensión más profunda de los casos de uso relevantes para apoyar este esfuerzo. Las organizaciones pueden utilizar los marcos existentes para la evaluación de riesgos con el objetivo de ayudar a guiar su trabajo, pero probablemente necesitarán aumentar o adaptar su enfoque para abordar los nuevos marcos emergentes de gestión de riesgos de la IA.

- **Incorporar la responsabilidad compartida de proteger la IA dependiendo de quién desarrolle los sistemas de IA, quién implemente los modelos desarrollados por el proveedor de modelos, quién ajuste los modelos o quién recurra a soluciones listas para usar**

  La seguridad de los sistemas de IA es una responsabilidad compartida entre los desarrolladores, implementadores y usuarios de esos sistemas. Las responsabilidades específicas de cada parte variarán según su función en el desarrollo y la implementación del sistema de IA. Por ejemplo, los desarrolladores de sistemas de IA son responsables de desarrollar sistemas de IA que sean seguros por diseño. Esto incluye el uso de prácticas de codificación segura, entrenar modelos de IA con datos limpios e implementar controles de seguridad para proteger los sistemas de IA frente a los ataques.

- **Hacer coincidir los casos de uso de IA con las tolerancias al riesgo**

  Esto significa comprender los riesgos específicos asociados con cada caso de uso de IA e implementar medidas de seguridad para mitigar esos riesgos. Por ejemplo, los sistemas de IA que se utilizan para ayudar a tomar decisiones que podrían afectar de manera considerable la vida de las personas, como la atención médica o las finanzas, probablemente necesitarán estar más protegidos que los sistemas de IA que se usan para tareas menos urgentes, como el marketing o el servicio al cliente.

En conclusión

La IA ha capturado la imaginación del mundo y muchas organizaciones están viendo oportunidades para impulsar la creatividad y mejorar la productividad aprovechando esta tecnología emergente. SAIF está diseñado para ayudar a elevar el nivel de seguridad y reducir el riesgo general al desarrollar e implementar sistemas de IA.

Cuatro fases de las comunicaciones ante las crisis de ciberseguridad

Comunicarse durante una crisis es difícil incluso para las organizaciones más inteligentes y mejor preparadas. Los atributos únicos de un ataque cibernético y el uso cada vez mayor del dominio público por parte de los actores significan que la forma en que una organización víctima se comunica con sus partes interesadas durante un incidente puede afectar su marca mucho después de que se complete la remediación técnica.

Para complicar aún más la respuesta, el proceso de gestión de las comunicaciones puede competir por el tiempo y la atención de los equipos de respuesta ante las crisis y los líderes ejecutivos mientras la organización trabaja para restaurar rápidamente las operaciones comerciales y corregir las redes durante un incidente cibernético. Además, a menudo, existe confusión sobre el alcance de las comunicaciones ante las crisis de ciberseguridad. Si bien las relaciones con los medios son una parte muy visible de la respuesta estratégica de comunicación, son solo una audiencia. En la práctica, las organizaciones deben desarrollar una estrategia de comunicación integral que informe a todas sus partes interesadas internas y externas.

Para evitar errores en la comunicación, especialmente en un momento estresante en el que cada segundo cuenta, es útil contar con expertos experimentados en comunicaciones ante las crisis de ciberseguridad que brinden asesoramiento y experiencia para ayudar a las organizaciones y sus juntas directivas a responder de manera adecuada. A medida que el panorama de amenazas evoluciona y los actores incorporan nuevas técnicas, Mandiant ahora ofrece especialistas en comunicaciones ante las crisis de ciberseguridad, junto con su equipo de respuesta a incidentes, para ayudar a los clientes a afrontar incidentes, evaluar la participación de las partes interesadas y elaborar estrategias para las comunicaciones en cascada asociadas.

¿Qué son las comunicaciones ante las crisis de ciberseguridad?

Las comunicaciones ante las crisis específicas de ciberseguridad son una combinación de respuesta a incidentes y operaciones de gestión de crisis, donde se desarrollan mensajes personalizados para distintas partes interesadas y canales, con una entrega en tiempos complejos. Durante una crisis, la estrategia de comunicación debe considerar varios factores más allá del impacto en las operaciones comerciales, el apetito de riesgo y el potencial de daños a la reputación o la marca. Por ejemplo, el comportamiento de los actores y las tendencias de inteligencia son consideraciones importantes a la hora de decidir cómo, qué y cuándo comunicarse. A veces, una "falta de respuesta estratégica" es la mejor respuesta, ya que ciertos mensajes pueden alertar al actor y obligarlo a cambiar sus tácticas, técnicas y procedimientos.

La confianza y la resiliencia de la marca se ponen a prueba especialmente durante un incidente cibernético, y en medio de un incidente no es el momento de empezar a generar confianza. Más bien, el enfoque de una organización respecto de las comunicaciones ante las crisis y la falta de información y transparencia pueden socavar aún más la confianza. Los errores de comunicación agravan la pérdida general y el impacto en las operaciones comerciales. Por lo tanto, es imperativo tener una comprensión sólida de qué son las comunicaciones ante las crisis, las mejores prácticas para responder ante tiempos de crisis y cómo prepararse y planificar la jornada.

¿Qué conduce a la mejor respuesta en comunicaciones ante las crisis de ciberseguridad?

Según la experiencia de Mandiant, el éxito comienza mucho antes del primer día de una vulneración o incidente. Más bien, es un ciclo continuo de revisión, análisis y perfeccionamiento de los planes de respuesta a incidentes y continuidad del negocio de la organización. Con atención específica a las comunicaciones ante las crisis, compartiremos las lecciones aprendidas de la experiencia de primera mano de nuestros especialistas al abordar la planificación de comunicaciones ante las crisis de ciberseguridad. El ciclo de estas actividades se agrupa en cuatro fases: preparación estratégica, seguridad, respuesta y revisión tras el incidente.

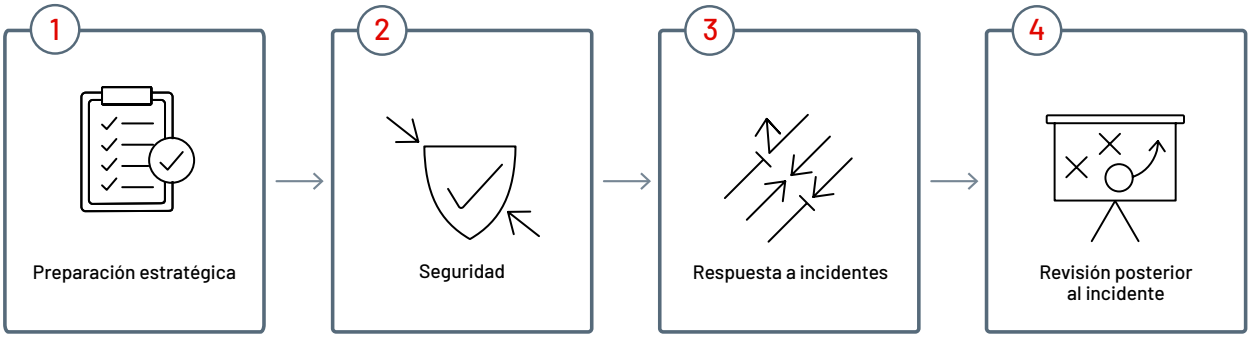
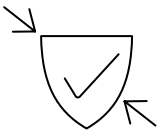


FIGURA 1: Fases de las comunicaciones ante las crisis de ciberseguridad

Fase 1: Preparación estratégica

La primera etapa del ciclo es la fase "Preparación estratégica" previa a la vulneración o, en pocas palabras, la fase de planificación. Esta fase es una actividad tradicional y esencial para todas las organizaciones, independientemente de su tamaño, sector o ubicación. El enfoque debe adaptarse a la organización, proporcionando un plan escrito y repetible con funciones y responsabilidades claramente definidas, una estructura de gobernanza con niveles formales de autoridad para tomar decisiones y un marco de respuesta. Como en el caso de los entrenadores deportivos, para el equipo de respuesta se trata de un manual de tácticas basado en actividades potenciales. Esto también debe considerarse y servir como un documento vivo que se revisa y comparte periódicamente con aquellas personas que formarán parte del equipo de respuesta durante un incidente.

