Arquitectura y tecnología

Cómo funciona Cyber Crucible internamente: IA Genética, funcionamiento a nivel de kernel, Rogue Process Prevention, independencia del sistema operativo y por qué las decisiones se toman en el endpoint.

¿Por qué responden suspendiendo los programas?

Los delincuentes han aprendido que operar únicamente en memoria, utilizando técnicas como la inyección de procesos, proporciona una variedad de ventajas:

1. Asumir el control de la identidad de un usuario para ese programa.
2. Evasión de las verificaciones de gestión de identidad.
3. Evasión de las reglas de firewall basadas en aplicaciones y de las listas blancas.
4. Evasión de las herramientas de prevención de pérdida de datos en los servidores.
5. Eliminación de casi toda la evidencia en memoria cuando se elimina el proceso.
6. Eliminación de casi toda la evidencia en memoria una vez que se reinicia el equipo.

Al suspender los programas bajo el control del atacante, la evidencia queda congelada en el tiempo.

El atacante también queda congelado.

La evidencia está disponible para las herramientas forenses como parte de un proceso de investigación.

Si la organización no está equipada para realizar análisis forenses en ese momento, un simple reinicio por parte del atacante devolverá el sistema a un estado saludable. Forzar la finalización del programa suspendido y reiniciarlo, por lo general, también produce el efecto deseado.

¿Cyber Crucible detiene toda inyección de procesos?

¡No!

Existe una variedad de actividades que ocurren en los sistemas que constituyen inyección de procesos, pero que no son necesariamente maliciosas.

Cyber Crucible hace visibles todas estas actividades de inyección de procesos para la caza de amenazas y el análisis de causa raíz, pero no detiene la actividad a menos que comiencen las actividades de extorsión de datos, ya sea robo de datos o cifrado por ransomware.

A continuación se muestra un ejemplo de inyección de procesos realizada por el software de seguridad de endpoints PC Matic, denominado PCPitStop. Es probable que se trate de un comportamiento que llevan a cabo para inspeccionar los procesos en ejecución.

5373968.png?width=680
Gracias al análisis de comportamiento a nivel de kernel de Cyber Crucible, usted cuenta con visibilidad completa de todas las actividades, independientemente del nivel de permisos del software (incluyendo software con altos privilegios de Microsoft o de proveedores de antivirus).

Si se realiza una respuesta automatizada, esta página y la página de Creaciones de Procesos pueden consultarse para obtener resultados de investigación rápidos.

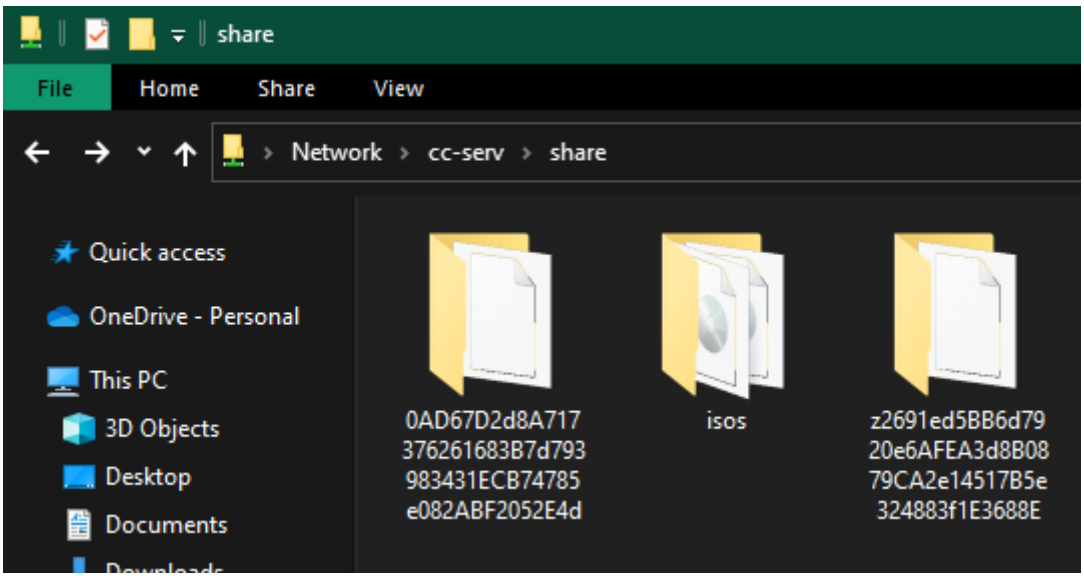
Además, los equipos de caza de amenazas pueden aprovechar la API REST para consultar indicadores de posibles técnicas maliciosas en memoria.

