



2

Sistemas operativos y redes para ciberdefensa

2.1 ARQUITECTURA DE LOS SISTEMAS OPERATIVOS

Cuando entrevisto a un candidato para mi equipo de detección, una de mis preguntas favoritas es sencilla en apariencia: ¿qué ocurre exactamente, a nivel de sistema operativo, cuando se hace doble clic sobre un ejecutable? La respuesta que dé separa con bastante fiabilidad a quien ha memorizado herramientas de quien entiende de verdad lo que estas herramientas están observando.

Un sistema operativo es, en esencia, un gestor de recursos limitados —CPU, memoria, disco, red— que debe repartir entre múltiples programas que compiten por ellos, manteniendo al mismo tiempo el aislamiento necesario para que ninguno interfiera con los demás sin autorización. Toda la seguridad del endpoint, desde el control de acceso hasta la detección de comportamiento anómalo, se construye sobre esta idea de aislamiento y arbitraje de recursos.

Esta función de arbitraje se apoya en un componente central, el núcleo o kernel, que se ejecuta con el máximo nivel de privilegio del procesador y que es el único con acceso directo al hardware. Todo lo demás —aplicaciones de usuario, servicios, controladores en modo usuario— opera en un anillo de privilegio inferior y debe solicitar al núcleo, mediante llamadas al sistema, cualquier operación que requiera acceso a un recurso protegido. Esta frontera entre modo núcleo y modo usuario es, en la práctica, la primera línea de defensa de cualquier sistema operativo moderno.

Los cinco subsistemas que desarrollo a continuación —procesos, hilos, memoria, sistema de archivos y entradas/salidas— no son un repaso académico gratuito. Son las cinco superficies sobre las que un atacante deja evidencia, y por tanto las cinco superficies que un EDR o un analista humano deben saber interpretar.

Conviene además distinguir, desde el primer momento, entre dos conceptos que se confunden con frecuencia incluso entre analistas con cierta experiencia: el modo de ejecución de un programa y su nivel de privilegio de usuario. El primero determina si el código corre en el anillo protegido del núcleo o en el espacio, más restringido, de las aplicaciones de usuario; el segundo determina qué puede hacer ese código en nombre de una cuenta concreta —administrador, usuario estándar, cuenta de servicio— dentro del sistema de permisos del propio sistema operativo. Un exploit de escalado de privilegios suele atacar precisamente la frontera entre ambos, elevando una ejecución que nace con privilegios de usuario estándar hasta alcanzar privilegios de administrador o de sistema.

2.1.1 Procesos

Un proceso es la unidad básica de ejecución: un programa en memoria con su propio espacio de direcciones, sus propios recursos asignados y un identificador único, el PID, que lo distingue de cualquier otro proceso en ejecución en ese instante. Cada proceso nace de otro proceso mediante una llamada de creación —`fork()` en sistemas Unix, `CreateProcess()` en Windows— y esta relación de filiación queda registrada por el sistema operativo.

Esa relación padre-hijo es, para un analista de Blue Team, una de las señales más valiosas que existen. Un árbol de procesos legítimo sigue patrones predecibles: `explorer.exe` lanza aplicaciones de usuario, `services.exe` lanza servicios del sistema, un servidor web lanza procesos hijos relacionados con su función. Cuando ese patrón se rompe —un proceso de Office lanzando `powershell.exe`, o `powershell.exe` lanzando `cmd.exe` que a su vez descarga un binario de internet— el árbol de procesos cuenta una historia que ningún antivirus basado en firmas es capaz de contar.

Cada proceso, además, mantiene un ciclo de vida con estados bien definidos: nuevo, listo, en ejecución, bloqueado y terminado. El planificador del sistema operativo decide en cada instante qué proceso listo pasa a ejecución, en función de algoritmos de prioridad que, aunque parezcan alejados de la seguridad, explican por ejemplo por qué un proceso de

criptominería malicioso configurado para consumir CPU de forma agresiva puede degradar de forma perceptible el rendimiento general de una máquina, convirtiéndose así en un indicio indirecto de compromiso.

Recomiendo a mi equipo revisar siempre tres atributos de cada proceso sospechoso: su línea de comandos completa, su proceso padre y la ruta desde la que se ejecuta el binario. La inmensa mayoría del malware que hemos investigado en los últimos años se ejecuta desde ubicaciones atípicas —carpetas temporales, rutas de usuario, directorios de descarga— precisamente porque las rutas legítimas del sistema suelen estar protegidas o monitorizadas con mayor rigor.

Recomendación: configurar el registro de creación de procesos (event ID 4688 en Windows, con línea de comandos habilitada mediante la directiva de auditoría correspondiente, o el subsistema auditd en Linux) es una de las inversiones de menor coste y mayor retorno en visibilidad que puede hacer cualquier SOC que empiece a madurar su capacidad de detección.

2.1.2 Hilos

Dentro de un mismo proceso pueden ejecutarse varios hilos, unidades de ejecución más ligeras que comparten el espacio de memoria y los recursos del proceso padre, pero que mantienen su propia pila y su propio contador de programa. La ventaja de los hilos es la concurrencia: un servidor web puede atender múltiples peticiones simultáneas sin necesidad de crear un proceso completo por cada una, con el ahorro de recursos que eso supone.

Desde la óptica defensiva, los hilos importan sobre todo por una técnica de evasión muy extendida: la inyección de código. Un atacante que ha comprometido un proceso puede crear un hilo remoto dentro de otro proceso legítimo —una técnica conocida como remote thread injection— y ejecutar así código malicioso bajo la identidad y los privilegios de un proceso de confianza, evitando de paso muchas reglas de detección basadas exclusivamente en el nombre del proceso.

He gestionado incidentes en los que el malware se inyectaba sistemáticamente en explorer.exe o en procesos de navegador, precisamente porque ningún analista sospecharía de una conexión de red saliente originada, en apariencia, por el propio explorador de Windows. Detectar esta

técnica exige herramientas capaces de monitorizar llamadas a funciones como `CreateRemoteThread` o `WriteProcessMemory`, no un simple listado de procesos en ejecución.

Existen variantes de esta técnica que conviene distinguir, porque cada una deja un patrón forense distinto: la inyección clásica de DLL, el vaciado de proceso o `process hollowing`, en el que se crea un proceso legítimo en estado suspendido y se sustituye su contenido en memoria antes de reanudarlo, y las técnicas más recientes de manipulación de hilos mediante APC (llamadas a procedimiento asíncronas), que evitan por completo la necesidad de escribir memoria ejecutable adicional. Un EDR maduro debe ser capaz de correlacionar estas tres variantes, y no limitarse a vigilar una sola de ellas.

