Capacidades técnicas, tácticas y de respuesta para equipos red team y blue team

Germán Alonso Ramírez Cleves

Asesor

Eduvin Trigos Sánchez

Universidad Nacional Abierta y a Distancia UNAD

Escuela de Ciencias Básicas, Tecnología e Ingeniería – ECBTI

Especialización en seguridad informática

2026

Resumen

El presente informe técnico consolida los resultados del Seminario Especializado en Equipos Estratégicos en Ciberseguridad: Red Team y Blue Team, aplicados al escenario organizacional de SecureNova Labs. El ejercicio integró las fases de reconocimiento, análisis de vulnerabilidades, explotación controlada, detección, respuesta y contención de amenazas sobre una infraestructura tecnológica de mediana escala. Las operaciones del Red Team identificaron una cadena de compromiso que avanzó desde una inyección SQL en una aplicación web expuesta hasta el control total del directorio activo, evidenciando brechas críticas en la gestión de parches, la configuración de servicios y la autenticación. El Blue Team implementó estrategias de detección basadas en análisis de tráfico, correlación de eventos y análisis forense digital, logrando contener las amenazas y formular planes de remediación estructurados. Se analizó el marco legal colombiano e internacional aplicable, incluyendo la Ley 1273 de 2009, la Ley 1581 de 2012 y los estándares ISO/IEC 27001 y NIST. La evaluación de madurez según el NIST Cybersecurity Framework ubicó a la organización en niveles iniciales, con brechas significativas en detección temprana. Los resultados constituyen una hoja de ruta técnica priorizada para elevar la postura de seguridad de la organización.

***Palabras clave*:** ciberseguridad, contención, detección, riesgo, vulnerabilidad.

Abstract

This technical report consolidates the results of the Specialized Seminar on Strategic Teams in Cybersecurity: Red Team and Blue Team, applied to the organizational scenario of SecureNova Labs. The exercise integrated the phases of reconnaissance, vulnerability analysis, controlled exploitation, detection, response, and threat containment over a medium-scale technology infrastructure. Red Team operations identified a compromise chain that progressed from a SQL injection in an exposed web application to full control of the active directory, revealing critical gaps in patch management, service configuration, and authentication. The Blue Team implemented detection strategies based on traffic analysis, event correlation, and digital forensic analysis, successfully containing the threats and formulating structured remediation plans. The applicable Colombian and international legal framework was analyzed, including Law 1273 of 2009, Law 1581 of 2012, and the ISO/IEC 27001 and NIST standards. The maturity assessment under the NIST Cybersecurity Framework placed the organization at initial levels, with significant gaps in early detection. The results constitute a prioritized technical roadmap to strengthen the organization's security posture.

Keywords**:** containment, cybersecurity, detection, risk, vulnerability.

Tabla de Contenido

Lista de Figuras

Lista de Tablas

Lista de Apéndices

Introducción

El panorama de la ciberseguridad ha experimentado una transformación profunda durante la última década, impulsada por la expansión de la superficie de ataque digital, la sofisticación creciente de los actores de amenaza y la digitalización acelerada de los procesos organizacionales. En este contexto, la capacidad de anticipar, detectar y responder eficazmente a los incidentes de seguridad se ha convertido en un factor estratégico para la continuidad operativa y la protección de los activos de información.

La metodología de ejercicios basada en equipos especializados, concretamente el Red Team y el Blue Team, representa uno de los enfoques más rigurosos para evaluar de manera integral la postura de seguridad de una organización. Mientras el Red Team adopta la perspectiva del adversario para identificar y explotar vulnerabilidades reales, el Blue Team opera desde la defensa, implementando controles de monitoreo, detección y contención. La interacción entre ambos equipos genera un ciclo de retroalimentación continua que eleva el nivel de madurez en ciberseguridad de la organización evaluada.

El presente informe se desarrolla en el marco del Seminario Especializado en Equipos Estratégicos en Ciberseguridad: Red Team & Blue Team, ofertado por la Universidad Nacional Abierta y a Distancia como parte de la Especialización en Seguridad Informática. El documento consolida los resultados obtenidos durante las etapas del seminario, articuladas en torno al escenario organizacional de SecureNova Labs, empresa que presentó una serie de vulnerabilidades técnicas y organizacionales sobre las cuales se ejecutaron los ejercicios de seguridad ofensiva y defensiva.

El escenario planteado en el Anexo 6 estableció la necesidad de actuar como experto en ciberseguridad durante un período de prueba en SecureNova Labs, con el objetivo de documentar

el proceso de los escenarios propuestos en cada una de las acciones del Blue Team, el Red Team y los aspectos legales. El informe resultante está destinado a ser analizado por los analistas seniors de seguridad de la organización como parte de un proceso de selección de expertos.

Desde el punto de vista metodológico, el trabajo se estructuró siguiendo el ciclo de vida de una evaluación de seguridad por equipos, que comprende las fases de reconocimiento, análisis de vulnerabilidades, explotación controlada, respuesta y contención, análisis forense y formulación de recomendaciones. Cada fase se documentó incorporando las herramientas utilizadas, los hallazgos obtenidos, las técnicas aplicadas y las lecciones aprendidas tanto en términos técnicos como estratégicos.

El documento se organiza, tras las secciones preliminares, en el marco conceptual y metodológico, la contextualización del escenario, las operaciones del Red Team, las actividades del Blue Team, el análisis del marco legal aplicable, la evaluación de riesgos, las conclusiones y las recomendaciones, acompañadas de las referencias bajo normas APA séptima edición y los apéndices correspondientes..

Justificación

La proliferación de amenazas cibernéticas dirigidas a organizaciones de todos los sectores ha hecho evidente que los enfoques tradicionales de seguridad basados exclusivamente en controles preventivos resultan insuficientes frente a adversarios sofisticados y persistentes. El tiempo promedio de permanencia de un atacante en un sistema comprometido antes de ser detectado sigue siendo de varios días, lo que subraya la necesidad de capacidades de detección y respuesta avanzadas que complementen los mecanismos de prevención existentes.

La realización de ejercicios estructurados de Red Team y Blue Team permite a las organizaciones obtener una evaluación objetiva de su postura de seguridad, identificar brechas que los controles automatizados no logran detectar y desarrollar la capacidad operacional necesaria para responder eficazmente ante incidentes reales. La integración de evaluaciones periódicas de seguridad ofensiva forma parte de las prácticas recomendadas para alcanzar niveles de madurez superiores en la gestión de la ciberseguridad organizacional (National Institute of Standards and Technology [NIST], 2018).