¿Qué son estos archivos extraños y de aspecto aleatorio que veo?

Nuestro software de Prevención de Extorsión de Datos recibe indicadores de comportamiento de varios motores antes de tomar una decisión en fracciones de segundo sobre si hay un ataque en curso. Uno de estos motores, el motor de monitoreo de archivos, utiliza archivos especiales en varias ubicaciones.

Esos archivos que a veces ve, y a veces no, están ahí por Cyber Crucible. Dependiendo de cómo una aplicación pueda listar archivos, es posible que los vea. Por ejemplo, un programa que use el Explorador de Archivos para listar archivos a guardar o abrir, ahora se los mostrará.

Llamamos a esos archivos “canarios”, como en la expresión “canario en la mina de carbón” (wiki).



¿Qué tiene de especial los archivos señuelo (canary files) de Cyber Crucible?

Los archivos señuelo (canary files) tienen datos especiales ocultos en ellos, en caso de que alguna vez sean robados por un atacante. Cyber Crucible puede rastrear el archivo hasta el propietario original.

Los nombres de los archivos son, en su mayoría, aleatorios (denominados pseudoaleatorios) de manera que los atacantes no puedan identificarlos ni evitarlos. Los nombres se crean con una fórmula matemática especial de Cyber Crucible, que permite que nuestro software reconozca un archivo como señuelo, incluso si los nombres difieren de una máquina a otra o de un archivo a otro.

La capacidad del software de Cyber Crucible para reconocer estos archivos también significa que solo se necesita un archivo señuelo por servidor, incluso si miles de estaciones de trabajo están utilizando el servidor.

¿Dónde existen los archivos canary?

Los canarios se encuentran en varios lugares de su sistema de archivos.

Cada vez que aparece una nueva fuente de archivos en su sistema, proveniente de cualquier usuario posible, Cyber Crucible verifica automáticamente si ya se han colocado canarios. Esto incluye unidades extraíbles como memorias USB, fuentes de archivos de red y almacenamiento de archivos basado en la nube.

¿Cómo utiliza Cyber Crucible la analítica de memoria?

Los análisis del kernel de Cyber Crucible realizan un seguimiento del estado de la memoria de un programa durante todo su ciclo de vida, desde el inicio hasta el final del programa. También realiza un seguimiento del estado de las interacciones en memoria con otros programas.

El monitoreo de comportamiento detecta estados inseguros en la memoria del programa. Esto podría significar un exploit, un error de software, o ataques conocidos en memoria (también conocidos como “file-less”) en curso.

Este monitoreo se realiza en tiempo real sobre los programas de manera altamente eficiente, para garantizar que las operaciones del sistema y del cliente no se vean afectadas por el análisis de Cyber Crucible.

En el momento de un posible comportamiento de atacante orientado al robo de identidad o de datos, o a la extorsión, este seguimiento y análisis del comportamiento de la memoria se incluye junto con el resto de la telemetría recopilada, para determinar si un programa debe suspenderse.

¿En qué se diferencia el Análisis de Comportamiento de Memoria de CC del escaneo de memoria?

¿Qué es el escaneo de memoria?