2.1.3 Memoria

La memoria de un sistema se organiza en espacios de direcciones aislados por proceso, gestionados por el sistema operativo mediante mecanismos de paginación que traducen direcciones virtuales a direcciones físicas reales. Este aislamiento es la primera línea de defensa contra que un proceso comprometido pueda leer o corromper la memoria de otro.

Sin embargo, ese aislamiento no es absoluto ni infalible, y buena parte del malware más sofisticado que hemos analizado en el SOC vive precisamente ahí, en memoria, sin llegar a escribir nunca un fichero en disco. Hablamos de malware *fileless*: código que se ejecuta directamente en el espacio de memoria de un proceso legítimo —a menudo mediante PowerShell o mediante técnicas de `process hollowing`— y que desaparece sin dejar rastro en disco cuando el sistema se reinicia.

Esta realidad ha cambiado radicalmente el enfoque de detección de cualquier SOC serio. Ya no basta con escanear el disco en busca de ficheros maliciosos; es imprescindible contar con capacidad de análisis de memoria en vivo, ya sea mediante volcados forenses con herramientas como Volatility o mediante sensores EDR capaces de inspeccionar el contenido de la memoria de procesos en ejecución.

Los sistemas operativos modernos incorporan además mecanismos de protección de memoria pensados específicamente para dificultar la explotación de vulnerabilidades, y que todo analista debería saber nombrar y explicar cuando revisa la configuración de bastionado de un equipo: la aleatorización del espacio de direcciones (ASLR), que impide predecir en qué

dirección de memoria se cargará un módulo determinado, y la prevención de ejecución de datos (DEP o NX bit), que marca como no ejecutables las regiones de memoria destinadas a datos. Ambos mecanismos elevan sustancialmente el coste de desarrollar un exploit fiable, aunque ninguno de los dos es, por sí solo, una barrera insalvable.

Riesgo: dimensionar la estrategia de detección de un SOC exclusivamente sobre análisis de disco es, a estas alturas, un error que deja a la organización ciega frente a una parte creciente y cada vez más relevante del malware que realmente se enfrenta en producción.

2.1.4 Sistema de archivos

El sistema de archivos organiza cómo se almacenan, nombran y localizan los datos en un dispositivo de almacenamiento persistente. Más allá de la jerarquía de carpetas visible para el usuario, cada sistema de archivos mantiene metadatos —permisos, propietario, marcas de tiempo de creación, modificación y acceso— que resultan esenciales para cualquier investigación forense.

Las marcas de tiempo, en particular, son uno de los primeros elementos que reviso al reconstruir la cronología de un incidente. Cuando se creó un fichero, cuándo se modificó por última vez y cuándo se accedió a él por primera vez permiten establecer con bastante precisión la secuencia de acciones de un atacante, siempre que se tenga presente que estas marcas también pueden manipularse mediante técnicas de timestomping.

Los permisos de acceso —el modelo de usuario, grupo y otros en sistemas Unix, o las listas de control de acceso discrecionales en Windows— determinan quién puede leer, escribir o ejecutar cada fichero, y constituyen la base técnica sobre la que se apoya cualquier estrategia seria de privilegio mínimo. Un hallazgo recurrente en mis auditorías es encontrar ficheros de configuración con permisos de escritura abiertos a todos los usuarios, una debilidad que cualquier atacante con acceso mínimo al sistema explota sin esfuerzo.

Conviene distinguir también, a efectos de detección, entre metadatos del sistema de archivos y contenido del fichero. Un atacante puede alterar las marcas de tiempo visibles desde el propio sistema de archivos, pero rara vez consigue eliminar por completo el rastro que deja en estructuras

de nivel inferior, como el registro de transacciones del sistema de archivos (el journal, en sistemas como NTFS o ext4) o los registros de auditoría del sistema operativo, que actúan como fuentes de verificación independientes frente a una posible manipulación.

La siguiente tabla resume, a modo de referencia rápida, los metadatos forenses más relevantes de un sistema de archivos y su utilidad principal en una investigación de incidentes.

Metadato	Qué indica	Utilidad forense principal
Marca de creación	Cuándo se generó el fichero en ese volumen	Acotar el inicio de la actividad del atacante
Marca de modificación	Última vez que se alteró el contenido	Detectar manipulación posterior a la creación
Marca de acceso	Última vez que se leyó el fichero	Confirmar ejecución o consulta, con matices por sistema
Propietario y permisos	Quién puede leer, escribir o ejecutar	Detectar configuraciones de privilegio excesivo
Registro de transacciones (journal)	Operaciones recientes sobre el volumen	Contrastar manipulación de marcas de tiempo visibles

Tabla 2.1. Metadatos forenses del sistema de archivos y su utilidad en la investigación.

Error: confiar en las marcas de tiempo de un sistema de archivos como única fuente de verdad temporal en una investigación, sin contrastarlas con registros de eventos independientes, es un error que he visto invalidar conclusiones completas de análisis forense.

2.1.5 Entradas y salidas

El subsistema de entradas y salidas gestiona la comunicación entre el sistema operativo y sus periféricos: discos, tarjetas de red, dispositivos USB, teclado y pantalla. Toda esta comunicación se canaliza a través de

controladores —drivers— que actúan como intermediarios entre el hardware y el resto del sistema.

Desde la perspectiva de un Blue Team, los controladores merecen una atención especial porque se ejecutan habitualmente con los privilegios más altos que ofrece el sistema, en modo núcleo (kernel mode). Un controlador malicioso o vulnerable compromete, por definición, todo el sistema, y no solo el espacio de un proceso aislado. Los ataques que instalan rootkits a nivel de kernel explotan precisamente esta posición de privilegio.

Los dispositivos de almacenamiento extraíble merecen mención aparte por su papel histórico en incidentes de exfiltración de información y de introducción de malware en redes aisladas. Cualquier política de bastionado sería que dirija debe contemplar el control granular de qué dispositivos USB pueden conectarse, y con qué permisos, en cada perfil de equipo dentro de la organización.

Un aspecto que suelo destacar en mis auditorías de madurez es la firma de controladores. Windows, por ejemplo, exige desde hace años que los controladores en modo núcleo estén firmados digitalmente por un fabricante verificado, precisamente para dificultar la instalación de rootkits no autorizados. Deshabilitar esta exigencia de firma, algo que ocurre con más frecuencia de la deseable en entornos de pruebas que después migran a producción sin revertir el cambio, elimina de un plumazo una de las barreras de bastionado más eficaces contra el compromiso del kernel.