Desde la perspectiva académica, el desarrollo de competencias en seguridad ofensiva y defensiva mediante escenarios simulados constituye una modalidad de formación altamente efectiva, dado que combina el conocimiento teórico con la aplicación práctica en entornos controlados que replican condiciones reales de ataque. Este tipo de experiencia resulta especialmente relevante en el contexto de la Especialización en Seguridad Informática, cuyo propósito es formar profesionales capaces de analizar, diseñar e implementar soluciones de ciberseguridad en organizaciones del sector público y privado.

El ejercicio aplicado al escenario de SecureNova Labs se justifica adicionalmente por la necesidad de que los profesionales desarrollen la capacidad de documentar y comunicar los resultados de sus evaluaciones de manera técnica y comprensible, tanto para audiencias especializadas como para tomadores de decisiones. La elaboración del presente informe contribuye a ese objetivo, integrando rigor técnico, claridad comunicativa y pertinencia estratégica.

Objetivos

Objetivo General

Consolidar mediante un informe técnico los resultados del Seminario Especializado en Equipos Estratégicos en Ciberseguridad, a partir del análisis del escenario de SecureNova Labs, describiendo las estrategias de seguridad ofensiva y defensiva aplicadas y formulando conclusiones y recomendaciones orientadas al fortalecimiento de la postura de ciberseguridad de la organización.

Objetivos Específicos

Identificar y documentar las vulnerabilidades técnicas presentes en la infraestructura de SecureNova Labs mediante la ejecución de ejercicios de reconocimiento y análisis propios del Red Team.

Describir las tácticas, técnicas y procedimientos aplicados durante las fases de explotación controlada, escalación de privilegios y persistencia, en correspondencia con el marco MITRE ATT&CK.

Analizar las estrategias de detección, contención y respuesta implementadas por el Blue Team frente a las amenazas identificadas, estableciendo la trazabilidad entre vulnerabilidades, impacto y controles.

Examinar el marco legal colombiano e internacional aplicable al contexto de seguridad informática de la organización.

Formular un análisis de riesgos que priorice las vulnerabilidades según su impacto sobre la confidencialidad, integridad y disponibilidad de los activos de información.

Proponer recomendaciones técnicas y estratégicas priorizadas, orientadas al fortalecimiento de la postura de seguridad de SecureNova Labs.

Marco Conceptual

La Disciplina de la Seguridad Ofensiva

La seguridad ofensiva constituye el conjunto de prácticas orientadas a evaluar la postura de seguridad de una organización mediante la simulación de las acciones que ejecutaría un adversario real. A diferencia de los enfoques puramente defensivos, que se concentran en la implementación de controles preventivos, la seguridad ofensiva adopta la perspectiva del atacante para identificar las debilidades que podrían ser explotadas. Esta aproximación reconoce que la única manera de validar de forma fehaciente la efectividad de los controles de seguridad consiste en someterlos a las mismas técnicas que emplearían los adversarios.

El Red Team representa la expresión más completa de la seguridad ofensiva en el ámbito organizacional. A diferencia de las pruebas de penetración tradicionales, que suelen centrarse en la identificación del mayor número posible de vulnerabilidades en un alcance acotado, el ejercicio de Red Team persigue objetivos específicos que replican las motivaciones de un adversario real, como el acceso a determinados datos sensibles o el compromiso de sistemas críticos. Esta orientación a objetivos confiere a los ejercicios de Red Team un carácter más realista y permite evaluar no solo la existencia de vulnerabilidades, sino también la capacidad de la organización para detectar y responder a un ataque en progreso.

Las tácticas, técnicas y procedimientos empleados por el Red Team se fundamentan en el conocimiento del comportamiento de los adversarios reales, sistematizado en marcos de referencia como MITRE ATT&CK. Este marco organiza las acciones adversarias en una matriz de tácticas, que representan los objetivos del atacante en cada fase del ataque, y técnicas, que describen los métodos específicos para alcanzar dichos objetivos. La adopción de este marco

como referencia común permite estructurar las operaciones ofensivas de manera sistemática y facilita la comunicación de los hallazgos a los equipos defensivos.

La Disciplina de la Seguridad Defensiva

La seguridad defensiva, materializada en el Blue Team, comprende el conjunto de prácticas orientadas a proteger los activos de información de una organización mediante la implementación de controles preventivos, de detección y de respuesta. El Blue Team opera de manera continua, monitoreando la infraestructura en busca de indicios de actividad maliciosa y respondiendo a los incidentes que se detectan. Su función no se limita a la reacción ante los ataques, sino que abarca también la mejora continua de los controles a partir de las lecciones aprendidas de cada incidente.

La capacidad de detección constituye el núcleo de la función defensiva. Un control preventivo, por robusto que sea, puede ser eludido por un adversario suficientemente determinado y capaz, por lo que la capacidad de detectar la actividad maliciosa que ha logrado superar los controles preventivos resulta esencial. La detección efectiva requiere la recopilación y correlación de información de seguridad proveniente de múltiples fuentes, así como la capacidad analítica para distinguir la actividad legítima de la maliciosa en grandes volúmenes de datos. Las plataformas de gestión de información y eventos de seguridad constituyen la herramienta central para esta función.

La respuesta a incidentes complementa la detección, transformando la identificación de una amenaza en acciones concretas de contención, erradicación y recuperación. Un proceso de respuesta maduro requiere procedimientos documentados, roles y responsabilidades claramente definidos y la capacidad de coordinar la actuación de múltiples equipos bajo la presión de un incidente activo. La preparación previa, mediante la elaboración de planes de respuesta y la

realización de ejercicios de simulación, determina en gran medida la efectividad de la respuesta ante un incidente real.

Tipología de Amenazas y Ciclo de Vida del Ataque

La comprensión de las amenazas que enfrentan las organizaciones constituye un requisito previo para el diseño de estrategias de seguridad efectivas. Las amenazas pueden clasificarse según diversos criterios, entre los que destacan el origen del actor, su motivación y su nivel de sofisticación. Los actores de amenaza abarcan desde atacantes oportunistas con conocimientos limitados hasta grupos organizados con recursos significativos y objetivos específicos, pasando por amenazas internas originadas en personal de la propia organización. Cada tipo de actor presenta capacidades, motivaciones y patrones de comportamiento diferenciados que condicionan las estrategias defensivas más apropiadas.