El escaneo de memoria es la práctica de tomar un bloque de memoria, normalmente copiándolo a una ubicación diferente en el sistema para su análisis, y escanearlo para su análisis.

Fotografía vs. acción en vivo

Esto no es un escaneo en vivo del programa mientras se comporta, sino una instantánea en el tiempo del programa. Piense en ello como una fotografía, en comparación con una transmisión en vivo. El estado del programa puede ser drásticamente diferente de un momento a otro, por lo que es más útil como ejercicio forense en un entorno controlado. Es menos útil en un entorno no controlado, como el entorno de un cliente. Cuanto más ocupada esté la máquina, o cuanto más esté haciendo un programa, más difícil es tener una imagen precisa del estado actual de un programa.

Cyber Crucible comenzó con un uso intensivo del escaneo de memoria, tratando de tomar muchas “fotografías” de los estados de memoria de los programas. A medida que nos volvimos más rápidos, encontramos múltiples cuellos de botella en cuanto a la rapidez con la que se podían tomar instantáneas de memoria, y a través de múltiples programas.

Preocupaciones sobre el consumo de recursos de CPU o GPU

Este escaneo de memoria puede ser extremadamente intensivo en recursos, y puede tomar algo de tiempo para programas grandes. Cuando Cyber Crucible utilizaba el escaneo de memoria, en lugar del monitoreo inteligente del comportamiento de la memoria, se dedicaba un núcleo completo de CPU al escaneo.

Intel tiene una tecnología innovadora llamada Intel TDT AMS, que está disponible en arquitecturas específicas basadas en Intel con una GPU Intel integrada.

Escanea la memoria de un programa cuando se inicia por primera vez, utilizando la GPU para ayudar a no sobrecargar la CPU. Además, la GPU tiene un mejor rendimiento para el tipo de escaneo que se realiza (más sobre esto más adelante).

Metodología de escaneo

El escaneo de memoria se realiza para buscar códigos o “cadenas” específicas. Estas suelen ser las mismas piezas de datos que las herramientas de antivirus/EDR buscan en archivos potencialmente maliciosos.

El escaneo de memoria específicamente es más limitado que algunas de las búsquedas complejas que se pueden realizar contra un archivo.

Por lo tanto, el escaneo, por defecto, busca elementos ya conocidos, como malware antiguo, pero su capacidad para hacerlo es limitada.

Cyber Crucible solía escanear segmentos de memoria en busca de material criptográfico, pero resultó demasiado limitado para los entornos modernos, y desarrollamos un método diferente para observar comportamientos, utilizando lógica basada en el kernel.

Metodologías basadas en comportamiento vs. escaneo

Cyber Crucible migró de una metodología basada en escaneo a una basada en comportamiento por un par de razones.

Desventajas del escaneo

La primera es que se necesitaba usar una imagen en vivo en todos los programas. Las instantáneas dificultaban enormemente completar incluso un análisis limitado antes de que se

necesitara otra. Se estaba perdiendo demasiada actividad del programa, casi como tener una cámara de video funcionando a una velocidad de fotogramas demasiado baja. Esto se agravaba con la necesidad de monitorear todos los programas simultáneamente. Los servidores o estaciones de trabajo con mucha actividad, posiblemente las computadoras que más necesitaban monitoreo y que eran las más importantes, terminaban con el menor porcentaje de vigilancia realizada. Independientemente de la empresa o el producto, las limitaciones tecnológicas eran evidentes e inviables.

La segunda es que, más allá del consumo de recursos, el escaneo requiere, por defecto, buscar **algo conocido**. Por lo tanto, el escaneo de memoria debe basarse en firmas. Además, el escaneo de firmas debe ser un subconjunto limitado del escaneo de firmas más completo que se encuentra en el escaneo de antivirus basado en archivos. Los atacantes ya han eludido las defensas basadas en firmas de archivos durante años; hay muy pocas probabilidades de que los atacantes sean descubiertos por un sistema basado en firmas menos capaz.

