Prueba y Evaluación

Evidencia, pruebas y guía de evaluación: resultados documentados, cómo realizar una prueba de concepto significativa y qué preguntar a cualquier proveedor de prevención.

¿Cyber Crucible supera las pruebas de simulación de ransomware?

La mejor respuesta es... depende de la calidad y precisión de las simulaciones de ransomware, pero no hemos visto muchas pruebas de alta calidad que coincidan con las herramientas y comportamientos de ataques reales.

Cyber Crucible examina los patrones de comportamiento subyacentes en el acceso a archivos, los comportamientos de memoria, los comportamientos de procesos y los comportamientos criptográficos para identificar y suspender las actividades de extorsión de datos en conjunción con un ataque.

Algunas simulaciones de software de extorsión de datos han evolucionado con el tiempo, para ser una representación más precisa de los ataques reales.

Hemos experimentado simulaciones que, al parecer, se centraban en verificar las contramedidas de extorsión de datos divulgadas por algunos proveedores, y no en las tácticas de las herramientas de extorsión y de los propios atacantes. Una forma de expresarlo podría ser: “Probar la contramedida, no el ataque”.

En respuesta, nunca rehuimos las pruebas frente a la emulación de atacantes, las pruebas de penetración o las herramientas de extorsión. Atacamos rutinariamente nuestro propio software y reproducimos las tácticas de los atacantes que observamos comentadas en línea o que encontramos (y detuvimos) en entornos de clientes.

¿Cyber Crucible tiene falsos positivos? ¿Cuál es la tasa de falsos positivos?

Cyber Crucible se esfuerza por lograr un entorno de producto con 0 falsos positivos. Dicho esto, en ocasiones ocurren falsos positivos. Analicemos de dónde provienen.

Actualmente, la tasa de falsos positivos es de aproximadamente 1 respuesta por mes, por cada 1800 implementaciones/agentes, con total trazabilidad sobre por qué ocurren esas respuestas.

Programa Corrupto Realizando Operaciones Rápidas de Archivos

Ya sea a propósito o debido a una mala programación, a veces los programas se corrompen a sí mismos, o son corrompidos por otros programas mientras se ejecutan. Vemos esto con frecuencia en nuevas funciones de suites de programas grandes, como Microsoft Office, o software personalizado de propiedad exclusiva.

Lo que sucede es que la memoria del programa se corrompe, lo que aumenta el nivel de inspección del programa. Si eso va seguido de un comportamiento de acceso a archivos similar al de una extorsión, Cyber Crucible se ve obligado a suspender el programa. Cada vez somos mejores identificando errores conocidos, y usted al menos podrá ejecutar su programa hasta que falle (algunos se reinician automáticamente, otros no).

En un mundo donde los atacantes no usan el sistema de archivos, y en su lugar secuestran la memoria de los programas en ejecución, debemos actuar con precaución.

Cyber Crucible ahora tiene la capacidad de crear un permiso, en el que, si el Programa A, con los Argumentos A, causa una corrupción en el Programa B, con los Argumentos B, podemos crear una exclusión temporal para:

(Programa A + Argumentos A) abre (Programa B + Argumentos B)

De esa manera no se asume un riesgo adicional de que un atacante esté involucrado, a menos que este utilice exactamente los mismos argumentos.

Un excelente ejemplo de esto fue cuando Microsoft introdujo por primera vez la apertura de documentos de Office desde dentro de su software de chat de escritorio. Corrompía el programa de Office…todas…y…cada…una de las veces…luego el programa de Office comenzaba a escanear todos los Documentos en la máquina aproximadamente el 25% de las veces. Eso… necesitaba una exclusión.

Si experimenta un fallo, use el botón de soporte, y le ayudaremos a crear una exclusión, hasta que el proveedor solucione el problema.

Programas de Seguridad

Cyber Crucible convive muy bien con programas de seguridad. Contrario a lo que la mayoría de los nuevos clientes piensan, las herramientas de seguridad más avanzadas suelen usar interacciones con el sistema más avanzadas que los programas antivirus que utilizan tecnologías más antiguas. De hecho, algunas herramientas de seguridad gratuitas o económicas ejecutan una variedad de scripts de Powershell inseguros como parte clave de su protección, en lugar de programas reales.