Las amenazas persistentes avanzadas representan la categoría de mayor sofisticación, caracterizada por la capacidad del adversario para mantener un acceso prolongado y encubierto al entorno comprometido, evadiendo los controles de detección durante períodos extensos.

Aunque el escenario de SecureNova Labs no involucró a un actor de este nivel, la simulación de mecanismos de persistencia durante el ejercicio permitió evaluar la capacidad de la organización para detectar y responder a las técnicas características de este tipo de amenazas, que constituyen el mayor desafío para las capacidades defensivas.

El ciclo de vida del ataque, conceptualizado en modelos como la cadena de eliminación cibernética, describe las fases sucesivas que un adversario recorre desde la fase inicial de reconocimiento hasta la consecución de sus objetivos. Estas fases comprenden el reconocimiento, la preparación de las herramientas de ataque, la entrega del vector inicial, la explotación de la vulnerabilidad, la instalación de mecanismos de control, el establecimiento de

canales de comando y control y la ejecución de las acciones sobre los objetivos. La comprensión de este ciclo resulta fundamental para el diseño de controles defensivos, dado que la interrupción del ataque en cualquiera de sus fases puede prevenir la consecución de los objetivos del adversario.

El ejercicio realizado sobre SecureNova Labs reprodujo de manera fiel las fases de este ciclo, lo que permitió evaluar la capacidad de la organización para detectar y responder en cada una de ellas. Los resultados evidenciaron que las capacidades de detección de la organización se concentraban en las fases más ruidosas del ciclo, como el reconocimiento activo, mientras que las fases más sigilosas, como la explotación de la aplicación web y el establecimiento de la persistencia, pasaron desapercibidas durante períodos prolongados. Esta distribución de las capacidades de detección orienta la priorización de las inversiones en controles hacia las fases del ciclo que actualmente presentan menor cobertura.

La Integración de Capacidades: el Modelo Purple Team

La evolución de las prácticas de seguridad ha conducido al reconocimiento de que la separación estricta entre los equipos ofensivo y defensivo puede limitar la efectividad de ambos. El modelo Purple Team surge como respuesta a esta limitación, proponiendo una colaboración estructurada entre el Red Team y el Blue Team en la que los hallazgos de las operaciones ofensivas se traducen de manera inmediata en mejoras de los controles defensivos. En este modelo, el Red Team no solo identifica las vulnerabilidades, sino que comparte de manera detallada las técnicas empleadas con el Blue Team, lo que permite a este último desarrollar y validar las capacidades de detección correspondientes.

La adopción del enfoque Purple Team resulta particularmente valiosa para organizaciones con una madurez en seguridad en desarrollo, como SecureNova Labs, dado que

maximiza la transferencia de conocimiento y acelera el desarrollo de las capacidades defensivas. La colaboración entre los equipos permite identificar de manera precisa las brechas de detección, priorizar las inversiones en controles y validar la efectividad de las mejoras implementadas mediante la repetición de los ejercicios ofensivos. Este ciclo de mejora continua constituye uno de los mecanismos más efectivos para elevar de manera sostenida el nivel de madurez en ciberseguridad.

Antecedentes y Estado del Arte

La práctica de evaluar la seguridad de los sistemas mediante la simulación de ataques tiene sus raíces en los ejercicios militares de equipos contrarios, en los cuales un grupo asumía el rol del adversario para poner a prueba las defensas del bando propio. La transposición de este concepto al ámbito de la seguridad informática dio origen a las metodologías de Red Team y Blue Team que hoy constituyen un estándar en la evaluación de la postura de seguridad organizacional. La evolución de estas prácticas ha estado marcada por la creciente sofisticación de las amenazas y por la necesidad de adoptar enfoques que reflejen las condiciones reales a las que se enfrentan las organizaciones.

La literatura especializada coincide en que las pruebas de penetración tradicionales, centradas en la identificación exhaustiva de vulnerabilidades, resultan insuficientes para evaluar la capacidad de respuesta de una organización ante un ataque dirigido. Los ejercicios de Red Team, en cambio, al perseguir objetivos específicos y simular las tácticas de adversarios reales, permiten una evaluación más realista que incorpora la dimensión temporal de la detección y la respuesta. Esta distinción resulta especialmente relevante en sectores regulados, donde la demostración de la capacidad de respuesta constituye un requisito de cumplimiento.

Los marcos de referencia desarrollados por organizaciones especializadas han contribuido a sistematizar estas prácticas. El marco MITRE ATT&CK, en particular, ha transformado la manera en que las organizaciones comprenden y comunican el comportamiento de los adversarios, al proporcionar un lenguaje común basado en tácticas y técnicas observadas en incidentes reales. De manera complementaria, los marcos de gestión de la ciberseguridad, como el NIST Cybersecurity Framework, han establecido modelos de madurez que permiten a las organizaciones evaluar su estado y planificar su evolución de manera estructurada.

El presente ejercicio se inscribe en esta tradición metodológica, aplicando los marcos y las prácticas reconocidas al escenario específico de SecureNova Labs. La contribución del trabajo radica no en la innovación metodológica, sino en la aplicación rigurosa de las prácticas establecidas a un contexto organizacional concreto, generando hallazgos accionables y recomendaciones fundamentadas que permitan a la organización elevar su nivel de madurez en ciberseguridad.

Marco Metodológico

El desarrollo del ejercicio de seguridad sobre SecureNova Labs siguió una metodología estructurada que integró estándares reconocidos en la disciplina, con el fin de garantizar el rigor, la reproducibilidad y la trazabilidad de los hallazgos. La metodología combinó elementos del Penetration Testing Execution Standard, la guía de pruebas del Open Web Application Security Project y el marco de respuesta a incidentes del NIST, articulados en un flujo de trabajo coherente con el ciclo de vida de una evaluación por equipos.

Enfoque y Alcance

El enfoque adoptado fue de carácter mixto, combinando técnicas cualitativas para la interpretación de los hallazgos con técnicas cuantitativas para la valoración de la severidad y el riesgo. El alcance del ejercicio comprendió la infraestructura tecnológica de SecureNova Labs definida en el Anexo 6, incluyendo los servicios expuestos a internet, la zona desmilitarizada y la red interna corporativa. Quedaron excluidas del alcance las pruebas de denegación de servicio y cualquier acción que pudiera comprometer la disponibilidad de los servicios productivos, en concordancia con los límites éticos y legales del ejercicio.