Es importante contar con el equipo adecuado, con sus funciones y responsabilidades claramente definidas. En este equipo deben participar representantes de toda la organización (incluidos los Departamentos de RR. HH., Adquisiciones, Comunicaciones, Legal, Logística y Operaciones, por nombrar algunos). No se puede anticipar lo que necesitará, especialmente cuando se trata de aprovisionar hardware, implementar comunicaciones en cascada y realizar evaluaciones detalladas del impacto de los datos. El equipo también debe implementar un modelo de gobernanza y gestión, con grupos de trabajo específicos alineados con las responsabilidades funcionales. Uno de los entregables desarrollados durante la fase de planificación es un anexo de comunicaciones ante las crisis incorporado en el manual de tácticas de respuesta a incidentes. Este manual de tácticas debe ser específico para la organización e incluir secciones sobre respuesta a incidentes y crisis, mensajes clave basados en escenarios hipotéticos e identificación de partes interesadas y asignación de canales. Una consideración adicional es la importancia de contar con mecanismos de comunicación alternativos, comúnmente conocidos como comunicaciones "fuera de banda", en caso de que su forma principal de comunicación se vea comprometida. En incidentes de ciberseguridad y vulneración de datos, un actor puede tener persistencia en la red, lo que requiere que los líderes y los equipos de respuesta utilicen estos métodos de comunicación alternativos.

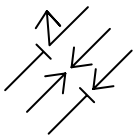


Fase 2: Seguridad

La segunda fase, que también forma parte de la respuesta proactiva y previa a la vulneración, es la fase "Seguridad" o ejercicio. Durante esta fase, las empresas deben ejercitar la respuesta de su equipo basándose en ataques y escenarios del mundo real. Algunos países incluso toman medidas para exigir esto como parte de la respuesta de la junta[1]. Sin duda, ayuda contar con especialistas bien capacitados para desarrollar y facilitar ejercicios basados en escenarios realistas y personalizados. Durante estos ejercicios, el equipo puede practicar su plan, probar su manual de tácticas e identificar brechas a remediar. Los miembros del equipo también desarrollan la memoria muscular en un entorno seguro y menos estresante. Cuando llega el momento del juego, las consecuencias de cometer un error son más importantes y más probables en una situación de mayor presión. Es mucho más fácil mantener la calma y responder con lucidez cuando puede anticipar lo que sigue en su entrega y ejecución esperadas.

También es imperativo, como parte de la fase de seguridad, que los equipos sean receptivos al asesoramiento y las sugerencias. Esta fase también debe ser una actividad recurrente, y no un ejercicio de "marque la casilla", con personas de toda la organización, en diversas funciones y niveles de trabajo, mucho más allá del equipo de liderazgo ejecutivo y la junta directiva. Por último, el ejercicio debe incluir el "refuerzo" o equipo de refuerzo, y este debe ser una gran reserva de talento. La planificación del equipo de respuesta debe tener en cuenta esfuerzos sostenidos que abarquen al menos los primeros 30 días, con turnos de personal. La respuesta inicial probablemente requerirá cobertura las 24 horas del día, los 7 días de la semana y, para evitar el cansancio y el agotamiento, es útil tener una lista con relevos capacitados y preparados para la respuesta. Es importante asegurarse de que su organización tenga un anexo o sección de comunicaciones en el plan de continuidad del negocio, el plan de recuperación ante desastres y el plan de respuesta a incidentes de la organización.

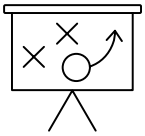
1. Departamento de Servicios Financieros de Nueva York, "DFS SUPERINTENDENT ADRIENNE A. HARRIS ANNOUNCES UPDATED CYBERSECURITY REGULATION", Noviembre de 2022



Fase 3: Respuesta a incidentes

La tercera es la fase reactiva de "Respuesta a incidentes". La ejecución de la respuesta estará definida por la prioridad y atención que usted ponga en las dos primeras fases. El dicho de que debe dedicar el 80 % de su tiempo a la planificación y el 20 % a la ejecución es verdaderamente cierto. Cuando llegue el día, es imperativo que las empresas puedan preparar rápidamente a sus equipos para responder. Conocerán sus funciones y responsabilidades y tendrán una estructura de gobernanza funcional para responder. Podrán organizar las sesiones de intercambio de información necesarias y realizar un seguimiento de las acciones y tareas. Ya habrán asignado sus partes interesadas y canales de comunicación y podrán evaluar rápidamente la preparación de los canales.

Los equipos de respuesta más fluidos y eficaces suelen ser aquellos que están bien capacitados, bien equipados y que han preparado las herramientas necesarias con antelación. Responden con dignidad, respeto por el equipo y consideración por el ritmo, reconociendo que es un maratón, no una carrera de velocidad. Colaboran estrechamente y comparten información previo a la toma de decisiones, pero tampoco se paralizan realizando análisis. Las organizaciones que fracasan o tropiezan suelen ser aquellas que no están abiertas a recibir asesoramiento o sugerencias, no reconocen sus fallas en el rendimiento o están mal organizadas y coordinadas en sus respuestas y comunicaciones.



Fase 4: Revisión posterior al incidente

Gestionar una vulneración es difícil, tanto desde un punto de vista emocional como operativo, y muchas personas no desean volver a hablar del incidente nunca más. Sin embargo, por más difícil que sea, es importante pasar a la fase final, la evaluación "post-mortem". Esta fase comienza justo cuando todo se calma: la investigación está completa, las actividades de remediación restablecen las operaciones comerciales y se han enviado notificaciones a los reguladores o a las víctimas. Algunos también pueden llamar a esto la fase "posterior a la acción" o "de lecciones aprendidas" y, después de la planificación, es una de las fases más importantes en las que hay que ser minucioso. Los especialistas pueden trabajar junto a los clientes para identificar deficiencias y soluciones que mitiguen el impacto de futuros incidentes.

Algunas de las mejores prácticas obtenidas de los casos de clientes de Mandiant surgieron durante las evaluaciones "post-mortem". Cada incidente y cada respuesta es diferente: algunos tienen salidas en falso y se recuperan bien; otros son ejemplos brillantes de las mejores prácticas de la industria. Lo importante es compartir las lecciones aprendidas en beneficio de los demás. Como afirmó Winston Churchill: "Aquellos que no aprenden de la historia están condenados a repetirla".

IoT empresarial atacado por grupos de espionaje

Se espera que el número de dispositivos activos conectados al Internet de las cosas (IoT) alcance casi 42 000 millones en 2023[2], lo que ayudará a acelerar la innovación y la automatización en sectores como la fabricación inteligente, la gestión de inventario minorista, los pagos digitales y la seguridad y vigilancia física. Como ocurre con casi todos los avances tecnológicos, el riesgo cibernético es un efecto secundario que toda empresa debe esperar.

En el pasado, Mandiant ha observado dispositivos IoT, dispositivos inteligentes y enrutadores atacados y utilizados para crear botnets y llevar a cabo operaciones de delitos cibernéticos a gran escala con motivación financiera. Una botnet es una red de dispositivos comprometidos que un actor puede utilizar para llevar a cabo una variedad de actividades de amenazas, como ataques de denegación de servicio distribuido (DDoS) y distribución de malware. Sin embargo, Mandiant evalúa con confianza moderada que los grupos de espionaje vinculados a países también han aprovechado las botnets para múltiples propósitos[3]. Este comportamiento de los atacantes subraya la oportunidad que presenta la adopción a gran escala de IoT y dispositivos inteligentes para los actores vinculados a países que buscan adquirir información estratégica y propiedad intelectual de empresas globales.

Se alienta a las organizaciones que buscan continuar su transformación digital, acelerar la automatización, recuperar las cadenas de valor perdidas después de los impactos económicos de la pandemia de COVID-19 o aprovechar el despliegue de redes de conectividad 5G[4] a trabajar en estrecha colaboración con sus equipos de ciberseguridad para garantizar que exista un plan integral de defensa cibernética y ayudar a proteger la organización.

Botnets de dispositivos IoT, dispositivos inteligentes y enrutadores útiles para ofuscar actividades

Mandiant evalúa que los grupos de espionaje vinculados a países utilizan botnets que consisten en IoT, dispositivos inteligentes y enrutadores para ofuscar actividades maliciosas, basándose en múltiples observaciones de campaña de Mandiant y otros investigadores de seguridad del sector público y privado. Los casos informados de uso de botnets de dispositivos comprometidos por parte de grupos de espionaje incluyen los siguientes.

- En abril de 2022, Mandiant informó[5] sobre una campaña de APT29 que utilizaba una botnet de cámaras IoT como parte de actividades de comando y control (C2) utilizando el malware QUIETEXIT (Figura 2). Los dominios utilizados en esta actividad C2 parecían diseñados para mezclarse con el tráfico legítimo de los dispositivos IoT infectados, aparentemente para ocultar la actividad a cualquiera que revisara los registros.