Ventajas del análisis de comportamiento

El monitoreo de memoria basado en comportamiento de Cyber Crucible no requiere firmas, y aprovecha el acceso a la memoria del kernel para rastrear los comportamientos de todos los programas simultáneamente. Vamos a desglosar esto.

Primero, el monitoreo de comportamiento significa que el malware nuevo, o el malware que ha sido editado por atacantes para parecer nuevo, puede descubrirse fácilmente en tiempo real. El escaneo basado en firmas, ya sea de memoria o de archivos, depende de haber visto antes esa pieza específica de malware. Los atacantes han estado eludiendo las defensas basadas en firmas durante años, a pesar de contar con un conjunto completo de técnicas de escaneo contra archivos. Las técnicas de escaneo disponibles contra la memoria son un subconjunto de eso, y son incluso menos efectivas.

Segundo, el novedoso monitoreo del comportamiento de memoria de Cyber Crucible puede mantenerse en tiempo real en todos los programas simultáneamente, durante toda la ejecución de los programas. El escaneo de memoria sufre de graves cuellos de botella en los recursos que solo empeoran a medida que se ejecutan más programas. Los atacantes normalmente ejecutan docenas de muestras de malware a la vez, ahogando el análisis de escaneo de memoria debido a la velocidad y cantidad de programas a examinar a la vez, sin importar si se trata de un día cero o malware antiguo.

Qué sucede cuando CC detecta ataques de memoria

En primer lugar, es muy importante señalar que las alteraciones de memoria ocurren por varias razones distintas:

1. Ataques en memoria
2. Interacción entre programas
3. Errores de software
4. Exploits

Por lo tanto, una inyección de proceso o una técnica de alteración de memoria utilizada, incluso entre programas, no es necesariamente una acción maliciosa.

Si un programa que ha sufrido un evento de memoria probablemente malicioso (es decir, el #1 o el #4 mencionados anteriormente) está involucrado en una respuesta automatizada por parte de Cyber Crucible, se produce una actividad de recopilación automatizada.

La parte de la memoria que fue alterada en un programa, así como el estado “anterior”, se guarda en un formato estructurado de Cyber Crucible llamado .ccdiff

Esa porción inyectada o alterada de la memoria se guarda y está disponible para su descarga en el portal web de Cyber Crucible.

¿Por qué no se envía esto automáticamente a Virustotal (etc.) para su análisis?

En algunas circunstancias, el código malicioso utilizado en un ataque contiene variables que son únicas para ese ataque, específicamente para una víctima.

Enviar este código malicioso a un motor de análisis público como Virustotal representaría una divulgación de información no deseada que identifica a la víctima ante la comunidad de seguridad.

¿Por qué sería útil enviar estos datos binarios a Virustotal (etc.)?

Las herramientas de antivirus y EDR basadas en firmas buscan cadenas o bytes de código vistos en ataques anteriores. Incluso si este malware nunca se ha visto antes, es posible que, después de un tiempo, sea conocido por la comunidad de seguridad. Estas herramientas de escaneo generalmente no necesitan que el código esté en un binario con el formato adecuado para señalar una asociación con una pieza de malware conocida.

¿Qué es la IA Genética y en qué se diferencia de un LLM?

Respuesta breve: La IA Genética es la IA de seguridad desarrollada a medida por Cyber Crucible. Modela cómo se comportan los programas — en memoria, entre procesos y frente a archivos — y detecta desviaciones maliciosas en tiempo real. No es un modelo de lenguaje de gran tamaño (LLM); los LLM analizan patrones de texto, lo cual es un problema completamente distinto.