Desde el punto de vista de la monitorización, conviene recordar que buena parte de la telemetría que alimenta a un EDR moderno procede precisamente de este subsistema: el registro de qué dispositivos se conectan y desconectan, qué volumen de datos se transfiere a través de cada interfaz y qué controladores se cargan durante el arranque o en caliente. Un salto inusual en el volumen de escritura hacia un dispositivo USB, fuera del patrón habitual de un puesto de trabajo, es con frecuencia el primer indicio observable de un intento de exfiltración de datos.

2.2 REDES

Si el sistema operativo es el terreno donde el malware se ejecuta, la red es el terreno donde se mueve. Ningún atacante serio se conforma con comprometer una única máquina; el objetivo casi siempre es progresar hacia activos de mayor valor —controladores de dominio, bases de datos,

sistemas financieros— y esa progresión exige atravesar una red que, bien diseñada, debería oponer resistencia en cada salto.

Cuando reviso la arquitectura de red de una organización antes de asumir la responsabilidad de su seguridad, mi primera pregunta nunca es qué firewall tienen instalado, sino cómo está segmentada la red y qué visibilidad existe sobre el tráfico que circula entre segmentos. Un firewall perimetral de última generación es de escaso valor si, una vez dentro, un atacante puede desplazarse libremente entre servidores críticos sin encontrar ningún control intermedio.

Los apartados que siguen recorren los fundamentos de redes imprescindibles para un analista de Blue Team: el modelo de referencia OSI, la pila de protocolos TCP/IP sobre la que funciona internet, y la segmentación mediante VLAN como primera herramienta de contención frente al movimiento lateral.

2.2.1 Modelo OSI

El modelo OSI describe la comunicación en red como una pila de siete capas, cada una responsable de una función concreta y cada una comunicándose únicamente con las capas inmediatamente superior e inferior. Aunque en la práctica ningún sistema implementa el modelo OSI de forma literal, sigue siendo el marco conceptual de referencia para razonar sobre dónde ocurre exactamente cada tipo de problema o de ataque.

De abajo arriba, las capas son: física, enlace de datos, red, transporte, sesión, presentación y aplicación, representadas en la Figura 2.1. Un analista experimentado aprende a situar mentalmente cada herramienta y cada tipo de ataque en su capa correspondiente, porque eso determina qué tipo de control aplica y qué tipo de evidencia puede esperar recoger.

Un ataque de suplantación ARP, por ejemplo, vive en la capa de enlace de datos y se combate con controles como la inspección dinámica de ARP en los switches. Un ataque de inyección SQL, en cambio, vive en la capa de aplicación y ningún control de capa 2 tendrá el menor efecto sobre él. Esta capacidad de ubicar un problema en su capa correcta es lo que distingue a un analista que diagnostica con precisión de uno que aplica controles genéricos sin entender por qué funcionan o dejan de funcionar.

La tabla siguiente recoge, capa a capa, un protocolo representativo y un ejemplo de ataque característico, como referencia rápida para el análisis diario.

Capa	Función principal	Ejemplo de ataque característico
7. Aplicación	Servicios directamente usados por el usuario (HTTP, DNS, SMTP)	Inyección SQL, phishing, abuso de API
6. Presentación	Formato, cifrado y compresión de los datos	Downgrade de cifrado, manipulación de certificados
5. Sesión	Establecimiento y control de sesiones de comunicación	Secuestro de sesión (session hijacking)
4. Transporte	Entrega fiable (TCP) o no fiable (UDP) de segmentos	Escaneo de puertos, inundación SYN
3. Red	Direccionamiento y enrutamiento de paquetes (IP)	Suplantación de IP, ataques de enrutamiento
2. Enlace de datos	Comunicación entre dispositivos de un mismo segmento	Suplantación ARP, saturación de tabla MAC
1. Física	Transmisión de bits sobre el medio físico	Intervención física del cableado, jamming

Tabla 2.2. Capas del modelo OSI, protocolos representativos y ataques característicos.

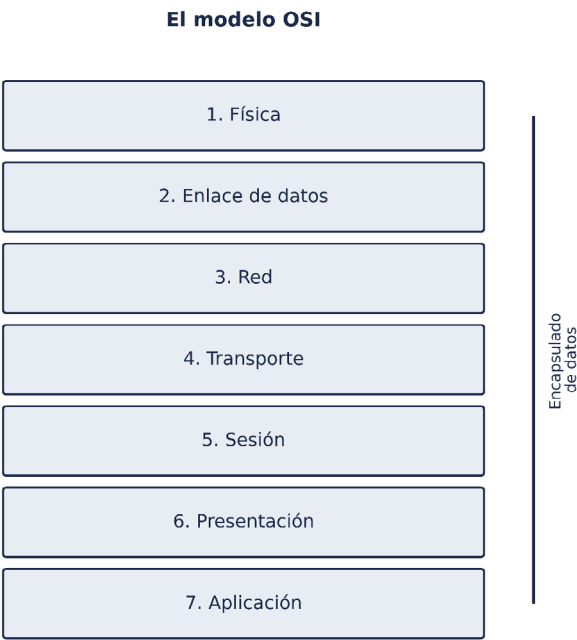


Figura 2.1. El modelo OSI y sus siete capas de comunicación, con el sentido de encapsulado de datos entre capas.

2.2.2 TCP/IP

En la práctica, la inmensa mayoría del tráfico que un analista de SOC investiga circula sobre la pila TCP/IP, un modelo más simplificado y de aplicación real que condensa las siete capas OSI en cuatro: acceso a red, internet, transporte y aplicación. Entender esta pila con solidez es, sin exagerar, el requisito mínimo para leer con criterio una captura de tráfico o un registro de firewall.

La capa de internet, gobernada por el protocolo IP, se encarga del direccionamiento y el enrutamiento de paquetes entre redes distintas. La capa de transporte introduce los dos protocolos que todo analista maneja a diario: TCP, orientado a conexión y con garantías de entrega, y UDP, sin conexión y sin esas garantías, pero más rápido y por tanto preferido por servicios como DNS o la transmisión de vídeo en tiempo real.

Uno de los ejercicios que más recomiendo a los analistas junior de mi equipo es acostumbrarse a leer el proceso de establecimiento de conexión TCP —el conocido saludo de tres vías, SYN, SYN-ACK, ACK— directamente en una captura de Wireshark, hasta que resulte tan natural como leer una frase. Un escaneo de puertos, por ejemplo, se reconoce con facilidad por un patrón de paquetes SYN sin respuesta ACK completa, repetidos contra múltiples puertos de destino en un intervalo de tiempo muy corto.