2. Frost and Sullivan, Internet of Things (IoT) Predictions Outlook, noviembre de 2022
3. Mandiant, Espionage Actors Lurk in Compromised Device Botnets, abril de 2023
4. Frost and Sullivan, The Top Growth Opportunities for IoT in 2023, marzo de 2023
5. Mandiant, https://www.mandiant.com/resources/blog/unc3524-eye-spy-email

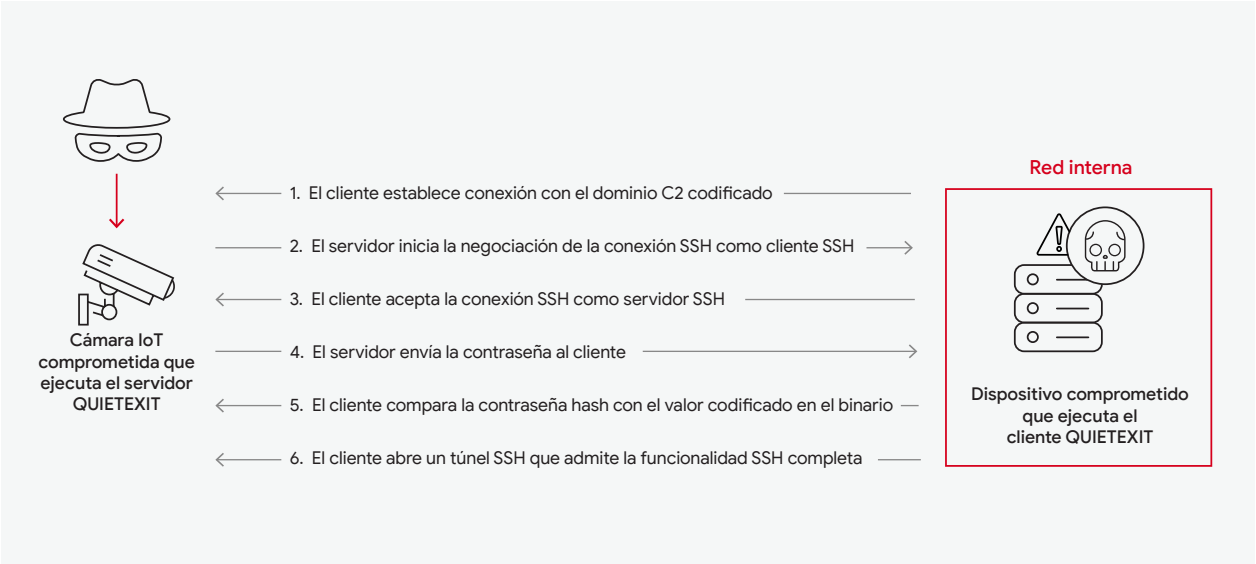


FIGURA 2: Cómo funciona QUIETEXIT con dispositivos IoT

- Un informe de 2021[6] de la Agence nationale de la sécurité des systèmes d'information (ANSSI, Agencia Nacional Francesa para la Seguridad de los Sistemas de Información) de Francia detalló una campaña vinculada al grupo chino APT31 que supuestamente utilizó una botnet de enrutadores y posiblemente otros dispositivos pequeños de oficinas y hogares para ofuscar actividades dentro de redes específicas.
- En 2022, PricewaterhouseCoopers informó[7] sobre el malware observado durante un encuentro al que denominaron "BPFDor", que Mandiant vinculó a APT41. En la campaña informada, el malware supuestamente recibió comandos de servidores privados virtuales (VPS) que estaban controlados por una red de enrutadores comprometidos con sede en Taiwán.
- La empresa de seguridad china Antiy informó[8] en 2022 que había observado una gran red de dispositivos IoT y dispositivos Linux comprometidos que enrutaban el tráfico entre servidores C2 y el malware Torii. Según la firma, pudieron atribuir la actividad a OceanLotus, denominado por Mandiant como APT32; sin embargo, Mandiant no ha confirmado esta atribución.
- En 2018, los investigadores informaron públicamente[9] del uso del malware VPNFILTER en campañas que atacaban a dispositivos de red y dispositivos de almacenamiento conectado a la red (network attached storage, NAS) a nivel mundial, con una gran concentración de dispositivos en Ucrania. Según se informa, algunas muestras integraban capacidades destructivas y de adversarios en el medio (adversary-in-the-middle, aitm), pero es posible que estos módulos estuvieran destinados a otros fines. Mandiant cree que este uso de VPNFILTER es consistente con una actividad de espionaje cibernético patrocinada por Rusia.

6. https://www.cert.ssi.gouv.fr/cti/CERTFR-2021-CTI-013/
7. https://www.pwc.com/gx/en/issues/cybersecurity/cyber-threat-intelligence/cyber-year-in-retrospect/yir-cyber-threats-report-download.pdf
8. https://mp.weixin.qq.com/s/2RIuW4O56UWiNSQB2hQtGA
9. https://blog.talosintelligence.com/vpnfilter/
10. https://thehackernews.com/2018/06/vpnfilter-router-malware.html

Los informes públicos y las observaciones de Mandiant indican que algunos actores han comprometido o utilizado botnets existentes creadas por otros actores. Mandiant sospecha que esta táctica es útil para los actores de espionaje en circunstancias muy limitadas y, por lo tanto, su uso no aumentará considerablemente en el futuro.

- En septiembre de 2022, Mandiant identificó[11] una campaña de UNC4210, que se sospecha que está vinculada al Turla Team, en la que los actores secuestraron al menos tres dominios C2 asociados con una botnet de malware ANDROMEDA. La versión de ANDROMEDA asociada con la botnet se cargó por primera vez en VirusTotal en 2013 y se propagó desde memorias USB infectadas. Después de volver a registrar los dominios C2 caducados, Turla aparentemente pudo utilizar las infecciones restantes que contactaron con los servidores para perfilar y seleccionar a las víctimas (Figura 3).

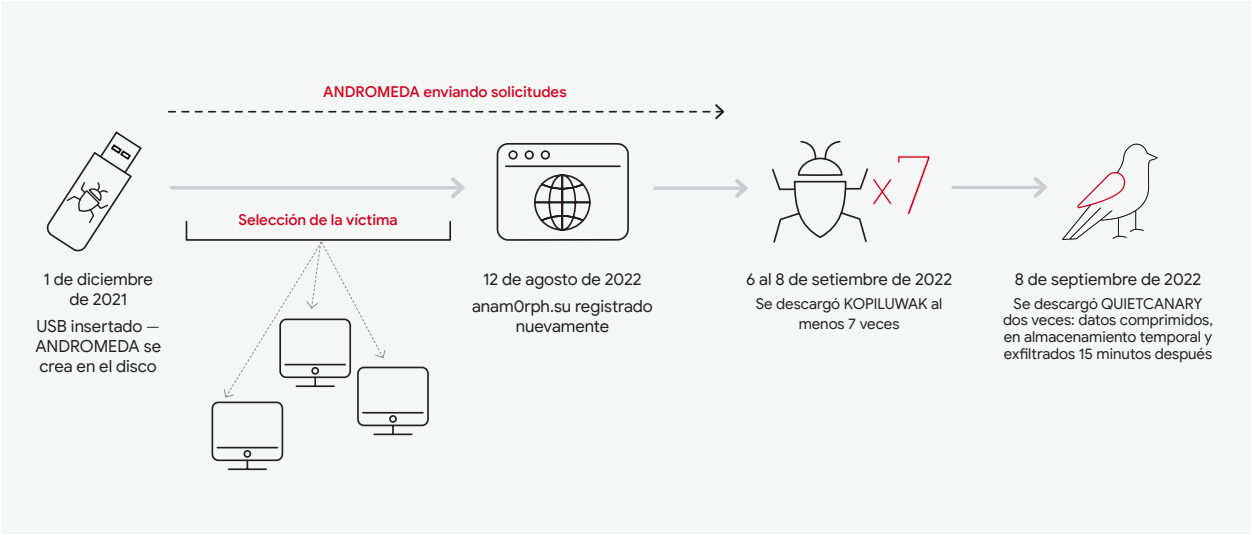


FIGURA 3: Cronología de ANDROMEDA en la intrusión de Turla

11. https://www.mandiant.com/resources/blog/turla-galaxy-opportunity

Seguridad de dispositivos IoT

Los dispositivos inteligentes y de IoT a menudo no están diseñados para ser seguros y, en ocasiones, tienen credenciales codificadas o son difíciles o imposibles de aplicar parches cuando se descubren vulnerabilidades de software. Las organizaciones que implementan activamente estos dispositivos, o que incluyen IoT en los planes de transformación digital, deben asegurarse de que puedan protegerse adecuadamente y verificarse periódicamente para detectar actividades sospechosas. En la figura 4, se describen los riesgos de seguridad relacionados con la fabricación y el funcionamiento de los dispositivos IoT que los propietarios de activos deben tener en cuenta junto con los planes para implementar estos dispositivos.