Como una herramienta de seguridad altamente resiliente basada en el kernel, es mejor considerar a Cyber Crucible como “más baja”, o “más cercana” al hardware que la mayoría de las demás herramientas. Cyber Crucible identifica automáticamente otras herramientas de seguridad y las agrega a los modelos de comportamiento. Mientras las herramientas de seguridad no muestren señales de estar comprometidas, se les permite operar normalmente.

Ocasionalmente, una herramienta de seguridad provocará inestabilidad en el sistema, cuando sus intentos de desactivar o interferir con el software de Cyber Crucible fallan. En esos casos, es mejor incluir en la lista blanca a Cyber Crucible en la otra herramienta de seguridad, para evitar que intente (sin éxito) deshabilitar o interferir con Cyber Crucible.

No dude en abrir un ticket de soporte con Cyber Crucible, a través del portal web, para discutir el asunto.

Herramientas Administrativas o de Copia de Seguridad

En algunas circunstancias, se debe crear un comportamiento personalizado para una herramienta administrativa o de copia de seguridad, mientras el equipo de Cyber Crucible desarrolla un modelo

de comportamiento mejorado. En este caso, el equipo de Cyber Crucible ayudará en la creación de una excepción en el modelo de comportamiento, que aislará el programa junto con sus argumentos, para una ventana de oportunidad extremadamente pequeña ante un posible ataque.

¿Cyber Crucible prueba de manera exhaustiva todo el ransomware posible?

Cyber Crucible opera a partir de un modelado conductual a nivel de kernel, para descubrir rápidamente comportamientos de robo de datos, robo de credenciales y cifrado de ransomware. Las decisiones conductuales se toman utilizando información del comportamiento de los procesos, comportamientos derivados de la memoria y ciertos tipos de comportamientos originados en archivos, para tomar una decisión justo antes de que ocurra la extorsión.

A pesar del gran número de muestras de software de ransomware y extorsión, estas pueden categorizarse conductualmente en un número relativamente pequeño de conjuntos.

Sin embargo, en la superficie, las defensas “superficiales” utilizadas por una variedad de herramientas de seguridad tienen que intentar contrarrestar un número muy grande de herramientas de extorsión. Es importante entender que nuestros análisis conductuales llegan mucho más a fondo en las herramientas y tácticas de los atacantes, reduciendo drásticamente la necesidad de algún tipo de prueba exhaustiva (imposible) de todo el malware posible en todo momento.

Los desarrolladores de Cyber Crucible categorizan el comportamiento a nivel de kernel de la memoria, los procesos y los archivos de una herramienta de extorsión, y luego se aseguran de que encaje en una de las capacidades defensivas conocidas. Ocasionalmente, de manera similar a una vacuna, la fórmula puede ajustarse ligeramente para producir una respuesta más precisa.

Muy rara vez aparece algo completamente nuevo.

El equipo de Cyber Crucible trabaja de manera exhaustiva para descubrir de forma preventiva técnicas novedosas que aún no se han observado en nuestra base de clientes ni en la inteligencia de amenazas.

¿El efecto?

1. La defensa de Cyber Crucible contra herramientas de extorsión es completa frente a todas las variantes de ransomware conocidas.

2. Las nuevas variantes de ransomware se bloquean incluso antes de afectar a los primeros clientes. Ni siquiera sabemos cómo llamar a la mayoría del software contra el que nos defendemos durante aproximadamente 120 días, hasta que los investigadores “se ponen al día”.

3. Los desarrolladores de ransomware y herramientas de extorsión a veces nos llaman para ver qué estamos haciendo, en un intento de encontrar un ángulo. Lamentablemente, la frustración causada por nuestra presencia parece a veces impulsar su evolución, lo que hace que otras herramientas sean incluso menos efectivas.

4. Damos la bienvenida a las pruebas de nuestro software. Cuando profesionales de pruebas de penetración, analistas de malware o especialistas en emulación de adversarios entran en el proceso de venta, ¡nos entusiasma!

¿Qué evidencia existe de que Cyber Crucible realmente detiene los ataques?

Respuesta breve: Implementaciones documentadas en clientes, pruebas independientes realizadas por una importante firma de contabilidad y asesoría, y repetidas instancias de detención de ataques de día cero meses antes de que otros proveedores los reportaran.