La definición precisa del alcance constituye una de las decisiones metodológicas de mayor relevancia en una evaluación de seguridad, dado que determina los límites de las acciones autorizadas y previene la afectación de sistemas no contemplados. En el caso de SecureNova Labs, el alcance se documentó de manera explícita antes del inicio de las actividades, estableciendo los rangos de direcciones, los sistemas objetivo y las ventanas de tiempo autorizadas para la ejecución de las pruebas activas.

Fases de la Evaluación

La evaluación se estructuró en seis fases secuenciales e interdependientes. La primera fase, de reconocimiento, comprendió la recopilación de información sobre el objetivo mediante técnicas pasivas y activas. La segunda fase, de análisis de vulnerabilidades, consistió en la identificación, clasificación y priorización de las debilidades de seguridad. La tercera fase, de explotación controlada, implicó el aprovechamiento de las vulnerabilidades identificadas para demostrar su impacto real. La cuarta fase, de detección y análisis, correspondió a las actividades del Blue Team orientadas a identificar la actividad maliciosa. La quinta fase, de contención, erradicación y recuperación, comprendió las acciones de respuesta al incidente. La sexta fase, de análisis y documentación, integró los hallazgos en el presente informe.

Cada fase generó productos específicos que alimentaron la fase siguiente, configurando un flujo de trabajo en el que la información se enriqueció progresivamente. Esta estructura permitió mantener la trazabilidad entre los hallazgos de las fases ofensivas y las acciones de las fases defensivas, aspecto fundamental para la coherencia del análisis y para la formulación de recomendaciones fundamentadas.

Criterios de Valoración

La severidad de las vulnerabilidades se valoró mediante el Common Vulnerability Scoring System versión 3.1, que asigna una puntuación numérica entre cero y diez a partir de un conjunto de métricas que consideran el vector de ataque, la complejidad de la explotación, los privilegios requeridos y el impacto sobre la confidencialidad, la integridad y la disponibilidad. Esta valoración permitió clasificar los hallazgos en los niveles crítico, alto, medio y bajo, facilitando la priorización de la remediación.

El riesgo se valoró mediante una metodología semicuantitativa basada en la relación entre la probabilidad de ocurrencia y el impacto potencial, en concordancia con los lineamientos de la norma ISO/IEC 27005. Esta metodología permitió ubicar cada escenario de riesgo en una matriz de cinco por cinco y determinar su nivel de prioridad, proporcionando una base objetiva para la toma de decisiones sobre la asignación de recursos de seguridad.

Contextualización del Escenario

Descripción de la Organización

SecureNova Labs es una organización del sector tecnológico que opera como laboratorio de investigación y desarrollo de soluciones de software para el sector financiero y gubernamental. De acuerdo con el escenario planteado en el Anexo 6, la organización cuenta con una infraestructura tecnológica de mediana escala que incluye servidores físicos y virtualizados, servicios expuestos a internet, redes corporativas y aplicaciones web desarrolladas a medida. La plantilla de personal de tecnología es reducida en relación con la complejidad de la infraestructura, lo que genera brechas en la cobertura operativa de los controles de seguridad.

El contexto de SecureNova Labs refleja una realidad común en empresas de tecnología de mediano tamaño, donde la presión por la entrega de productos suele relegar las prácticas de seguridad a un plano secundario. Esta dinámica favorece la acumulación de deuda técnica en materia de seguridad, expresada en la coexistencia de sistemas desactualizados, configuraciones inseguras por defecto y la ausencia de procesos formales de gestión de vulnerabilidades.

La solicitud de la organización de contar con un informe técnico que documente el proceso de evaluación se enmarca en su necesidad de demostrar ante sus clientes del sector financiero y gubernamental que dispone de controles adecuados para la protección de los datos que procesa y almacena. Esta motivación añade una dimensión regulatoria al ejercicio, dado que los sectores mencionados se encuentran sujetos a marcos de cumplimiento exigentes.

Tabla 1

Activos críticos identificados en SecureNova Labs

Activo	Criticidad	Dimensión afectada
Base de datos de clientes	Crítica	Confidencialidad e integridad
Directorio activo	Crítica	Confidencialidad y disponibilidad
Repositorios de código fuente	Alta	Confidencialidad e integridad
Servidores web de producción	Alta	Disponibilidad
Servidor de correo electrónico	Media	Confidencialidad y disponibilidad

Nota. La clasificación de la criticidad corresponde al resultado del análisis de riesgos realizado para el escenario evaluado, considerando el impacto potencial sobre las dimensiones de confidencialidad, integridad y disponibilidad de cada activo de información.

Arquitectura de la Infraestructura

La infraestructura evaluada comprende un perímetro externo delimitado por un firewall de próxima generación que gestiona el tráfico entre internet y la zona desmilitarizada. En la zona desmilitarizada se alojan los servidores web que exponen las aplicaciones públicas, así como el servidor de correo electrónico. La red interna corporativa se distribuye en subredes de usuarios finales, servidores de aplicaciones, bases de datos y gestión de tecnología. Durante el reconocimiento se constató la ausencia de segmentación efectiva entre la red de usuarios y la red de servidores internos, lo que amplió considerablemente la superficie de ataque disponible una vez logrado el acceso inicial. La Figura 1 representa la topología identificada.

Figura 1

Arquitectura de la infraestructura evaluada

Nota. La conectividad directa entre el servidor web de la zona desmilitarizada y la red interna constituye un riesgo de movimiento lateral.

La organización no contaba con un sistema de gestión de información y eventos de seguridad operativo al momento de la evaluación, lo que limitaba su capacidad de correlación de eventos y detección temprana de amenazas. Esta carencia, sumada a la segmentación insuficiente, configuró un entorno con baja resiliencia frente a un compromiso inicial.

Operaciones del Red Team

Fase de Reconocimiento

La fase de reconocimiento constituye el punto de partida de toda operación de Red Team y, en muchos sentidos, es la etapa con mayor impacto sobre el éxito de las fases subsiguientes. Su objetivo central es la recopilación sistemática de información sobre el objetivo que permita comprender la arquitectura del entorno, identificar posibles vectores de entrada y planificar las acciones de explotación con la mayor precisión posible. El reconocimiento efectivo reduce el ruido generado durante las fases activas y minimiza el riesgo de detección prematura (Alhamed et al., 2023).