FIGURA 4: Consideraciones para proteger dispositivos IoT

Qué significa esto para las organizaciones dentro de una transformación digital

Mandiant prevé que los actores del espionaje cibernético seguirán utilizando esta táctica porque proporciona a los atacantes una ventaja táctica eficaz por una inversión relativamente baja de tiempo y recursos, ya que los dispositivos IoT e inteligentes suelen estar mal protegidos y siguen proliferando. Mandiant también especula con la posibilidad de que, a medida que aumente la popularidad del IoT y de los dispositivos inteligentes y que las herramientas dirigidas específicamente a estos dispositivos estén más disponibles en los mercados clandestinos y libremente en línea, los actores del espionaje muestren un mayor interés en el uso de botnets como medio para disfrazar la actividad de recopilación de inteligencia como delincuencia cibernética benigna u oportunista con motivaciones financieras.

Establecer resiliencia contra ataques a dispositivos perimetrales

En los últimos 10 años, las organizaciones han aumentado la visibilidad en todos sus entornos digitales. Como resultado, están detectando a los atacantes con mayor rapidez[12] y han logrado avances considerables en la protección proactiva de sus entornos frente a amenazas como la reutilización de contraseñas y los ataques de fuerza bruta, a medida que siguen avanzando hacia una arquitectura de estilo de defensa en profundidad.

Si bien esta tendencia avanza en la dirección correcta, la mayoría de las organizaciones centran la detección y la respuesta en torno a la visibilidad proporcionada por sus soluciones de detección y respuesta del endpoint (EDR). Sin embargo, las soluciones EDR se implementan, como su nombre indica, en endpoints. En otras palabras, los firewalls, los dispositivos IoT, las VPN, los hipervisores y muchos otros dispositivos no suelen ser compatibles con EDR y, por lo tanto, comúnmente se los denomina "dispositivos perimetrales". ¿Qué sucede cuando los actores maliciosos comienzan a atacar esos dispositivos?

Debido a que los dispositivos perimetrales, por definición, se encuentran fuera del rango de detección típico de la mayoría de las organizaciones, brindan a los atacantes un valor enorme durante las intrusiones. Los dispositivos perimetrales siempre serán un objetivo para los adversarios, pero de maneras diferentes. Estos dispositivos perimetrales prestan muchos servicios valiosos a las organizaciones, como la supervisión de las herramientas de seguridad internas, pero históricamente no han sido compatibles con las soluciones EDR y rara vez se monitorean a nivel de sistema. Este tipo de supervisión a nivel del sistema es necesario para identificar si se han instalado cambios en el código o malware específico.

Los dispositivos perimetrales se utilizan para la búsqueda y protección de seguridad y no se protegen inherentemente a sí mismos. Más concretamente, los proveedores normalmente no permiten a los usuarios el acceso directo al sistema operativo o al sistema de archivos. Debido a que las detecciones no se extienden a estos sistemas y dispositivos perimetrales, los defensores tienen una capacidad limitada para realizar análisis del comportamiento subyacente y potencialmente anómalo.

12. M-Trends 2023, Mandiant abril de 2023

En los últimos cinco años, Mandiant ha visto cada vez más evidencia que sugiere que adversarios con respaldo de naciones están atacando a dispositivos perimetrales. Este enfoque en los dispositivos perimetrales es tan preocupante para los defensores como ventajoso para los atacantes. Las intrusiones maliciosas atacan a dispositivos perimetrales que pueden establecer una presencia o persistir en el entorno objetivo. Más allá de una simple presencia, los dispositivos perimetrales ofrecen una serie de ventajas a los actores maliciosos. La primera de ellas es que los dispositivos perimetrales cuentan con visibilidad y privilegios elevados dentro del entorno para proporcionar supervisión de red o un punto de acceso seguro. El acceso a estos dispositivos también permite al atacante controlar el momento de la operación y puede reducir las posibilidades de detección. Los dispositivos perimetrales, por definición, no son visibles para las soluciones EDR, lo que significa que todas estas ventajas se confieren a los ataques, así como la capacidad de permanecer ocultos a los defensores.

Los adversarios con respaldo de naciones a menudo dedican tiempo y esfuerzo considerables a ciclos extensos de investigación y desarrollo para identificar y crear exploits para vulnerabilidades previamente desconocidas. Mandiant ha investigado docenas de intrusiones a lo largo de los años en las que presuntos grupos con nexo chino explotaron vulnerabilidades de día cero e implementaron malware personalizado para robar credenciales de usuario y mantener el acceso a largo plazo a los entornos de las víctimas. Por ejemplo, en 2022, UNC3886 atacó dispositivos perimetrales, como firewalls y, más adelante en el ciclo de vida del ataque, tecnologías de hipervisores.

Caso práctico de UNC3886

UNC3886 atacó múltiples componentes[13] del ecosistema de Fortinet antes de que se desplazara lateralmente a la infraestructura de VMWare. Estos componentes y sus versiones asociadas, en el momento del ataque, se incluyen a continuación:

- **FortiGate: 6.2.7** – Las unidades FortiGate son dispositivos de firewall de red que permiten el control y la supervisión del tráfico de red que pasa a través de los dispositivos.
- **FortiManager 6.4.7** – FortiManager actúa como una plataforma de gestión centralizada para gestionar dispositivos Fortinet.
- **FortiAnalyzer 6.4.7** – FortiAnalyzer actúa como una solución de gestión de registros centralizada para dispositivos Fortinet, así como una plataforma de informes.