Resultados documentados

- **Servicios financieros:** casi 10,000 procesos maliciosos interceptados de forma autónoma en 60 días, el 98% en una granja de servidores Microsoft SQL con altos privilegios, sin alertas de tres EDR implementados, un MDR o un SOC subcontratado del Top 50.
- **Pruebas independientes:** una de las firmas de contabilidad y asesoría más grandes de Norteamérica sometió a Cyber Crucible a su propia evaluación frente a escenarios de ransomware, robo de datos y fraude de identidad.
- **A la vanguardia de la industria:** Cyber Crucible ha detenido múltiples ataques de día cero en redes de clientes hasta 90 días antes de que cualquier otro proveedor los reportara o respondiera a ellos.
- **Ataques basados en navegador:** desde el cuarto trimestre de 2023 en adelante, las respuestas automatizadas contra actividad de Chrome, Edge y Chromium aumentaron de cero a miles, con los CVE relacionados publicados posteriormente por Google y Microsoft.
- **Pagos de rescate:** aproximadamente el 90% de las víctimas de ransomware pagaron en 2023. Los clientes de Cyber Crucible pagaron $0.

Resiliencia ante ataques directos

En una implementación marítima, los adversarios escalaron hasta secuestrar credenciales de Windows y bloquear sistemas a nivel de BIOS cuando los intentos convencionales fallaron. Con otros dos clientes, los atacantes que no pudieron vencer el motor de decisiones del kernel intentaron en su lugar bloquear la notificación al cliente atacando la pila de red del sistema operativo, y fallaron, porque las comunicaciones de red de Cyber Crucible son independientes del sistema operativo.

¿Cómo debemos ejecutar una prueba de concepto para una herramienta de prevención?

Respuesta breve: impleméntela junto a su stack existente en un entorno de producción real y compare lo que detecta cada herramienta. El resultado más informativo no es una prueba de laboratorio: es la diferencia entre lo que sus herramientas actuales alertan y lo que realmente se intercepta.

Por qué implementarla en paralelo funciona mejor

La empresa de servicios financieros de nuestro caso mejor documentado hizo exactamente esto. No eliminaron nada. Añadieron Cyber Crucible específicamente para descubrir qué se le estaba escapando a su inversión existente, y obtuvieron una respuesta definitiva en un plazo de 60 días.

Qué medir

- **Intercepciones sobre las que su stack actual nunca alertó.** Esta es la señal principal.
- **Tiempo hasta la decisión.** La prevención debe completarse antes de que se produzca el daño, no antes de que se cierre un ticket.
- **Interrupción del negocio.** Cuente el tiempo de inactividad, los reinicios y las interrupciones provocadas por falsos positivos.
- **Horas de analista consumidas.** La prevención autónoma debe reducir la carga de trabajo, no añadir una cola.

Una advertencia sobre las pruebas de laboratorio

Probar con muestras extraídas de bases de datos de malware a menudo mide muy poco, porque los atacantes diseñan el malware para que permanezca inactivo a menos que reciba una señal de

validación de un servidor de mando y control en vivo que ellos aún operan. Una muestra que no hace nada en su laboratorio no demuestra nada sobre un ataque real.

¿Qué preguntas debemos hacerle a cualquier proveedor de seguridad para endpoints?

Respuesta breve: Pregunte dónde se toman las decisiones, qué ocurre sin conectividad, de qué depende la herramienta para funcionar y si previene o simplemente informa. Las respuestas separan rápidamente la arquitectura del marketing.

Preguntas que vale la pena hacer

1. **¿Dónde se toma la decisión de protección: en el endpoint o en su nube?** Un intercambio de red en la ruta crítica no puede superar a un ataque de escala de milisegundos.
2. **¿Qué ocurre cuando el equipo está fuera de línea o aislado (air-gapped)?** Si la protección se degrada, no era local.
3. **¿Su agente depende de las bibliotecas del sistema de Windows para funcionar?** Si es así, un atacante que comprometa esas bibliotecas en memoria puede cegarlo o silenciarlo.
4. **¿Requieren firmas o fuentes de inteligencia de amenazas?** Si es así, la protección va, por definición, un paso detrás del atacante.
5. **¿Una respuesta requiere intervención humana?** Si es así, el tiempo de respuesta está limitado por la biología humana.
6. **¿Almacenan nuestras claves de cifrado o cargan datos sensibles para su análisis?** El almacenamiento centralizado crea un objetivo y un riesgo de divulgación forzada.
7. **¿Cuál es su tasa de falsos positivos, y qué hace realmente una respuesta?** El bloqueo total del sistema como única opción de contención resulta costoso en un hospital o una planta industrial.