Reconocimiento pasivo

El reconocimiento pasivo se desarrolló utilizando exclusivamente fuentes de información públicamente accesibles, sin interacción directa con los sistemas del objetivo. La recopilación de inteligencia de fuentes abiertas incluyó la consulta de registros DNS, la revisión de información de registro de dominios, la enumeración pasiva de subdominios y la revisión de certificados publicados en los registros de transparencia de certificados. Estos recursos permitieron identificar catorce subdominios activos, varios de los cuales no aparecían referenciados en la documentación interna proporcionada como contexto del ejercicio.

La consulta de plataformas de indexación de dispositivos conectados identificó varios hosts con puertos abiertos accesibles desde internet, entre los que destacan instancias de servicios de acceso remoto expuestos en puertos no estándar, un servidor con una versión desactualizada de su servidor web y una interfaz de administración accesible sin autenticación previa.

Adicionalmente, la revisión de metadatos de documentos corporativos publicados reveló nombres de usuario internos, versiones de software y rutas de red, información útil en fases posteriores.

La búsqueda en repositorios de código público arrojó hallazgos de particular relevancia: se encontraron fragmentos de código de proyectos internos, publicados por personal actual y anterior, que contenían credenciales de acceso a bases de datos de entornos de desarrollo codificadas directamente en el código fuente. Aunque los entornos de desarrollo y producción se encuentran separados, la presencia de estas credenciales representa un riesgo de reutilización de contraseñas en los entornos productivos.

Tabla 2

Subdominios y servicios expuestos identificados

Serivicio	Estado	Observación
Aplicación web principal	Expuesto	Versión de framework obsoleta
Servidor de correo	Expuesto	Vulnerabilidades de ejecución remota
Interfaz de administración	Expuesto	Accesible sin autenticación
Servicio de acceso remoto	Expuesto	Puerto no estándar
Servidor FTP	Expuesto	Credenciales por defecto

Nota. Síntesis de los hallazgos del reconocimiento pasivo y activo.

Reconocimiento activo

Concluido el reconocimiento pasivo, se procedió a las actividades de reconocimiento activo, que implican la interacción directa con los sistemas del objetivo. El escaneo de red se realizó con configuraciones destinadas a equilibrar la velocidad con la evasión de los

mecanismos de detección. Se ejecutaron escaneos de identificación de versiones de servicios y detección de sistema operativo sobre los rangos definidos en el alcance. Los resultados revelaron veintitrés hosts activos y un total de ciento ochenta y siete puertos abiertos distribuidos entre protocolos de transporte.

El escaneo de aplicaciones web identificó la exposición de directorios de trabajo, la presencia de versiones obsoletas de frameworks y la ausencia de cabeceras de seguridad estándar. La enumeración de servicios DNS internos, posible debido a una configuración incorrecta que permitía transferencias de zona, proporcionó un mapa completo de la infraestructura interna, incluyendo nombres de host, direcciones y registros de alias de sistemas no accesibles directamente desde el exterior. El detalle de comandos utilizados se presenta en el Apéndice A.

Fundamentos metodológicos del reconocimiento

El reconocimiento se sustenta en estándares metodológicos ampliamente reconocidos en la disciplina de las pruebas de penetración. El Penetration Testing Execution Standard estructura esta fase en categorías que abarcan la inteligencia de fuentes abiertas, el análisis de la presencia digital de la organización y la identificación de la infraestructura tecnológica. La adopción de un enfoque estructurado garantiza la reproducibilidad del proceso y facilita la posterior correlación de los hallazgos con las fases de explotación. En el caso de SecureNova Labs, la aplicación de este enfoque permitió construir un mapa de activos progresivo que se enriqueció en cada iteración del reconocimiento.

La distinción entre reconocimiento pasivo y activo no es meramente técnica, sino también estratégica desde la perspectiva de la evasión de la detección. El reconocimiento pasivo, al no generar interacción directa con los sistemas del objetivo, resulta indetectable para los

controles de monitoreo de la organización, lo que permite al adversario construir un conocimiento detallado del entorno sin alertar a los defensores. El reconocimiento activo, en cambio, genera tráfico que puede ser detectado, por lo que su ejecución debe equilibrar la profundidad de la información obtenida con el riesgo de detección. Esta tensión entre información y sigilo constituye uno de los aspectos centrales de la planificación de una operación ofensiva.

La interpretación de los hallazgos del reconocimiento de SecureNova Labs revela un patrón característico de las organizaciones con baja madurez en la gestión de su superficie de ataque. La presencia de catorce subdominios activos, varios de ellos no documentados, evidencia la ausencia de un inventario centralizado y actualizado de los activos expuestos a internet. Esta carencia, conocida en la disciplina como expansión no controlada de la superficie de ataque, constituye uno de los factores de riesgo más relevantes en los entornos corporativos modernos, dado que cada activo expuesto y no gestionado representa un potencial vector de entrada para un adversario.

Procedimiento aplicado y consideraciones operativas

El procedimiento de reconocimiento se ejecutó de manera escalonada, comenzando por las técnicas de menor riesgo de detección y avanzando progresivamente hacia las de mayor interacción con los sistemas objetivo. Esta secuencia permitió maximizar la información obtenida en las etapas iniciales, reduciendo la necesidad de actividades activas que pudieran alertar a los defensores. Cada hallazgo del reconocimiento pasivo se utilizó para orientar y focalizar las actividades de reconocimiento activo, optimizando el uso de los recursos y reduciendo la huella generada.

Durante el reconocimiento se prestó especial atención a la correlación de la información proveniente de fuentes diversas. La combinación de los datos obtenidos de los registros de dominios, los certificados de transparencia y la indexación de dispositivos permitió construir un mapa de la superficie de ataque más completo que el que habría resultado de cualquiera de estas fuentes de manera aislada. Esta práctica de correlación constituye uno de los aspectos que distinguen un reconocimiento profesional de una simple recopilación de datos, dado que el valor de la información reside en gran medida en las relaciones que pueden establecerse entre los distintos elementos.

La documentación rigurosa de cada hallazgo, incluyendo su fuente, el momento de su obtención y su relevancia para las fases posteriores, constituyó una práctica transversal a toda la fase. Esta documentación no solo facilitó la planificación de las actividades de explotación, sino que también proporcionó la trazabilidad necesaria para que el Blue Team pudiera posteriormente comprender el alcance de la información expuesta y adoptar las medidas correctivas correspondientes, como la reducción de la información disponible públicamente y el fortalecimiento de la gestión de la superficie de ataque.