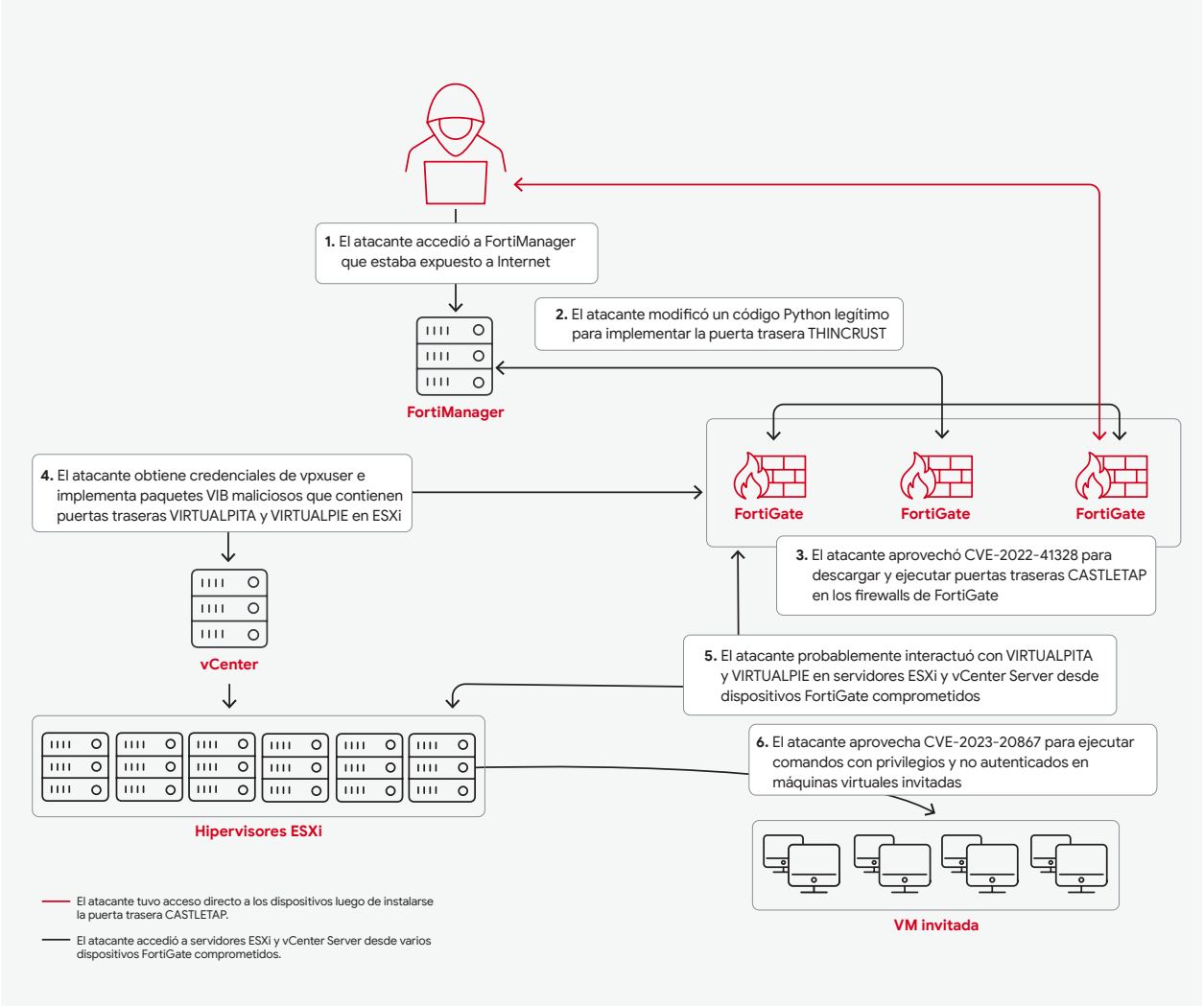


FIGURA 5: Actividad luego de implementarse restricciones de acceso a Internet en FortiManager

13. https://www.mandiant.com/resources/blog/fortinet-malware-ecosystem

En 2022, Mandiant comenzó a rastrear a UNC3886, un grupo con un presunto nexo chino. Este grupo atacó específicamente en el ecosistema de Fortinet y, finalmente, se desplazó lateralmente para acceder a la infraestructura de VMWare en entornos objetivo. Para obtener este acceso, UNC3886 demostró tener conocimiento suficiente de varias soluciones de Fortinet, incluidas FortiGate (firewall), FortiManager (solución de gestión centralizada) y FortiAnalyzer (plataforma de informes, análisis y gestión de registros). Con este conocimiento, UNC3886 implementó una puerta trasera, rastreada por Mandiant como **THINCRUST**, en los dispositivos FortiManager y FortiAnalyzer para ganar persistencia. Luego, UNC3886 aprovechó el acceso a los scripts nativos de FortiManager para explotar CVE-2022-41328 a fin de descargar y ejecutar otra puerta trasera, **CASTLETAP**, en dispositivos FortiGate y así mantener aún más el acceso dentro del entorno.

Mandiant observó conexiones SSH desde los dispositivos Fortinet a servidores ESXi dentro del entorno objetivo, seguidas de la instalación de paquetes de instalación de vSphere[14] que contenían las puertas traseras **VIRTUALPITA** y **VIRTUALPIE**.

En otro escenario donde FortiManager estaba restringido desde Internet, UNC3886 aprovechó el acceso previamente establecido para instalar una utilidad de redireccionamiento de tráfico de red, que Mandiant rastrea como **TABLEFLIP**, y una variante de puerta trasera de shell inverso, **REPTILE**, en FortiManager. Este uso combinado de malware permitió a UNC3886 eludir las listas de control de acceso a la red (ACL) implementadas para restringir el acceso externo.

En ambos escenarios, se detectó actividad maliciosa luego de una vulneración total tanto del ecosistema de Fortinet como del hipervisor de VMware, una vez que UNC3886 comenzó a realizar comandos de reconocimiento y a filtrar datos utilizando procesos legítimos del sistema.

Para obtener una descripción detallada de este caso práctico, consulte el blog "Fortinet Zero-Day and Custom Malware Used by Suspected Chinese Actor in Espionage Operation".

CASTLETAP es un binario de Linux que escucha pasivamente los paquetes y activa la funcionalidad de puerta trasera cuando recibe un paquete ICMP Echo. Dentro de estos paquetes, el malware también busca información del servidor C2 al que pueda conectarse nuevamente a través del conector SSL. Sus capacidades incluyen cargar y descargar archivos, generar shells normales y basados en Busybox.

THINCRUST es una puerta trasera de Python integrada en el código de una biblioteca de terceros que permite la ejecución remota de comandos, lectura y escritura de archivos a través de solicitudes HTTP. Los comandos cifrados se almacenan en cookies HTTP.

VIRTUALPITA es una puerta trasera pasiva de 64 bits para Linux y VMware ESXi que crea un listener en números de puerto TCP o VMCI codificados. Admite la ejecución de comandos arbitrarios, carga y descarga de archivos y la capacidad de iniciar y detener vmsyslogd.

VIRTUALPIE es una puerta trasera escrita en Python que genera un listener IPv6 preparado para trabajar en segundo plano en un puerto TCP codificado. Admite transferencia de archivos, ejecución de comandos arbitrarios y capacidades de shell inverso. Se comunica mediante un protocolo personalizado y los datos se cifran mediante RC4.

TABLEFLIP es una utilidad de Linux que realiza el redireccionamiento del tráfico. Escucha pasivamente en todas las interfaces activas los paquetes de comandos especializados. Estos paquetes contienen la dirección IP codificada XOR y el número de puerto al que redireccionar el tráfico mediante comandos iptable.

REPTILE es un rootkit de Linux disponible públicamente escrito en C. Admite la funcionalidad de puerta trasera que se puede activar a través de paquetes ICMP, UDP o TCP mediante la activación de puertos. Las capacidades adicionales incluyen shell inverso y transferencia de archivos.

14. https://blogs.vmware.com/vsphere/2011/09/whats-in-a-vib.html

Estudio de caso de APT29

Mandiant también ha observado a actores con respaldo de una nación, como APT29, atacando a tipos similares de dispositivos perimetrales con un tunneler novedoso.

A principios de 2022, luego de obtener acceso al entorno objetivo, APT29 implementó **QUIETEXIT** en endpoints en todo el entorno. En un caso, APT29 secuestró scripts de inicio de aplicaciones específicas legítimas para permitir que **QUIETTEXT** se ejecutara al inicio, ya que no tiene mecanismos de persistencia nativos. **QUIETEXIT** admite la funcionalidad SSH completa y APT29 aprovechó un túnel SOCKS hacia el entorno objetivo. Esto permitió a APT29 ejecutar herramientas para robar datos con poca o ninguna evidencia en la computadora objetivo. APT29 atacó al almacenamiento conectado a la red (NAS) disfrazando el nombre binario para mezclarlo con archivos legítimos en el sistema de archivos. Para mantener acceso adicional, APT29 implementó una puerta trasera secundaria, el web shell **REGEORG**, en un servidor web DMZ. Esto, combinado con la falta de soluciones antivirus o EDR compatibles, contribuyó a un tiempo de permanencia prolongado.

QUIETEXIT es un tunneler SSH inverso que se conecta a un C2 remoto, pero requiere una contraseña para autenticarse. QUIETEXIT puede ejecutar comandos o interceptar tráfico a través de SOCKS. QUIETEXIT se deriva del software cliente-servidor SSL DROPBEAR de código abierto.

REGEORG es una utilidad de código abierto que se utiliza para canalizar el tráfico de webshell.

QUIETEXIT. Los sistemas de comando y control (C2) observados por Mandiant eran principalmente sistemas de cámaras de salas de conferencias heredados, que probablemente estaban infectados con el componente del servidor de QUIETTEXT. Al atacar estos sistemas de confianza, APT29 permaneció sin ser detectado en los entornos objetivo durante al menos 18 meses.