Cada segmento TCP y cada datagrama UDP se identifican, además del par de direcciones IP origen y destino, mediante un puerto de origen y un puerto de destino, cuya combinación con el protocolo forma lo que se conoce habitualmente como una tupla de sesión. Esta tupla —IP origen, puerto origen, IP destino, puerto destino, protocolo— es la unidad mínima que registra cualquier firewall con estado, y comprender su estructura es indispensable para leer con soltura un registro de conexión o construir una regla de bloqueo precisa.

Los números de puerto, por su parte, no se asignan de forma arbitraria. Los puertos comprendidos entre el 0 y el 1023 se consideran puertos privilegiados o “bien conocidos” —el 80 y el 443 para tráfico web, el 22 para SSH, el 53 para DNS— y en sistemas Unix tradicionalmente solo pueden ser abiertos por procesos con privilegios de administrador, una convención que en sí misma constituye una señal de alerta cuando se observa un servicio no privilegiado escuchando en uno de estos puertos. Los puertos por encima de ese rango quedan disponibles para aplicaciones de usuario y para las conexiones salientes efímeras que el propio sistema operativo asigna de forma dinámica a cada nueva sesión.

Consejo: exigir a todo analista que se incorpore a mi equipo que sepa explicar, sin ayuda de ninguna herramienta, la diferencia práctica entre TCP y UDP y sus implicaciones para la detección, es uno de los filtros más eficaces que aplico en el proceso de selección.

2.2.3 VLAN

Una VLAN, o red de área local virtual, permite segmentar lógicamente una red física en múltiples dominios de difusión independientes, sin necesidad de desplegar cableado ni switches adicionales por cada segmento. El tráfico de cada VLAN queda aislado del de las demás salvo que exista un dispositivo de capa 3 —un router o un switch multicapa— configurado explícitamente para enrutarlo entre segmentos.

Esta capacidad de aislamiento lógico es una de las herramientas de contención más infravaloradas por organizaciones que aún no han alcanzado un nivel de madurez suficiente en su estrategia de defensa en profundidad. Separar la red de usuarios, la red de servidores, la red de gestión de infraestructura de seguridad y la red de invitados en VLAN distintas limita de forma drástica el alcance de cualquier movimiento lateral, incluso antes de introducir controles más sofisticados.

Un error que he encontrado con demasiada frecuencia en auditorías de organizaciones de tamaño medio es la existencia de una única VLAN plana que agrupa estaciones de trabajo, servidores críticos y, en ocasiones, hasta equipos de invitados conectados a la misma infraestructura de red. En ese escenario, comprometer un único portátil de usuario equivale, en la práctica, a tener visibilidad directa sobre activos que deberían estar completamente fuera de su alcance.

La identificación de cada VLAN se realiza mediante una etiqueta numérica —el VLAN ID, según el estándar IEEE 802.1Q— que se inserta en la trama Ethernet cuando esta atraviesa un enlace troncal (trunk) entre switches. Un error de configuración habitual, y que he visto explotado en ejercicios de red team, es dejar habilitado por defecto el protocolo de negociación de enlaces troncales en puertos de acceso de usuario final, lo que en determinadas condiciones permite a un atacante con acceso físico a un puerto de red inyectar tráfico etiquetado y saltar entre VLAN sin pasar por el dispositivo de enrutamiento previsto, una técnica conocida como VLAN hopping.

Riesgo: una VLAN, por sí sola, no sustituye a un firewall ni a una política de control de acceso entre segmentos; sin reglas de filtrado explícitas en el dispositivo de capa 3 que las interconecta, la segmentación lógica se convierte en una barrera meramente administrativa, fácilmente sorteable por un atacante que alcance ese punto de enrutamiento.

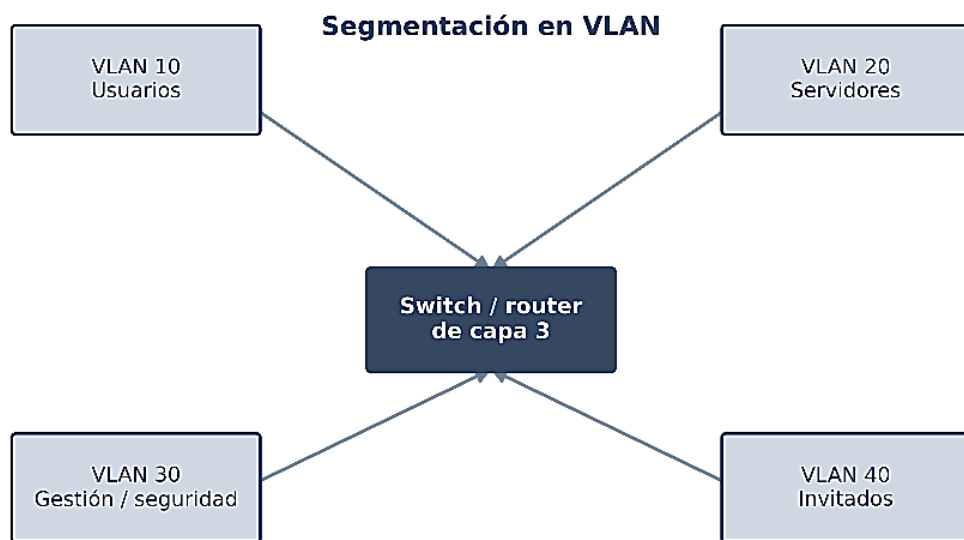


Figura 2.2. Segmentación de una red plana en VLAN independientes para usuarios, servidores, gestión de seguridad e invitados, con el router o switch de capa 3 como único punto de paso controlado entre segmentos.

2.2.4 Enrutamiento (routing)

Una vez segmentada la red en VLAN independientes, es necesario decidir cómo y cuándo se permite que el tráfico pase de un segmento a otro. Esa función corresponde al enrutamiento: el proceso mediante el cual un dispositivo de capa 3 —típicamente un router o un switch multicapa— examina la dirección IP de destino de cada paquete y decide por qué interfaz debe reenviarlo para acercarlo a su destino final.