Qué la hace diferente

- **Diseñada específicamente para la seguridad**, no adaptada de un modelo de texto de propósito general.
- **Basada en comportamiento, no en firmas** — sin firmas, sin fuentes de inteligencia de amenazas, sin esperar a que alguien más descubra primero el ataque.
- **Totalmente autónoma** — las herramientas tradicionales dependen de la colaboración entre humano y máquina; la IA Genética decide y actúa por sí sola.
- **Previa a la ejecución** — bloquea el comportamiento malicioso antes de que se ejecute, en lugar de describirlo después.

IA vertical

La IA Genética es un ejemplo de IA vertical: inteligencia artificial personalizada para los desafíos específicos de un sector, en lugar de un razonamiento de propósito general. En el ámbito de la seguridad, esa especialización es lo que permite tomar una decisión sobre la intención en menos de 200 milisegundos.

Los resultados son medibles: 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.

¿Cómo toma Cyber Crucible una decisión en menos de 200 milisegundos?

Respuesta breve: Ejecuta un ciclo de tres pasos —Detect-Decide-Respond— completamente en el endpoint, a nivel de kernel, utilizando un motor patentado de análisis de edge computing. Como nada tiene que viajar hacia un servicio en la nube y regresar, y no hay ningún humano en el ciclo, todo el proceso se completa en menos de 200 milisegundos.

El ciclo

1. **Detect (Detectar)** — sensores novedosos a nivel de kernel identifican patrones de comportamiento malicioso en la memoria, la actividad de procesos y el acceso a archivos.
2. **Decide (Decidir)** — el modelo de comportamiento mejorado con IA evalúa instantáneamente la intención del acceso a datos del programa.
3. **Respond (Responder)** — si la intención es maliciosa, el programa es interceptado y suspendido.

Por qué el procesamiento local no es negociable

Enviar telemetría a un servidor en la nube, esperar el análisis y recibir una respuesta añade milisegundos que el atacante aprovechará con gusto. Para dimensionar la escala involucrada: esta decisión ocurre más rápido de lo que un impulso nervioso viaja desde el ojo hasta el cerebro. Cualquier arquitectura con un viaje de ida y vuelta por la red en la ruta crítica ya ha perdido esa carrera.

La telemetría bruta se sigue copiando posteriormente a los servidores para fines de evidencia y enriquecimiento, pero la decisión de protección nunca espera por ello.

¿Qué es Rogue Process Prevention (RPP)?

Respuesta breve: Rogue Process Prevention es el enfoque patentado de Cyber Crucible que consiste en interceder contra un proceso malicioso en lugar de limitarse a observarlo. Ya sea que se trate de un ransomware que comienza a cifrar datos o de un infostealer que busca obtener un token de sesión, el proceso se detiene antes de que los datos salgan del dispositivo.

Observar frente a intervenir

La mayor parte del mercado de endpoints está diseñada para observar. RPP está diseñado para intervenir. La distinción es la diferencia entre una cámara de seguridad que graba un robo para su revisión posterior y un guardia que detiene al ladrón antes de que llegue a la bóveda.

Contra qué protege

RPP aborda las tres amenazas principales de manera conjunta, ya que los ataques modernos suelen tener múltiples vectores:

- **Robo de identidad** — acceso a claves, tokens y contraseñas, a menudo sin ninguna interacción por parte del empleado.
- **Robo de información** — exfiltración masiva de datos a velocidad de máquina.
- **Ransomware** — análisis de comportamiento junto con tecnología de engaño para hackers que impide el cifrado incluso de un solo archivo.

Una defensa centrada exclusivamente en el ransomware pasaría por alto las dos primeras etapas, que es donde suele comenzar el daño real.

¿Qué papel desempeña la analítica de memoria en toda la plataforma?

Respuesta breve: La analítica de memoria es la base sobre la que se construye todo lo demás. Es el sensor y la fuente analítica que alimenta todo el modelado conductual: la protección de datos, la protección de identidad y FortressAI por igual. Cada decisión que toman esas capacidades se apoya en saber qué está haciendo realmente la memoria de un programa.