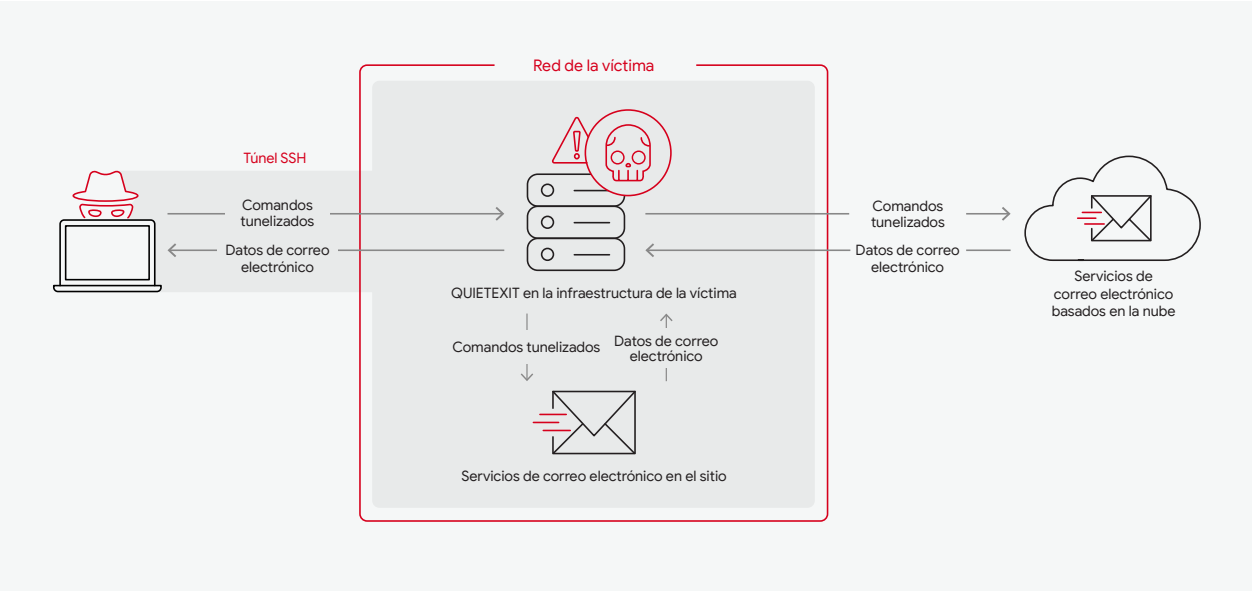


FIGURA 6: Tunneling a través de QUIETTEXIT

Al completar la misión, APT29 obtuvo con éxito credenciales con privilegios para el entorno de correo electrónico del objetivo y centró sus esfuerzos en equipos ejecutivos y empleados que trabajan con desarrollo corporativo, fusiones y adquisiciones, o el personal de seguridad de TI. En algunos casos, APT29 aprovechó las mismas herramientas API eDiscovery y Graph utilizadas para realizar búsquedas programáticas y tener acceso a datos de correo electrónico que los investigadores utilizan para las iniciativas de respuesta. Estas herramientas permitieron a APT29 realizar una exfiltración masiva de correos electrónicos.

Para obtener una descripción detallada de este caso práctico, consulte el blog
"Eye Spy on Your Email".

Estudio de caso de APT28

En 2022, Mandiant observó que APT28 se desvió de su actividad histórica. Este grupo demostró una preferencia por atacar la infraestructura perimetral para realizar una variedad de operaciones, una técnica que se denomina "Vivir al límite" (Living on the Edge). Desde el comienzo de la guerra en Ucrania, la inteligencia militar rusa, conocida como la GRU, ha intentado llevar a cabo campañas sucesivas y casi constantes de espionaje cibernético y disrupción dirigidas contra servicios y organizaciones clave dentro de Ucrania. Este equilibrio de acceso y acciones contra las organizaciones objetivo se basa en el compromiso de la infraestructura perimetral, como enrutadores y otros dispositivos conectados a Internet.

Para obtener una descripción detallada de este caso práctico, consulte:
M-Trends 2023, "La invasión de Ucrania: Operaciones cibernéticas durante tiempos de guerra".

Conclusiones clave

En estos casos prácticos, se detectó evidencia de vulneración dentro del entorno durante la actividad posterior a la explotación, ya que, por diseño, los actores atacan a las redes perimetrales para permanecer sin ser detectados. Durante las investigaciones, Mandiant realizó revisiones exhaustivas de los sistemas afectados para identificar el vector de entrada inicial. En estos casos, existían pruebas para rastrear el acceso hasta las direcciones IP de los dispositivos perimetrales. Esto llevó a los investigadores a trabajar con proveedores para recopilar imágenes forenses de estos dispositivos y realizar análisis más detallados. La comunicación y la colaboración entre organizaciones son fundamentales para que tanto los fabricantes puedan conocer con antelación, antes de que se hagan públicos, los nuevos métodos de ataque como para que los investigadores dispongan de los conocimientos necesarios para esclarecer mejor estos nuevos ataques.

Qué puede hacer para protegerse contra estos ataques

Los actores relacionados con el espionaje cibernético han aumentado su inversión en investigación y desarrollo de herramientas y exploits contra sistemas que generalmente no soportan EDR. Este tipo de herramientas y exploits requieren un conocimiento profundo de los sistemas operativos objetivo. Si bien las organizaciones continúan construyendo centros de operaciones de seguridad (SOC), también deben continuar ampliando la visibilidad más allá de la detección del endpoint. Las brechas de visibilidad permiten a los actores evadir la detección con un mínimo esfuerzo. Determinar esas brechas de visibilidad es el siguiente paso para construir un SOC eficiente que respalde la seguridad de la organización. Las organizaciones deben inventariar los dispositivos en la red y evaluar si hay herramientas de supervisión disponibles para cada uno. Es probable que cada dispositivo que no admita herramientas de supervisión tenga acciones de refuerzo específicas del proveedor para garantizar que se habilite el registro adecuado. Las organizaciones también deben asegurarse de que estos registros específicos de proveedores se envíen a un repositorio central. También se debe evaluar la utilización de controles de acceso a la red para limitar o restringir completamente el tráfico de salida de estos dispositivos. La implementación de supervisión de red adicional y la búsqueda de tráfico anómalo hacia y desde dispositivos perimetrales y otras tecnologías no habilitadas para EDR permite mayores capacidades de detección si estos controles de red no son factibles.

Para obtener recursos adicionales, consulte lo siguiente:

Guía de refuerzo de Microsoft 365 de Mandiant

Detección y refuerzo dentro de hipervisores ESXi

Seis consejos para implementar la gestión de activos con privilegios

La mayor adopción de servicios en la nube y aplicaciones SaaS está aumentando exponencialmente la cantidad de cuentas que las organizaciones deben operar y gestionar. Por ejemplo, hoy en día el empleado promedio puede acceder a 30 cuentas y aplicaciones corporativas. Además, las identidades de las máquinas, los certificados digitales y las claves ahora superan en número a las identidades humanas por un factor de 45 veces[15].

Para las organizaciones que luchan por reducir las cuentas innecesarias y eliminar privilegios excesivos para humanos y sistemas que no los requieren, implementar una gestión de accesos con privilegios (Privileged Access Management, PAM) puede ser útil.

PAM es una práctica de controlar y asegurar el acceso a los activos dentro de una empresa, mediante:

La creación de flujos de trabajo de autorización

El almacenamiento y cifrado de secretos de forma segura

La auditoría, la supervisión y el registro de eventos de acceso con privilegios

El establecimiento de políticas para la gestión de secretos (por ejemplo, cambios de contraseña)

La protección y el aislamiento del acceso a los sistemas objetivo a través de un administrador de sesión

Tradicionalmente, las organizaciones se apoyan en soluciones de autenticación multifactor (MFA) como tecnología principal en su enfoque de PAM. Si no se implementa y no se mantiene adecuadamente, la solución MFA puede presentar un riesgo no deseado para la organización.

15. 5 Reasons to Prioritize Privileged Access Management, CyberArk, 2022

2017	2019	2021	2022
Equifax Los atacantes llegan a la información de identificación personal (PII) de aproximadamente 147 millones de consumidores.	**Universidad Nacional Australiana** Los atacantes acceden a 19 años de PII de personal y estudiantes.	**Verkada** Un ataque a la cadena de suministro en el que los atacantes acceden al sistema de cámaras de seguridad Verkada utilizado en hospitales, escuelas y prisiones.	**Departamento de Asuntos de los Veteranos de EE. UU.** Credenciales confidenciales para sistemas que contienen registros médicos expuestos en GitHub.

FIGURA 7: Cronología de las vulneraciones causadas cuando los atacantes explotan las soluciones de gestión de accesos con privilegios.

Históricamente, los atacantes han aprovechado las vulnerabilidades en las soluciones de gestión de accesos con un alto grado de éxito. Las vulneraciones de datos notables en tamaño y alcance citan vulnerabilidades de PAM que se remontan a 2017 con Equifax, donde los atacantes obtuvieron acceso a información personal y confidencial de 147 millones de consumidores[16]. Seguido por la Universidad Nacional Australiana, donde se expusieron en GitHub credenciales confidenciales de sistemas que contienen registros médicos[17]. En 2021, un ataque a la cadena de suministro contra el proveedor de seguridad física Verkada expuso el acceso a los sistemas de cámaras de seguridad utilizados en hospitales, escuelas y cárceles[18]. Finalmente, en 2022, el Departamento de Asuntos de los Veteranos de EE. UU. fue víctima de la exposición de datos de credenciales de cuentas con privilegios por parte de un contratista[19].