Cada dispositivo de enrutamiento mantiene una tabla de rutas que asocia rangos de direcciones de red con la interfaz o el siguiente salto por el

que deben encaminarse los paquetes destinados a ellas. En redes pequeñas esta tabla puede construirse manualmente mediante rutas estáticas; en redes de mayor tamaño se recurre a protocolos de enrutamiento dinámico —OSPF o BGP son los más habituales en entornos corporativos— que permiten a los dispositivos intercambiar información sobre la topología de la red y adaptarse automáticamente a cambios o caídas de enlace.

Desde la óptica de un Blue Team, el enrutamiento no es solo una cuestión de conectividad, sino una herramienta de control activo. Las listas de control de acceso (ACL) aplicadas en un dispositivo de enrutamiento permiten permitir o denegar tráfico entre segmentos en función de direcciones de origen y destino, puertos y protocolos, convirtiendo cada punto de enrutamiento entre VLAN en un auténtico punto de control de seguridad, no solo de conectividad.

He revisado configuraciones de organizaciones que, tras invertir en una segmentación de red razonablemente bien diseñada sobre el papel, no habían aplicado ninguna ACL real entre segmentos, limitándose a enrutar todo el tráfico sin restricción alguna. En ese escenario, la segmentación existe únicamente a efectos de organización lógica del direccionamiento, pero no aporta ninguna barrera de contención real frente a un atacante que alcance el dispositivo de enrutamiento o cualquier segmento adyacente.

Íntimamente relacionada con el enrutamiento aparece la traducción de direcciones de red (NAT), mecanismo mediante el cual un dispositivo perimetral sustituye las direcciones IP privadas de la red interna por una o varias direcciones públicas antes de enviar el tráfico hacia internet, y realiza la operación inversa con el tráfico de respuesta. Más allá de su función original de ahorro de direcciones IPv4 públicas, el NAT tiene un efecto colateral de seguridad relevante: oculta la estructura interna de direccionamiento de la organización frente a un observador externo, dificultando el reconocimiento previo a un ataque dirigido.

Conviene no confundir, sin embargo, esta ocultación incidental con un control de seguridad diseñado como tal. He revisado organizaciones que confiaban en el NAT como si fuera un firewall, cuando en realidad su función es de traducción de direcciones, no de filtrado; un dispositivo de NAT sin reglas de filtrado explícitas asociadas puede seguir permitiendo tráfico entrante no deseado a través de reglas de redirección de puertos mal configuradas, un error que he encontrado más de una vez en revisiones de infraestructura de pequeñas delegaciones.

Riesgo: dar por hecho que la existencia de rutas entre segmentos, o la presencia de NAT en el perímetro, implica automáticamente un control de seguridad es un error de bulto; sin una política explícita de ACL y de filtrado revisada y probada periódicamente, el enrutamiento y la traducción de direcciones son pura conectividad, no defensa.

2.2.5 DNS (Sistema de Nombres de Dominio)

El DNS traduce nombres de dominio legibles por personas —como el nombre interno de un servidor corporativo o el dominio público de una organización— en direcciones IP que los equipos utilizan realmente para establecer conexiones. Su función parece, a primera vista, puramente operativa, pero en la práctica el DNS es una de las superficies más ricas en telemetría de seguridad de toda la infraestructura de red.

Prácticamente todo malware moderno necesita resolver, en algún momento de su ciclo de vida, el nombre de dominio de su servidor de mando y control. Esa resolución DNS, si se registra y se analiza correctamente, ofrece una de las señales más tempranas y fiables de compromiso disponibles para un SOC, mucho antes de que el atacante llegue a ejecutar acciones más ruidosas sobre el sistema.

Existen patrones de abuso del DNS que todo analista debería reconocer sin dudar. El más citado en la literatura de detección es la exfiltración de datos mediante túneles DNS, en la que un atacante codifica información robada dentro de subdominios de consultas DNS aparentemente legítimas, aprovechando que este tráfico rara vez se inspecciona con el mismo rigor que el tráfico HTTP. Otro patrón habitual es el uso de dominios generados algorítmicamente (DGA), mediante los cuales el malware genera de forma automática cientos de nombres de dominio candidatos para su servidor de mando y control, dificultando el bloqueo mediante listas negras estáticas.

La siguiente tabla recoge los tipos de registro DNS más relevantes desde una óptica de seguridad, junto con su función y un ejemplo de abuso asociado.

Tipo de registro	Función	Relevancia o abuso en seguridad
A / AAAA	Asocia un nombre de dominio a una dirección IPv4 o IPv6	Resolución hacia infraestructura de mando y control
MX	Identifica los servidores de correo de un dominio	Suplantación en campañas de phishing dirigido
TXT	Almacena texto arbitrario asociado al dominio	Canal de verificación SPF/DKIM, también de abuso en túneles DNS
NS	Indica los servidores autoritativos de un dominio	Objetivo de ataques de secuestro de dominio
CNAME	Crea un alias hacia otro nombre de dominio	Encadenamiento usado para ocultar infraestructura maliciosa

Tabla 2.3. Tipos de registro DNS relevantes para la ciberdefensa.

Otro aspecto que suelo incluir en cualquier revisión de madurez de una organización es el estado de DNSSEC, la extensión de seguridad del protocolo DNS que añade firmas criptográficas a las respuestas, permitiendo a un resolutor verificar que una respuesta no se ha manipulado en tránsito. Su adopción real sigue siendo desigual incluso entre organizaciones grandes, en parte por la complejidad operativa de gestionar correctamente la rotación de claves, pero su ausencia deja abierta la puerta a ataques de envenenamiento de caché DNS que redirigen a los usuarios hacia infraestructura maliciosa sin que exista ningún indicio visible de manipulación.

Recomendación: registrar y analizar de forma centralizada todas las consultas DNS salientes de la organización, correlacionándolas con fuentes de inteligencia de amenazas actualizadas, es una de las capacidades de detección con mejor relación entre coste de implantación y valor defensivo que puede desplegar cualquier SOC.

2.2.6 DHCP (configuración dinámica de host)

El protocolo DHCP automatiza la asignación de direcciones IP y de otros parámetros de configuración de red —máscara de subred, puerta de enlace, servidores DNS— a los dispositivos que se conectan a una red,

evitando la necesidad de configurar manualmente cada equipo. El proceso se resume habitualmente en cuatro mensajes: descubrimiento, oferta, petición y confirmación, conocidos por su acrónimo en inglés DORA.

Desde la perspectiva de seguridad, el servidor DHCP es un componente de infraestructura crítico que rara vez recibe la atención que merece. Un servidor DHCP no autorizado —conectado, de forma accidental o deliberada, por un empleado o por un atacante con acceso físico a la red— puede responder a las peticiones de otros equipos antes que el servidor legítimo, entregando una puerta de enlace o unos servidores DNS bajo control del atacante y redirigiendo así, de forma transparente para la víctima, todo su tráfico de salida.