Por qué importan estas preguntas

La mayoría de las herramientas para endpoints se ven similares en una ficha técnica. Estas preguntas ponen de manifiesto las decisiones arquitectónicas que determinan si una herramienta puede actuar a tiempo, o si solo puede explicar, después del hecho, lo que ocurrió.

¿Qué le sucedió al modelo de "recolector tonto" de la industria de la seguridad?

Respuesta breve: Durante años, los proveedores aconsejaron que el procesamiento local en el endpoint estaba obsoleto e impulsaron agentes ligeros que reenvían telemetría sin procesar a la analítica en la nube. Los ataques modernos en memoria contra bibliotecas centrales del sistema operativo han dejado al descubierto ese modelo: los agentes siguen ejecutándose, pero quedan efectivamente ciegos y mudos.

Por qué se adoptó el modelo

Era más barato de construir. La ingeniería a nivel de kernel es difícil y costosa; enviar la telemetría a la nube permitía a los proveedores distribuir software de endpoint económico, monetizar grandes reservas de almacenamiento en la nube y comercializar capacidades de IA en la nube. El modelo se optimizó para la economía del desarrollo, no para la defensa.

Por qué está fallando ahora

Estos agentes dependen por completo de componentes nativos del sistema operativo para funcionar y recopilar datos. Cuando los atacantes comprometen esos componentes en memoria, el agente no se bloquea; sigue ejecutándose y no reporta nada, dejando un panel de control que indica que todo está bien mientras se exfiltran los datos.

También existe una interdicción deliberada documentada: atacantes, y en algunos casos verificados competidores que buscan ocultar sus propias deficiencias, explotan activamente estas dependencias de bibliotecas para forzar la terminación o sabotear las defensas del endpoint.

El callejón sin salida arquitectónico

Corregir esto exigiría que los proveedores heredados abandonen las canalizaciones de telemetría orientadas a la nube y reconstruyan para la computación local a nivel de kernel en el borde ("edge computing"), un esfuerzo de varios años. Las adquisiciones recientes y la inversión de capital de riesgo muestran que la industria avanza aún más hacia la analítica en la nube y los lagos de datos centralizados, en lugar de hacerlo en la dirección contraria.

¿Por qué la prevención basada en la intención genera menos ruido de alertas que la detección basada en firmas?

Respuesta breve: Porque plantea una pregunta más acotada. Las herramientas de firmas y heurísticas generan alertas cuando algo *se asemeja* a un patrón conocido como malicioso y dejan que los humanos determinen si es relevante. La prevención basada en la intención pregunta si un programa específico está, en ese momento, realizando algo malicioso en un punto de entrada conocido de robo de datos, una pregunta con muchas menos respuestas ambiguas.

> Para conocer la tasa de falsos positivos medida por Cyber Crucible, consulte el artículo existente **"¿Tiene falsos positivos? ¿Cuál es su tasa de falsos positivos?"** Esta página explica la razón arquitectónica por la que el perfil de ruido difiere.

Por qué el perfil de ruido es diferente

Las herramientas de firmas y heurísticas generan alertas cuando algo *se asemeja* a un patrón conocido como malicioso, y luego dependen de la clasificación humana para distinguir lo real de lo irrelevante. Eso es lo que produce más de 100 alertas por endpoint al día en muchos entornos.

La evaluación basada en la intención plantea una pregunta más acotada: ¿está este programa, en este momento, realizando algo malicioso en un punto de entrada conocido de robo de identidad o de datos? Esa pregunta tiene muchas menos respuestas ambiguas.

Qué significa esto en términos operativos

- Los analistas no tienen que perseguir alertas de baja confianza.

- No se requiere ajuste de reglas YARA ni búsqueda manual de amenazas para mantener el ruido bajo control.
- Dado que la respuesta suspende únicamente el proceso infractor, incluso una intercepción incorrecta no provoca la caída de un sistema.

Para conocer las tasas actuales medidas en un entorno como el suyo, consulte a su representante de Cyber Crucible, y vea el artículo existente "¿Tiene falsos positivos?" para más detalles a nivel de producto.