Análisis de Vulnerabilidades

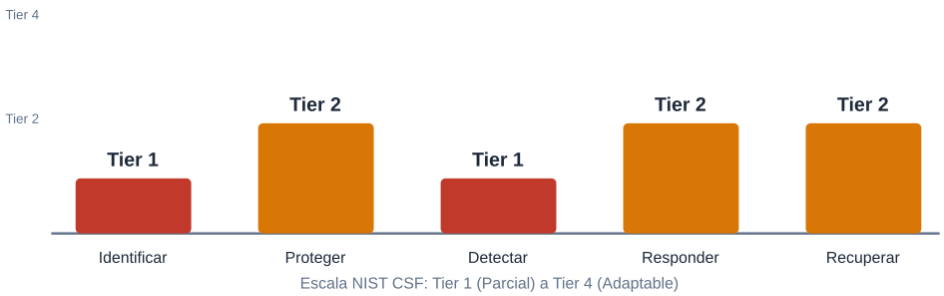
El análisis de vulnerabilidades siguió un proceso estructurado de identificación, clasificación y priorización, utilizando como referente el Common Vulnerability Scoring System versión 3.1 para la cuantificación de la severidad. El proceso se dividió en dos etapas complementarias: el escaneo automatizado y la validación manual de los hallazgos, depurando los falsos positivos y priorizando las vulnerabilidades de mayor impacto en el contexto específico del entorno. La gestión efectiva de parches requiere un proceso continuo que incluya la identificación, la evaluación del riesgo, la planificación de la remediación y la verificación de

su efectividad (Scarfone & Mell, 2022); en SecureNova Labs este proceso se encontraba en un estado incipiente.

El análisis arrojó cuarenta y siete hallazgos distribuidos en los niveles de severidad crítica, alta, media y baja. Los hallazgos críticos y de alta severidad, diecinueve en total, constituyeron el foco de las actividades de explotación. La Figura 2 presenta la distribución por severidad

Figura 2

Distribución de vulnerabilidades por severidad



Nota. Clasificación de los 47 hallazgos según la escala CVSS 3.1.

Tabla 3

Hallazgos críticos y de alta severidad

CVSS	Vulnerabilidad	Activo afectado
9.8	Inyección SQL en módulo de autenticación	Aplicación web
9.1	Ejecución remota de código	Servidor de correo
8.8	Credenciales por defecto sin modificar	Servidor FTP
8.2	Cross-Site Scripting almacenado	Panel de administración
7.8	Escalación de privilegios por cron job	Servidor web
7.4	Cifrado débil (TLS 1.0 / SSL 3.0)	Comunicaciones internas

Nota. Hallazgos priorizados para la fase de explotación controlada.

Entre las vulnerabilidades críticas, la inyección SQL en el módulo de autenticación, con una puntuación de 9.8, permite a un atacante no autenticado extraer el contenido completo de la base de datos de usuarios. El servidor de correo presentó vulnerabilidades de ejecución remota de código con una puntuación promedio de 9.1. El servidor FTP operaba con credenciales por defecto, lo que permitió el acceso a un directorio con archivos de configuración que contenían contraseñas en texto plano de otros sistemas.

Análisis individual de las vulnerabilidades críticas

El detalle de cada vulnerabilidad crítica permite comprender no solo su severidad técnica, sino también su contexto, su explotabilidad y las condiciones que la habilitaron. Este análisis individualizado resulta esencial para orientar la remediación de manera precisa y para extraer lecciones aplicables a la prevención de debilidades similares en el futuro.

La inyección SQL en el módulo de autenticación, con una puntuación de 9.8, constituyó la vulnerabilidad de mayor criticidad y el vector de acceso inicial del ejercicio. Su origen se encuentra en la construcción dinámica de consultas a la base de datos a partir de la entrada del usuario sin la aplicación de mecanismos de saneamiento o de consultas parametrizadas. La explotabilidad de esta vulnerabilidad es elevada, dado que no requiere autenticación previa y puede ejecutarse mediante herramientas automatizadas ampliamente disponibles. Su impacto abarca la confidencialidad, al permitir la extracción de la totalidad de los datos almacenados; la integridad, al habilitar la modificación de registros; y la disponibilidad, al posibilitar la eliminación de información. La remediación requiere la sustitución de las consultas dinámicas por consultas parametrizadas y la implementación de validación de la entrada en el lado del servidor.

La ejecución remota de código en el servidor de correo, con una puntuación promedio de 9.1, representó una vulnerabilidad de criticidad equivalente, aunque de naturaleza distinta. Su origen se encuentra en la operación de una versión desactualizada del software del servidor de correo, afectada por vulnerabilidades conocidas y documentadas para las cuales existían parches disponibles. La explotabilidad de esta vulnerabilidad es alta, dado que las técnicas de explotación se encuentran públicamente documentadas. Su impacto es severo, dado que la ejecución de código en el servidor de correo otorga al adversario el control de un sistema con acceso a comunicaciones sensibles y una posición privilegiada en la red. La remediación requiere la aplicación inmediata de los parches disponibles y el establecimiento de un proceso de gestión de parches que prevenga la acumulación de versiones vulnerables.

Las credenciales por defecto sin modificar en el servidor FTP, con una puntuación de 8.8, ejemplifican una categoría de vulnerabilidad cuyo origen no radica en un defecto del software, sino en una deficiencia en los procesos de configuración. El servidor operaba con las credenciales establecidas por el fabricante, públicamente conocidas, lo que permitió el acceso directo sin necesidad de técnicas de explotación. El impacto de esta vulnerabilidad se vio agravado por el almacenamiento, en el servidor accesible, de archivos de configuración que contenían contraseñas en texto plano de otros sistemas, lo que la convirtió en un punto de pivote hacia activos adicionales. La remediación es directa en términos técnicos, mediante el cambio de las credenciales por defecto, pero requiere el establecimiento de procesos de endurecimiento que prevengan la recurrencia de este tipo de deficiencias.

El Cross-Site Scripting almacenado en el panel de administración, con una puntuación de 8.2, introdujo un vector de ataque dirigido a los usuarios legítimos de la aplicación. Esta vulnerabilidad permite la inyección de código que se almacena en el servidor y se ejecuta en el

navegador de los usuarios que acceden a la sección afectada, lo que habilita el robo de sesiones, la ejecución de acciones en nombre del usuario y la distribución de código malicioso. Su remediación requiere la codificación adecuada de la salida y la implementación de políticas de seguridad de contenido que limiten la ejecución de código no autorizado.