Mandiant ha observado que los actores evadieron los controles de MFA en varias instancias con éxito. En un caso, grupos de amenazas persistentes avanzadas (Advanced Persistent Threat, APT) con sede en Rusia realizaron ataques de fatiga a la MFA[20] enviando repetidamente solicitudes de autenticación de segundo factor al correo electrónico, teléfono o dispositivos registrados de la víctima objetivo para obtener acceso a cuentas de correo electrónico, lo que provocó incidentes de fraude electrónico.

En otro ejemplo, Mandiant observó que APT29[21] se aprovechaba del proceso de registro automático para MFA en una organización, que permitía inscribir un dispositivo a cualquier persona con un nombre de usuario y contraseña. APT29 realizó ataques de suposición de contraseña para intentar encontrar cuentas sin dispositivos registrados y añadió las suyas propias.

Independientemente del tamaño de la organización o de la madurez del programa PAM, los líderes de seguridad deben tomarse el tiempo para revisar estos 7 consejos al implementar PAM para ayudar a proteger su negocio.

16. Wallix Cybersecurity, Equifax Breach: Preventing Data Breaches with Privileged Access Management
17. Australian National University, Incident Report on the Breach of the Australian national Universities Administrative Systems, 2019
18. Verkada, Summary: 9 de marzo de 2021, Security Incident Report, 2021
19. FedScoop, VA investigates breach after federal contractor publishes source code, septiembre de 2022
20. Mandiant, Suspected Russian Activity Targeting Government and Business Entities Around the Globe, diciembre de 2021
21. Mandiant, You Can't Audit Me: APT29 Continues Targeting Microsoft 365, agosto de 2022

01

Comprender las cuentas con privilegios

A los líderes de seguridad y TI a menudo se les pregunta: "¿Qué es una cuenta con privilegios?" Si bien la respuesta genérica es que todas las cuentas pueden tener algún nivel de privilegio, a continuación se presentan varias categorías de cuentas que brindan privilegios más altos:

- **Administradores de dominio:** Usuarios que tienen control total sobre un dominio.
- **Cuentas personales con privilegios:** Cuentas de usuario con más privilegios que un usuario normal. Los usuarios utilizan esto caso por caso.
- **Cuentas predeterminadas:** Cuentas creadas automáticamente por el sistema o la aplicación (por ejemplo, SA, Root, mysql, ec2-user).
- **Cuentas de servicio:** Cuentas que se asignan a máquinas y brindan acceso a sistemas, servicios y aplicaciones corporativos.
- **Raíz, superadministrador o administrador global (nube):** Cuentas de administrador adicionales para un sistema que otorga al usuario control total sobre el dispositivo local.
- **Cuentas de emergencia:** Cuentas utilizadas para obtener acceso a los sistemas en caso de un incidente de seguridad.
- **Cuentas de seguridad:** Cuentas utilizadas por el personal de seguridad para acceder a los sistemas y realizar auditorías e investigaciones de seguridad.

Es importante comprender el riesgo asociado con el uso indebido de cuentas que brindan acceso con privilegios. Comience con privilegios mínimos para garantizar que cada usuario solo pueda realizar las acciones definidas por su función.

Preste especial atención a las funciones que acceden a información de identificación personal (PII) o propiedad intelectual (IP).

Acciones a tomar

- Realice una evaluación de riesgos del acceso con privilegios dentro de su organización. Identifique cuentas tanto para personas como para sistemas que presenten riesgos para la información y los activos críticos. Considere los tipos de permisos, incluidos los procedimientos de identificación, como el inicio de sesión interactivo. Priorice las cuentas PAM de alto riesgo.
- Obtenga la aceptación de la alta dirección y la gerencia ejecutiva para impulsar la implementación de herramientas que reducirán el riesgo general dentro de una organización.
- Asegúrese de que los equipos de seguridad y tecnología de la información colaboren en la implementación para tener en cuenta las necesidades de varios grupos de usuarios y las comunicaciones y la gestión de cambios requeridas en las implementaciones de PAM.
- Realice la recertificación y validación de permisos asignados a las cuentas. Al mantener esta cuenta, no se debe cambiar; si hay un nuevo caso de uso, entonces se debe crear el tipo correcto de cuenta o se debe otorgar una política de acceso por tiempo limitado tras las autorizaciones correspondientes.

02

Establecer un proceso continuo para la creación, el descubrimiento y la incorporación de cuentas

A medida que PAM se implementa en todo el entorno, los equipos de PAM deben abordar de manera proactiva las brechas de seguridad dentro de la organización. Un aspecto crítico de este esfuerzo es incorporar todas las cuentas necesarias que hayan sido identificadas por los equipos de aplicaciones y que se comprendan las implicaciones de gestionar estas cuentas a través de la solución PAM.

No incorporar todas las cuentas necesarias puede tener como resultado la proliferación de cuentas con privilegios no seguras, lo que deja a una organización vulnerable. Los atacantes pueden explotar estas cuentas para obtener acceso a sus sistemas, elevar privilegios, desplazarse lateralmente y establecer persistencia.

El problema de las cuentas con privilegios no seguras es particularmente desafiante cuando se agregan nuevas cuentas y servicios a un entorno. Estas incorporaciones aumentan aún más la superficie de ataque y el alcance del descubrimiento de cuentas con privilegios, lo que agrava el riesgo de una vulneración de seguridad.

Al mantener un inventario completo de todas las cuentas antiguas y nuevas dentro de un entorno, las organizaciones pueden identificar rápidamente cuáles son las cuentas en riesgo durante un incidente de seguridad. Esto ayuda a proteger esas cuentas, identificar los sistemas a los que tienen acceso y crear rutas confiables para acceder a activos críticos. Esto puede aliviar la presión sobre su equipo de seguridad, los equipos de respuesta a incidentes y los administradores de incidentes al responder a incidentes de seguridad.

Acciones a tomar

- Incorpore cuentas cuando se crean y evite aumentar el "alcance de descubrimiento"
- Utilice herramientas de descubrimiento para identificar cuentas que se han omitido, incorpórelas e implemente controles eficaces para gestionar el ciclo de vida de la cuenta.
- Comprenda el alcance de cuentas con privilegios. Considere dónde se almacena la propiedad intelectual, PII o PHI y cómo se accede a ella.
- Establezca un proceso continuo para el descubrimiento de cuentas. Trabaje con los equipos para adoptar la automatización para la creación, el descubrimiento y la incorporación.
- Aproveche el marco MITRE ATT&CK® para revisar docenas de técnicas adversarias que por lo general se usan indebidamente en ataques de escalación de privilegios.

03

Garantizar controles de acceso adecuados para la implementación de PAM

PAM está diseñado para proteger las llaves del reino. Por lo tanto, se debe gestionar el acceso a las soluciones PAM, ya que estas cuentas de usuario y sistemas pueden convertirse en objetivos.

Las autorizaciones para la administración de PAM no deben estar vinculadas a los directorios que protege. Utilice el directorio integrado de la herramienta PAM para gestionar este acceso.

Considere también el flujo de trabajo para la autorización. Este control mejora la capacidad de la organización para defenderse contra amenazas internas y ayuda al equipo de PAM a comprender cómo controlar el sistema implementado para PAM.

Acciones a tomar

- **Incorpore este tipo de cuentas**
  - **Administradores de herramientas PAM**
  - **Cuentas de aplicaciones**
  - **Cuentas de servidor**
  - **Automatización y scripting**
- **Revise el modelo de acceso para utilizar la solución PAM e identifique combinaciones tóxicas de privilegios.**
- **Asegúrese de que las cuentas de administración de la solución PAM tengan los permisos correctos, limitando el acceso a las credenciales que protege la aplicación.**
- **Configure flujos de trabajo de autorización y acceso para proteger cuentas críticas.**
- **Proporcione a sus equipos conocimientos de seguridad esenciales para mejorar continuamente la postura de seguridad de la organización a través de la capacitación, la habilitación y el soporte de expertos.**