Una base, tres capacidades

Cyber Crucible rastrea el estado de la memoria de un programa a lo largo de todo su ciclo de vida, desde el inicio hasta el final, incluidas sus interacciones en memoria con otros programas. Ese único flujo de análisis es lo que hace posible el resto:

- **Protección de datos** — ¿este proceso se comporta como algo que está a punto de cifrar o exfiltrar?
- **Protección de identidad** — ¿este proceso es el propietario legítimo de la credencial a la que está accediendo, o ha sido tomado por otro?
- **FortressAI** — ¿esta herramienta de IA sigue siendo la misma, y su acceso a datos está dentro de la política?

Estos no son tres motores separados con tres conjuntos de sensores separados. Son tres políticas expresadas sobre el mismo modelo conductual de memoria.

Por qué la memoria es la base correcta

Los ataques modernos cada vez con más frecuencia no dejan ningún archivo que inspeccionar. El código se inyecta en procesos de confianza y se construye en memoria, por lo que el único lugar donde la actividad es visible es el estado de memoria de un programa en ejecución. Una defensa que no puede ver ahí es ciega ante toda esa clase de ataques.

El análisis de memoria también responde a la pregunta que más importa antes de conceder acceso a cualquier cosa sensible: **¿sigue siendo este programa lo que dice ser?** Una aplicación de confianza cuyo proceso o bibliotecas han sido manipulados ya no es confiable, y solo la inspección a nivel de memoria lo revela.

No es escaneo de memoria

Esto es seguimiento conductual en vivo, no escaneo periódico. El escaneo copia un bloque de memoria y analiza la instantánea: una fotografía en lugar de una grabación en vivo. En una máquina con mucha actividad, el estado del programa puede ser completamente diferente un momento después, lo que hace que las instantáneas sean útiles para la forense controlada, pero mucho más débiles para la defensa en tiempo real.

> Consulte también los artículos existentes de *Analítica de memoria* para conocer el detalle técnico de en qué se diferencia esto del escaneo de memoria, y qué sucede cuando se detecta un ataque de memoria.

¿Por qué Cyber Crucible opera a nivel de kernel?

Respuesta breve: El kernel es la capa más profunda y privilegiada del sistema operativo. Operar allí permite a Cyber Crucible ver y controlar toda la actividad del sistema, identificar la intención de un proceso malicioso y detenerlo en milisegundos, además de hacer que la herramienta sea mucho más difícil de manipular o deshabilitar para un atacante.

Qué proporciona el acceso a nivel de kernel

- **Visibilidad completa** del comportamiento de la memoria, el comportamiento de los procesos y el acceso a archivos, incluida la actividad que nunca toca el disco.
- **La capacidad de actuar**, no solo de observar: los permisos pueden revocarse y los procesos pueden suspenderse en el momento en que se detecta la intención maliciosa.
- **Resiliencia**, porque los atacantes inevitablemente intentarán deshabilitar la propia herramienta de seguridad.

Sobre las críticas al desarrollo a nivel de kernel

Algunos actores establecidos y analistas desaconsejan el desarrollo a nivel de kernel. La ingeniería de kernel es genuinamente difícil y costosa, precisamente por eso gran parte del mercado la evitó, y por eso los endpoints quedaron dependientes de la telemetría en el espacio de usuario, que los ataques a nivel raíz pueden neutralizar.

La calidad y la resistencia a la manipulación de Cyber Crucible están validadas a través del Windows Hardware Compatibility Program (WHCP) de Microsoft.

¿Qué significa que Cyber Crucible sea independiente de las bibliotecas de Windows?

Respuesta breve: Cyber Crucible no depende de las bibliotecas del sistema Windows que está protegiendo. Fue desacoplado deliberadamente de ellas, de modo que cuando los atacantes comprometen, bloquean o secuestran esas bibliotecas, Cyber Crucible sigue funcionando, sigue aplicando controles y sigue informando.