Este ataque, conocido como DHCP spoofing o rogue DHCP, es en especial peligroso precisamente porque no requiere vulnerabilidad técnica alguna: basta con que el servidor no autorizado responda más rápido que el legítimo. Los switches gestionados permiten mitigar este riesgo mediante una función llamada DHCP snooping, que restringe qué puertos de la red pueden actuar legítimamente como origen de respuestas DHCP, descartando el resto.

Error frecuente: desplegar DHCP snooping únicamente en los switches de acceso más recientes de una red y dejar sin esa protección los segmentos antiguos o los puntos de acceso inalámbrico, generando una falsa sensación de cobertura completa que un atacante mínimamente atento detecta con rapidez.

2.2.7 VPN (red privada virtual)

Una VPN establece un túnel cifrado sobre una red no confiable —típicamente internet— que permite a un equipo o a una sede remota comunicarse con la red corporativa como si estuviera físicamente conectado a ella. Esta capacidad, imprescindible desde la generalización del trabajo remoto, ha convertido a las VPN en uno de los puntos de entrada más vigilados, y también más atacados, del perímetro corporativo.

Existen dos modelos principales de VPN que conviene distinguir con claridad. Las VPN de acceso remoto conectan un dispositivo individual —el portátil de un empleado— con la red corporativa, habitualmente mediante protocolos como IPsec o SSL/TLS. Las VPN de sitio a sitio, en cambio, conectan de forma permanente dos redes completas, típicamente la sede

central de una organización con una delegación o con un proveedor externo, sin intervención del usuario final en cada conexión.

En mi experiencia dirigiendo la respuesta a incidentes, una parte significativa de los compromisos de gran alcance que he investigado en los últimos años comenzó con el uso de credenciales robadas contra un servicio VPN sin autenticación multifactor habilitada. La VPN, precisamente por otorgar a quien se conecta un nivel de confianza equiparable al de estar físicamente dentro de la red corporativa, es un objetivo de altísimo valor para cualquier atacante que haya conseguido credenciales válidas mediante phishing o mediante la reutilización de contraseñas filtradas en otros servicios.

Una decisión de configuración que reviso siempre en cualquier despliegue VPN corporativo es si se habilita el split tunneling, esto es, si se permite que parte del tráfico del dispositivo remoto —por ejemplo, la navegación web general— salga directamente a internet sin atravesar el túnel corporativo, mientras que solo el tráfico dirigido a recursos internos viaja cifrado hacia la organización. Esta configuración mejora el rendimiento y reduce la carga sobre la infraestructura VPN, pero también reduce la visibilidad del SOC sobre una parte del tráfico del dispositivo remoto, y puede exponerlo a amenazas de internet sin el filtrado perimetral que sí recibiría el tráfico corporativo.

VPN de acceso remoto frente a VPN de sitio a sitio

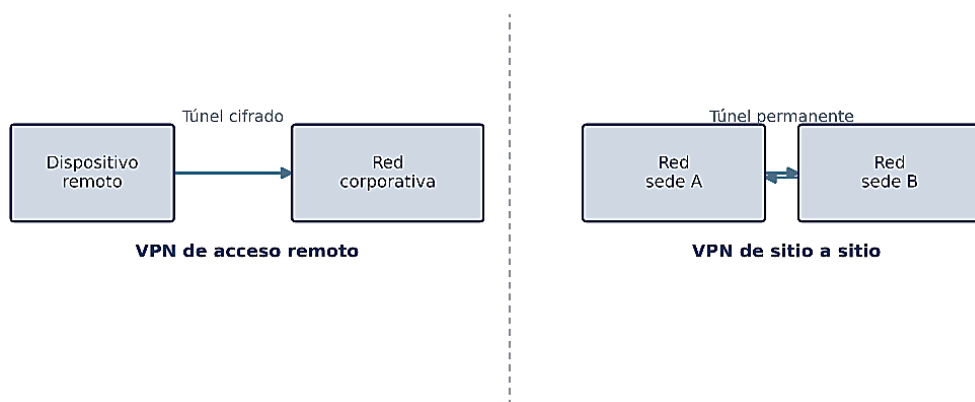


Figura 2.3. Comparativa esquemática entre una VPN de acceso remoto, que conecta un único dispositivo con la red corporativa a través de un túnel cifrado, y una VPN de sitio a sitio, que interconecta de forma permanente dos redes completas.

Consejo: exigir autenticación multifactor en cualquier servicio VPN corporativo no es una recomendación opcional en el panorama de amenazas actual; es, sencillamente, uno de los controles de menor coste y mayor impacto que puede implantar cualquier responsable de seguridad, y suelo situarlo entre las primeras prioridades de cualquier plan de mejora que presento a un consejo de administración.

2.3 DIRECTORIO ACTIVO

Si hay un componente de infraestructura que concentra, de forma casi obsesiva, la atención de cualquier atacante que consiga un primer punto de apoyo dentro de una red corporativa, ese componente es el Directorio Activo. Prácticamente cualquier organización de tamaño medio o grande que utilice Windows como sistema operativo predominante gestiona la identidad de sus usuarios, equipos y recursos a través de este servicio, lo que lo convierte en el objetivo natural de cualquier progresión hacia el compromiso total de la infraestructura.

No exagero cuando afirmo que, en la inmensa mayoría de los incidentes graves que he dirigido a lo largo de mi carrera, el momento decisivo de la intrusión —aquel en el que un atacante deja de ser una molestia localizada para convertirse en una amenaza existencial para la organización— coincide con el instante en que consigue privilegios de administrador de dominio dentro del Directorio Activo. A partir de ese punto, prácticamente cualquier sistema, cualquier fichero y cualquier cuenta de la organización queda, en la práctica, bajo su control.

Los apartados siguientes desarrollan los cuatro conceptos de Directorio Activo que marcan la diferencia entre entender superficialmente esta tecnología y ser capaz de razonar con solidez sobre un incidente que la involucra: el concepto de dominio, el protocolo LDAP que sustenta sus consultas, los objetos de directiva de grupo que distribuyen configuración y, por último, el bastionado que cierra este capítulo.

2.3.1 Dominio

Un dominio de Directorio Activo es un límite administrativo y de seguridad que agrupa un conjunto de usuarios, equipos y recursos bajo

una base de datos común y unas políticas de seguridad compartidas. Todos los objetos de un dominio confían en un mismo conjunto de controladores de dominio, servidores especializados que almacenan la base de datos del directorio y que autentican cada intento de inicio de sesión dentro de ese dominio.