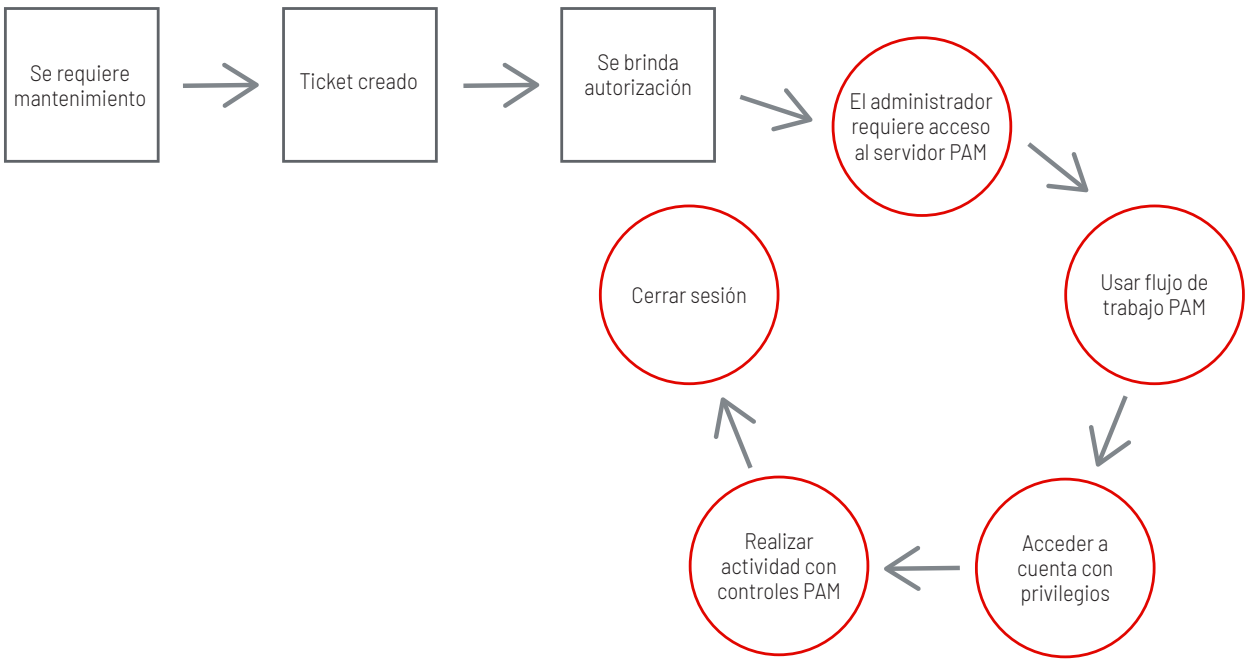


FIGURA 8: Flujo de trabajo de autorización PAM

04

Mapa y rutas de acceso seguro

Al acceder a los activos dentro de la organización, existen actividades clave para asegurar el acceso con privilegios. Una acción es evitar que las contraseñas queden expuestas en un sistema de endpoint. También hay consideraciones sobre lo que sucede entre el usuario y el sistema objetivo.

¿El tráfico se envía a través de una ruta confiable?
¿Cómo se puede aplicar esto?

Otras preguntas para hacer incluyen:

- ¿La conexión implica un recurso web?
- ¿El recurso web está conectado directamente desde una estación de trabajo?
- ¿Se utiliza HTTPS para las conexiones?

Comprender el flujo y la ruta hacia los objetivos ayuda a calcular el riesgo de una amenaza. También se debe considerar cómo las aplicaciones acceden a los secretos y cómo se utilizan esos secretos entre aplicaciones y servidores.

Acciones a tomar

Cree un mapa de acceso dentro del entorno y verifique que la arquitectura corresponda con lo que se proporciona con las herramientas PAM.

- **Identifique los protocolos seguros que se deben utilizar e identifique las rutas que se deben tomar para obtener acceso a los administradores y objetivos de sesión.**
  - **Conéctese a través de HTTPS en lugar de RDP**
  - **Considere utilizar un listener inverso para mantener claras las NACL y permitir solo el acceso saliente.**
  - **Establezca un modelo de niveles de credenciales y acceso.**
- **Asegúrese de que los terceros estén protegidos con mecanismos de autenticación seguros y que la autorización se proporcionó correctamente.**
- **Confirme que los terceros se están conectando a través de un método seguro.**
- **Identifique las rutas de acceso:**
  - **Ruta a través de redes internas y públicas**
  - **Sistemas de fuente limpia**
  - **Fuertes niveles de acceso**
  - **Protección de credenciales**
  - **Rutas de acceso de prueba**

05

Implementar el registro con retención adecuada

El registro y la supervisión brindan información valiosa y esencial tras un ataque cibernético. El registro y la supervisión tienen como objetivo ayudar a identificar el alcance y el impacto de un incidente cibernético. Los investigadores forenses utilizan fuentes de registros para responder varias preguntas durante un incidente cibernético. Por ejemplo, para identificar la exfiltración de datos, los investigadores forenses dependen en gran medida de los datos de registro de firewall o NetFlow. Estos registros pueden responder preguntas sobre la exfiltración de datos y la cantidad de datos que han salido de la red. Otro ejemplo implica el seguimiento de la actividad del usuario. Si una cuenta con privilegios se ve comprometida, el registro puede ayudar a rastrear las acciones realizadas por ese usuario.

Acciones a tomar

- Implemente una solución de registro y supervisión que capture las actividades de cuentas con privilegios en toda la organización.
- Garantice una retención de registros adecuada: Desarrolle una política de registro y supervisión que describa los tipos de actividades que deben registrarse y la retención de esos registros.
- Confirme que los datos de registro se envíen a los sistemas apropiados, pero sobre todo que los datos se utilicen para enriquecer las defensas dentro de la organización. Los equipos de seguridad pueden aprovechar el análisis de amenazas para mejorar aún más los controles que se han implementado y como indicadores de un actor que aterriza en el entorno.
- Busque anomalías en la actividad del usuario, incluido el acceso al sistema.
- Comprenda el plan de acción cuando ocurre un incidente. Cree un plan para revisar los registros y datos de auditoría si no los almacena en un SIEM.

06

Implementar MFA

La autenticación multifactor (MFA) es importante para comprobar la identidad digital y proteger el acceso a un sistema a través de un único actor o entidad.

La MFA requiere que los usuarios proporcionen múltiples factores de autenticación para acceder a una aplicación. Dos de las formas más comunes de MFA son los códigos de un solo uso (one-time passcode, OTP) y las notificaciones automáticas. Los OTP son códigos que los usuarios reciben en sus dispositivos móviles a través de aplicaciones MFA (por ejemplo, Google Authenticator), que pueden usarse para autenticarse. Las notificaciones automáticas envían una notificación al dispositivo móvil de un usuario para aprobar o rechazar un intento de inicio de sesión.

Ambos métodos son vulnerables a intentos de phishing o ataques de intermediario. Ataques recientes han demostrado que las notificaciones automáticas de MFA o los códigos enviados por SMS no son suficientes para proteger el acceso. Por ejemplo, un ataque de fatiga a la MFA[22] ocurre cuando los actores bombardean a un usuario con notificaciones automáticas con la esperanza de que el usuario se frustre y presione aceptar.

Acciones a tomar

- Implemente MFA sólida y resistente al phishing
  - Autenticación FIDO2 mediante el uso de claves biométricas o de hardware (por ejemplo, YubiKey)
  - Utilice mecanismos de respuesta a desafíos que no permitan simplemente aceptar (coincidencia de números)
- Capacite a todos los empleados para garantizar que tengan el conocimiento disponible y utilicen la MFA correctamente
- Introduzca alertas y etiquetado basado en riesgos para cuentas que parezcan estar bajo ataque
  - Ubicación geográfica
  - Horarios de acceso anormales
  - Solicitudes excesivas de respuesta ante desafíos de MFA
- Revise la seguridad de autenticación de los autenticadores que se utilizan
- Realice búsqueda de amenazas contra estas identidades

22. Mandiant, Suspected Russian Activity Targeting Government and Business Entities Around the Globe, diciembre de 2021

En conclusión

A medida que más empresas avanzan hacia la implementación de soluciones PAM para mejorar su postura de seguridad, es fundamental que estas herramientas se configuren e implementen adecuadamente. Las configuraciones incorrectas, la falta de una gestión de acceso adecuada y el no utilizar plenamente las capacidades integradas son algunos de los factores clave que pueden generar una falsa sensación de seguridad y, en algunos casos únicos, un mayor riesgo empresarial.

Más información en www.mandiant.com

Mandiant

11951 Freedom Dr, 6th Fl, Reston, VA 20190
(703) 935-1700
833.3MANDIANT (362.6342)
info@mandiant.com

Acerca de Mandiant

Mandiant es un líder reconocido en defensa cibernética dinámica, inteligencia de amenazas y servicios de respuesta a incidentes. Gracias a décadas de experiencia en el frente de batalla, Mandiant ayuda a las organizaciones a confiar en su preparación para defenderse y responder a las amenazas cibernéticas. Mandiant ahora parte de Google Cloud.

MANDIANT
NOW PART OF Google Cloud