La escalación de privilegios mediante el cron job inseguro, con una puntuación de 7.8, constituyó el eslabón que permitió la transición desde un acceso de bajo privilegio hasta el control administrativo del servidor web. Su origen se encuentra en la combinación de una tarea programada ejecutada con privilegios elevados y la asignación de permisos de escritura sobre el script invocado a usuarios sin privilegios. Esta vulnerabilidad ilustra la importancia del principio de mínimo privilegio y de la auditoría periódica de la configuración del sistema. Su remediación requiere la corrección de los permisos del sistema de archivos y la revisión de todas las tareas programadas ejecutadas con privilegios elevados.

Interpretación y discusión de los hallazgos

La distribución de los hallazgos por nivel de severidad ofrece una lectura significativa sobre el estado de seguridad de SecureNova Labs. La presencia de seis vulnerabilidades críticas y trece de severidad alta, que en conjunto representan el cuarenta por ciento del total de hallazgos, indica una concentración de riesgo elevada que excede lo esperable para una organización de este tamaño y sector. Esta concentración no es producto del azar, sino la consecuencia directa de la ausencia de un proceso sistemático de gestión de vulnerabilidades que permita identificar y remediar las debilidades antes de que se acumulen.

El análisis de la naturaleza de las vulnerabilidades críticas revela que la mayoría de ellas no corresponde a fallos sofisticados o de difícil identificación, sino a debilidades bien conocidas y documentadas que disponen de soluciones de remediación establecidas. La inyección SQL, por

ejemplo, figura entre las vulnerabilidades de aplicaciones web más antiguas y mejor comprendidas de la disciplina, y su presencia en un módulo de autenticación sugiere deficiencias en las prácticas de desarrollo seguro. De manera similar, la existencia de credenciales por defecto sin modificar y de protocolos de cifrado obsoletos refleja la ausencia de líneas base de configuración de seguridad y de procesos de endurecimiento de los sistemas.

La validación manual de los hallazgos automatizados resultó fundamental para depurar los falsos positivos y para contextualizar el riesgo real de cada vulnerabilidad en el entorno específico de SecureNova Labs. Un hallazgo automatizado puede reportar una vulnerabilidad de severidad teórica elevada que, en el contexto particular de la organización, presenta un riesgo real reducido debido a la existencia de controles compensatorios o a las características de la arquitectura. Inversamente, ciertas vulnerabilidades de severidad teórica moderada pueden representar un riesgo crítico en el contexto específico, como ocurrió con la transferencia de zona DNS, que individualmente no constituye una vulnerabilidad crítica pero que, combinada con la ausencia de segmentación, habilitó un mapeo completo de la infraestructura interna.

La interrelación entre las vulnerabilidades identificadas constituye uno de los aspectos más relevantes del análisis. En seguridad ofensiva, el riesgo agregado de un conjunto de vulnerabilidades suele ser superior a la suma de los riesgos individuales, dado que las debilidades pueden encadenarse para construir una ruta de ataque completa. En SecureNova Labs, la combinación de la inyección SQL, la carga de archivos sin validación, el cron job inseguro y la reutilización de credenciales conformó precisamente una cadena de este tipo, que permitió escalar desde un acceso no autenticado hasta el control total del dominio.

Proceso de validación y priorización

El proceso de validación manual de los hallazgos automatizados constituyó una etapa crítica del análisis de vulnerabilidades, orientada a depurar los falsos positivos y a confirmar la explotabilidad real de cada debilidad en el contexto específico del entorno. La validación se realizó mediante la verificación manual de cada hallazgo de severidad crítica y alta, comprobando la existencia de la vulnerabilidad y las condiciones necesarias para su explotación. Este proceso permitió descartar varios hallazgos que, si bien fueron reportados por las herramientas automatizadas, no representaban un riesgo real debido a la existencia de controles compensatorios o a las características particulares de la configuración.

La priorización de las vulnerabilidades validadas se realizó considerando no solo su puntuación de severidad, sino también su explotabilidad en el contexto del entorno, su posición en las posibles cadenas de ataque y el valor de los activos afectados. Este enfoque contextual permitió orientar las actividades de explotación hacia las vulnerabilidades de mayor impacto potencial, optimizando el uso del tiempo disponible y maximizando el valor demostrativo del ejercicio. La priorización contextual representa una práctica superior a la simple ordenación por puntuación de severidad, dado que incorpora las particularidades del entorno evaluado.

La interrelación entre las vulnerabilidades identificadas se analizó de manera explícita, con el fin de identificar las posibles cadenas de ataque que un adversario podría construir mediante el encadenamiento de debilidades individuales. Este análisis de cadenas resultó fundamental para comprender el riesgo agregado del conjunto de vulnerabilidades y para anticipar las rutas de ataque que se materializarían posteriormente durante la fase de explotación controlada.

Clasificación según el marco MITRE ATT&CK

Los hallazgos fueron mapeados al marco MITRE ATT&CK para identificar las tácticas y técnicas que cada vulnerabilidad habilitaría para un adversario real. Este ejercicio resulta de utilidad para el Blue Team, ya que permite orientar la implementación de controles de detección hacia los vectores de ataque más probables. La Tabla 4 resume el mapeo.

Tabla 4

Mapeo de hallazgos al marco MITRE ATT&CK

Táctica	Identificador	Vector habilitado
Acceso inicial	TA0001	Explotación de aplicación web
Ejecución	TA0002	RCE en servidor de correo
Persistencia	TA0003	Creación de cuentas y acceso remoto
Escalación de privilegios	TA0004	Debilidades en cuentas privilegiadas
Movimiento lateral	TA0008	Segmentación insuficiente
Exfiltración	TA0010	Ausencia de controles de prevención de fuga

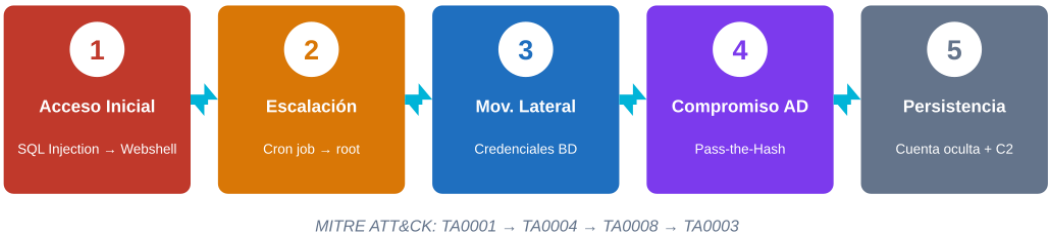
Nota. Correspondencia entre los hallazgos y las tácticas adversarias del marco MITRE ATT&CK (MITRE Corporation, 2023).