Los dominios pueden agruparse, a su vez, en árboles y bosques, estructuras jerárquicas que permiten a organizaciones grandes o con múltiples filiales mantener dominios administrativamente separados mientras comparten, en distinto grado, relaciones de confianza entre ellos. Comprender estas relaciones de confianza es esencial en cualquier investigación que involucre movimiento lateral entre dominios distintos de un mismo bosque, un escenario cada vez más habitual tras fusiones y adquisiciones empresariales.

La autenticación dentro de un dominio de Directorio Activo se apoya, desde hace más de dos décadas, en el protocolo Kerberos, que sustituye el envío de contraseñas por un sistema de tickets cifrados con una validez temporal limitada. Cuando un usuario inicia sesión, obtiene un ticket inicial (TGT) que después utiliza para solicitar, sin necesidad de volver a introducir su contraseña, tickets de servicio específicos para cada recurso al que necesita acceder. Este diseño reduce de forma significativa la exposición de credenciales en tránsito, pero no es inmune a abusos: técnicas como el Kerberoasting o el Golden Ticket explotan precisamente debilidades en la gestión de estos tickets para obtener acceso persistente sin necesidad de conocer contraseñas en texto claro.

Las relaciones de confianza entre dominios, mencionadas anteriormente, pueden configurarse además con distinto alcance y direccionalidad: una confianza unidireccional permite que los usuarios de un dominio accedan a recursos de otro, pero no a la inversa, mientras que una confianza bidireccional habilita el acceso en ambos sentidos. Auditar qué confianzas existen realmente entre los dominios de una organización, y en qué sentido operan, es un ejercicio que recomiendo repetir con cierta frecuencia, porque las confianzas heredadas de fusiones o reestructuraciones antiguas tienden a sobrevivir mucho más tiempo del que su justificación original habría aconsejado.

Cada controlador de dominio replica de forma continua su base de datos con el resto de controladores del mismo dominio, garantizando que cualquier cambio —la creación de un usuario, la modificación de un grupo, el cambio de una contraseña— se propague en un plazo razonable a toda la

infraestructura. Esta replicación, aunque habitualmente transparente para el usuario, es también una superficie de ataque: un atacante con control sobre un único controlador de dominio comprometido puede inyectar cambios maliciosos que se propagarán automáticamente al resto de la infraestructura de directorio.

Todo controlador de dominio aloja además un recurso compartido especial, SYSVOL, que distribuye scripts de inicio de sesión y la configuración de las directivas de grupo a todos los equipos del dominio. Precisamente por su papel de distribución masiva de contenido, SYSVOL ha sido en el pasado objetivo de ataques que sustituían scripts legítimos por versiones maliciosas, con la garantía de que se ejecutarían de forma automática en un número muy elevado de equipos.

2.3.2 LDAP

El protocolo LDAP (Lightweight Directory Access Protocol) es el mecanismo estándar mediante el cual las aplicaciones y los propios equipos de la organización consultan la información almacenada en el Directorio Activo: qué usuarios existen, a qué grupos pertenecen, qué atributos tiene cada objeto. Prácticamente cualquier operación de autenticación o de consulta de pertenencia a grupos que ocurre en un entorno Windows corporativo se apoya, en algún punto de su ejecución, en una consulta LDAP.

Esta ubicuidad convierte a LDAP en una de las herramientas de reconocimiento favoritas de cualquier atacante que haya obtenido un primer punto de apoyo dentro de la red, incluso con las credenciales de una cuenta de usuario estándar sin privilegios especiales. Una simple consulta LDAP autenticada con una cuenta de dominio de bajo privilegio puede revelar la estructura completa de grupos administrativos, la relación de confianza entre dominios, los equipos con sesiones de administrador activas y, con relativa frecuencia, contraseñas o claves almacenadas en atributos de objetos configurados de forma incorrecta.

Herramientas ampliamente utilizadas tanto por equipos ofensivos como defensivos, como BloodHound, se apoyan precisamente en la recolección masiva de datos LDAP para construir un grafo de relaciones de privilegio dentro del dominio, identificando automáticamente rutas de escalado de privilegios que resultarían muy difíciles de detectar mediante inspección manual. Recomendando encarecidamente a cualquier equipo de Blue Team que se familiarice con estas mismas herramientas

desde la perspectiva defensiva, porque es exactamente lo que un atacante competente utilizará contra la organización.

Un descuido de bastionado que reviso de forma sistemática en cualquier auditoría es si el tráfico LDAP viaja cifrado. Por defecto, muchas consultas LDAP se transmiten en texto claro o con una firma opcional, lo que permite a un atacante situado en la misma red capturar credenciales o manipular respuestas mediante ataques de intermediario. Habilitar LDAPS, la variante cifrada mediante TLS, o exigir la firma obligatoria de todas las consultas LDAP, cierra una vía de ataque que en mi experiencia sigue abierta en una proporción sorprendentemente alta de organizaciones que, en otros aspectos, cuidan razonablemente bien su seguridad.

Error frecuente: asumir que una consulta LDAP anónima o de bajo privilegio no supone un riesgo relevante porque “no modifica nada” ignora que el reconocimiento de la estructura del directorio es, con enorme frecuencia, el paso previo indispensable para cualquier escalado de privilegios posterior.

2.3.3 GPO (objetos de directiva de grupo)

Los objetos de directiva de grupo, conocidos habitualmente por sus siglas en inglés GPO, son el mecanismo mediante el cual un administrador de Directorio Activo distribuye configuración de forma centralizada a conjuntos de usuarios o de equipos: políticas de contraseñas, restricciones de software, configuración de firewall local, scripts de inicio de sesión y prácticamente cualquier otro parámetro de configuración de un sistema Windows unido al dominio.

Esta capacidad de distribución centralizada es, al mismo tiempo, una de las herramientas de bastionado más potentes de cualquier organización y una de las superficies de ataque más peligrosas si cae en manos equivocadas. Un atacante que consiga privilegios suficientes para modificar una GPO aplicada a un gran número de equipos puede, en una sola operación, desplegar un script malicioso, deshabilitar el antivirus corporativo o crear una cuenta administrativa local en toda la organización de forma simultánea.