Por qué se construyó así

En la segunda mitad de 2025, los atacantes cambiaron su enfoque hacia la infección de bibliotecas centrales de Windows en memoria. Dado que casi todos los proveedores de seguridad dependen de esas mismas bibliotecas para funcionar, ese único movimiento dejó ciega a una gran parte del mercado.

Cyber Crucible dedicó aproximadamente nueve meses a inventar sensores novedosos y un esfuerzo adicional de ingeniería de cinco meses para eliminar por completo las dependencias de las bibliotecas de Windows, incluida la eliminación de las dependencias volátiles de MiniFilter.

Cómo se ve esto a nivel arquitectónico

La arquitectura ejecuta siete subsistemas totalmente independientes dentro del kernel, incluidos su propio cifrado aislado, redes, análisis de metadatos y seguimiento de eventos. Una fase adicional de este trabajo reemplaza la biblioteca estándar de redes de Windows (`ws32_2.dll`) por una pila de red personalizada a nivel de kernel.

Comprobado en el campo

Con dos clientes, los atacantes bloqueados por el motor de decisión exclusivo del kernel intentaron luego impedir que se notificara al cliente atacando la capa de red del sistema operativo. Dado que las comunicaciones de red de Cyber Crucible son independientes del sistema operativo, la protección se mantuvo y la inteligencia igualmente llegó al cliente.

¿Funciona Cyber Crucible sin conexión o en entornos aislados de la red (air-gapped)?

Respuesta breve: Sí. Todo el análisis y la aplicación de medidas se realizan localmente en el endpoint, por lo que Cyber Crucible ofrece protección completa sin conexión a internet, sin servicio en la nube y sin señal de ningún tipo. Es adecuado para redes aisladas (air-gapped), entornos OT e ICS, y operaciones de campo desconectadas.

Por qué funciona sin conectividad

Dado que el ciclo Detect-Decide-Respond se ejecuta completamente en el kernel del dispositivo, nunca se requiere conectividad para tomar una decisión de protección. La telemetría se sincroniza de forma oportunista para fines de evidencia e informes cuando hay conexión disponible, pero la protección no depende de ello.

Comprobado en el mar

En un despliegue marítimo, Cyber Crucible operó de forma independiente durante travesías de varios días sin personal de TI a bordo, se adaptó a enlaces satelitales de baja capacidad e intermitentes con interrupciones por mal tiempo, y conservó registros forenses detallados durante largos períodos sin conexión para su análisis posterior a la travesía. Los sistemas protegidos abarcaron desde terminales recreativos orientados a los pasajeros hasta controles marítimos de misión crítica.

Esta es también la razón por la que el mismo agente resulta adecuado para plantas de fabricación, flotas logísticas y otros entornos periféricos donde la conectividad a la nube es poco confiable o está deliberadamente ausente.

¿Por qué Cyber Crucible responde suspendiendo un proceso en lugar de apagar el sistema?

Respuesta breve: Suspender únicamente el proceso malicioso —a veces llamado Congelamiento Selectivo (Selective Freeze)— neutraliza el ataque mientras todo lo demás sigue funcionando. No hay bloqueo total del sistema, no hay reinicio y no hay interrupción del trabajo legítimo.

El problema con las respuestas de tipo bloqueo

Cuando muchas herramientas de detección identifican una anomalía, las opciones de contención son poco precisas: aislar la máquina o desconectarla. En un hospital, una planta o una operación de negociación (trading), esa respuesta puede ser casi tan perjudicial como el propio ataque.

Lo que logra la suspensión selectiva

El ataque se suspende en memoria. Las transacciones legítimas nunca se detienen. En la implementación de servicios financieros donde se interceptaron casi 10,000 procesos maliciosos, las transacciones SQL legítimas continuaron sin interrupción durante todo el proceso; las operaciones nunca se detuvieron.

Esto es lo que hace que la prevención sea compatible con la continuidad del negocio, en lugar de representar una disyuntiva frente a ella.