Las GPO se organizan y se aplican siguiendo una jerarquía de herencia —sitio, dominio, unidad organizativa— y esta herencia, mal diseñada, es una fuente habitual de configuraciones inconsistentes o directamente contradictorias entre distintas partes de la organización. Auditar periódicamente qué GPO existen, a quién se aplican realmente y quién tiene permisos para modificarlas es una de las tareas de higiene de seguridad más descuidadas en organizaciones que, por lo demás, cuentan con un nivel de madurez razonable en otros ámbitos.

Más allá de la herencia jerárquica, las GPO admiten mecanismos de filtrado adicionales —filtrado de seguridad basado en pertenencia a grupos y filtrado WMI basado en características del propio equipo— que permiten aplicar configuraciones muy específicas a subconjuntos reducidos de la organización. Esta flexibilidad es potente, pero también incrementa la complejidad de auditar con precisión qué configuración final recibe en realidad cada equipo, especialmente cuando distintas GPO con filtros solapados interactúan entre sí de forma poco documentada.

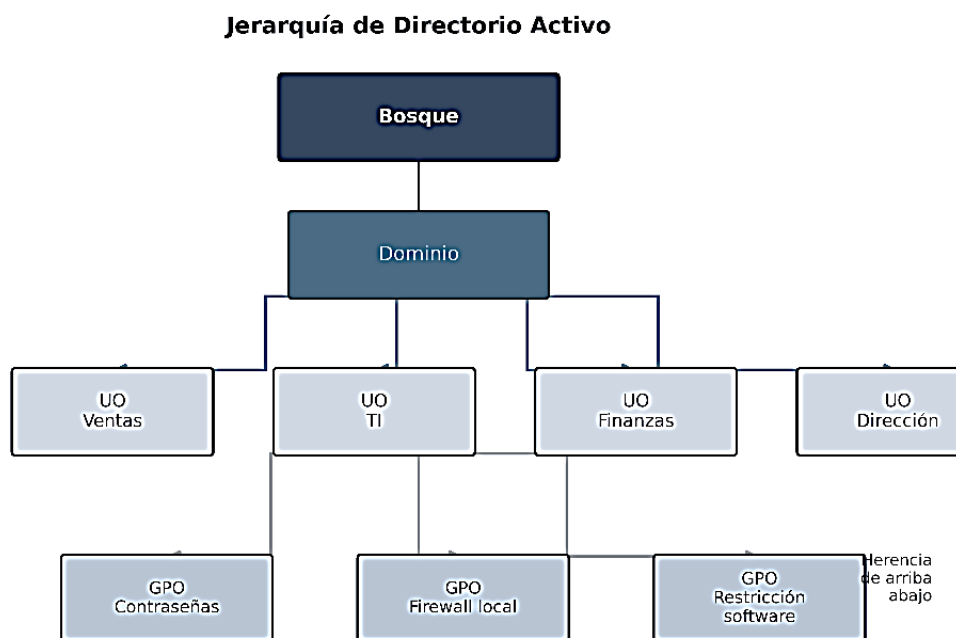


Figura 2.4. Jerarquía de dominio, unidades organizativas y objetos de directiva de grupo dentro de un bosque de Directorio Activo, con el sentido de herencia de configuración de arriba hacia abajo.

2.3.4 Bastionado (hardening)

El bastionado, o hardening, es el conjunto de medidas orientadas a reducir la superficie de ataque de un sistema, eliminando o restringiendo todo aquello que no resulte estrictamente necesario para su función: servicios innecesarios, puertos abiertos sin uso, cuentas con privilegios excesivos, protocolos obsoletos o inseguros habilitados por defecto. Aplicado al Directorio Activo, el bastionado adquiere una importancia crítica precisamente por la posición central que este servicio ocupa en toda la infraestructura de identidad de la organización.

Entre las medidas de bastionado de Directorio Activo que considero irrenunciables en cualquier organización que dirija, destaco tres por su impacto práctico demostrado en las investigaciones que he liderado. La primera es la limitación estricta del número de cuentas con privilegios de administrador de dominio, aplicando el principio de privilegio mínimo también a los propios administradores, mediante cuentas administrativas separadas de las cuentas de uso diario. La segunda es la implantación de un modelo de niveles (tiering) que impide que las credenciales de administración de estaciones de trabajo puedan reutilizarse para administrar servidores críticos o controladores de dominio. La tercera es la desactivación de protocolos de autenticación heredados y débiles, como NTLM en sus versiones más antiguas, en favor de Kerberos con configuraciones actualizadas.

La siguiente tabla resume estas y otras medidas de bastionado habituales, junto con el riesgo concreto que mitigan.

Medida de bastionado	Riesgo que mitiga
Modelo de niveles (tiering) administrativo	Reutilización de credenciales de alto privilegio en equipos de menor confianza
Cuentas administrativas separadas de las de uso diario	Robo de credenciales de administrador mediante phishing en el uso cotidiano
Desactivación de NTLM heredado	Ataques de retransmisión de credenciales (relay) y de fuerza bruta offline
Auditoría periódica de membresía de grupos privilegiados	Escalado de privilegios silencioso no detectado
Contraseñas robustas y rotación de cuentas de servicio	Ataques de Kerberoasting contra cuentas de servicio

Tabla 2.4. Medidas de bastionado del Directorio Activo y riesgo que mitigan.

A estas tres medidas añado habitualmente una cuarta que suele pasar desapercibida en organizaciones que no han sufrido todavía un incidente grave: la gestión de contraseñas de administrador local mediante una solución como LAPS, que asigna y rota de forma automática una contraseña distinta para la cuenta de administrador local de cada equipo. Sin esta medida, es extremadamente habitual encontrar la misma contraseña de administrador local reutilizada en cientos o miles de equipos de una organización, de modo que comprometer un único puesto de trabajo mediante esa credencial equivale, en la práctica, a comprometer administrativamente todos los demás equipos que comparten la misma contraseña.

Ninguna de estas medidas, aplicada de forma aislada, ofrece una protección completa. Es la combinación disciplinada de todas ellas, revisada con una periodicidad razonable y ajustada a medida que evoluciona la organización, lo que realmente reduce de forma sostenida la probabilidad de que un atacante progrese desde un primer punto de apoyo hasta el control total del dominio.

Qué debes recordar: el sistema operativo y la red no son un requisito previo anecdótico para trabajar en un SOC, sino el terreno mismo sobre el que ocurre cada incidente; y dentro de ese terreno, el Directorio Activo concentra un nivel de riesgo tan desproporcionado respecto al resto de la infraestructura que su bastionado merece, por sí solo, una atención prioritaria y continuada por parte de cualquier responsable de ciberdefensa.