Explotación Controlada

La transición del análisis a la explotación controlada requirió una planificación que garantizara la ejecución ordenada de los ataques, la documentación de cada acción y el respeto estricto de los límites éticos y legales del ejercicio. Se elaboró un plan que priorizó las vulnerabilidades de mayor severidad y mayor probabilidad de éxito, con el objetivo de demostrar el impacto potencial de las debilidades sin causar interrupciones en los servicios. La Figura 3 sintetiza la cadena de explotación ejecutada.

Figura 3

Cadena de explotación controlada



Nota. Secuencia de cinco fases desde el acceso inicial hasta la persistencia, con su correspondencia en el marco MITRE ATT&CK.

Acceso inicial

El vector de acceso inicial se obtuvo mediante la explotación de la inyección SQL identificada en el módulo de autenticación de la aplicación web principal. La inyección de una consulta maliciosa en el campo de usuario permitió eludir el mecanismo de autenticación y obtener acceso a la interfaz de administración con privilegios elevados. Desde esa interfaz se identificó una funcionalidad de carga de archivos con validaciones insuficientes, que se aprovechó para cargar una webshell que proporcionó ejecución de comandos en el servidor, con los privilegios del usuario bajo el que operaba el servidor web.

Escalación de privilegios

Con el acceso inicial como usuario de bajo privilegio, se exploraron las vías de escalación disponibles. La enumeración del sistema incluyó la revisión de programas con permisos especiales incorrectos, la búsqueda de tareas programadas ejecutadas con privilegios elevados y la verificación de versiones del núcleo. La escalación se logró aprovechando una tarea programada configurada para ejecutarse con privilegios de administrador, que invocaba un script ubicado en un directorio con permisos de escritura para usuarios no privilegiados. La

sustitución del script permitió obtener una sesión con privilegios de administrador en el servidor web.

Movimiento lateral y compromiso del dominio

El movimiento lateral desde el servidor comprometido hacia otros sistemas se facilitó por la ausencia de segmentación efectiva. Las credenciales del administrador de base de datos, recuperadas tras la escalación, permitieron la conectividad directa con el servidor de base de datos que almacenaba la información de clientes. El análisis de los hashes extraídos reveló el uso de algoritmos débiles que facilitaron la recuperación de un número significativo de contraseñas. El movimiento hacia el directorio activo se realizó por reutilización de credenciales: una de las contraseñas recuperadas resultó válida para una cuenta del dominio, desde la cual se ejecutaron técnicas de volcado de credenciales que proporcionaron acceso a hashes de múltiples cuentas

Post-explotación y persistencia

Alcanzado el nivel de acceso objetivo, se implementaron mecanismos de persistencia para demostrar la capacidad de un adversario de mantener el acceso de manera encubierta. Estos mecanismos, documentados exhaustivamente, fueron comunicados al Blue Team al concluir la fase con el fin de que pudieran ser detectados y eliminados. Incluyeron la creación de una cuenta de dominio de apariencia legítima con privilegios administrativos, la instalación de un agente de comando y control en modalidad simulada y la configuración de una tarea programada de reconocimiento. La revisión de archivos accesibles identificó certificados y claves privadas almacenados sin protección, copias de seguridad con contraseñas en texto claro y documentación técnica de la infraestructura productiva.

Análisis de la cadena de ataque

La cadena de explotación ejecutada en SecureNova Labs ilustra de manera didáctica el concepto de progresión del compromiso, en el cual un adversario avanza de manera incremental desde un punto de entrada inicial de bajo privilegio hasta el control de los activos de mayor valor. Cada fase de la cadena dependió del éxito de la fase anterior y, a su vez, habilitó la fase siguiente, configurando una secuencia en la que la interrupción de cualquier eslabón habría detenido el avance del ataque. Esta característica tiene implicaciones defensivas significativas, dado que sugiere que la implementación de controles efectivos en cualquiera de las fases podría haber prevenido el compromiso total.

El acceso inicial mediante la inyección SQL demuestra la criticidad de las vulnerabilidades en las aplicaciones expuestas a internet, que constituyen el punto de contacto más directo entre la organización y los potenciales adversarios. La posterior carga de la webshell evidencia cómo una vulnerabilidad de validación de archivos, que individualmente podría considerarse de severidad moderada, se transforma en un vector crítico cuando se combina con un acceso administrativo previo. Esta observación refuerza la importancia de evaluar las vulnerabilidades no de manera aislada, sino en el contexto de las rutas de ataque que habilitan.

La escalación de privilegios a través del cron job inseguro pone de manifiesto la relevancia del principio de mínimo privilegio y de la gestión adecuada de los permisos del sistema de archivos. La ejecución de tareas programadas con privilegios elevados que invocan scripts ubicados en directorios modificables por usuarios sin privilegios constituye un error de configuración recurrente que los procesos de endurecimiento de sistemas deberían identificar y corregir. La detección de este tipo de configuraciones requiere auditorías periódicas de la configuración de seguridad, ausentes en SecureNova Labs al momento de la evaluación.

El movimiento lateral y el compromiso del directorio activo representan la fase de mayor impacto de la cadena, dado que el control del directorio activo equivale, en términos prácticos, al control de la totalidad de los recursos del dominio. La técnica de reutilización de credenciales que habilitó este movimiento subraya la importancia de las políticas de contraseñas únicas y de la implementación de autenticación multifactor, que habría neutralizado el vector incluso en presencia de credenciales comprometidas. La ausencia de segmentación entre las zonas de la red amplificó el impacto, al permitir la conectividad directa entre sistemas que, en una arquitectura adecuadamente segmentada, deberían estar aislados.

La fase de persistencia, aunque ejecutada en modalidad simulada y controlada, demuestra la capacidad de un adversario real para establecer un acceso duradero y encubierto al entorno comprometido. Los mecanismos de persistencia implementados habrían permitido a un atacante mantener el control del entorno incluso tras la remediación de la vulnerabilidad de acceso inicial, lo que subraya la importancia de los procesos de análisis forense exhaustivo y de erradicación completa tras la detección de un incidente. La detección de mecanismos de persistencia constituye uno de los aspectos más desafiantes de la respuesta a incidentes y requiere capacidades de análisis avanzadas que SecureNova Labs deberá desarrollar.

Análisis de impacto por dimensión de seguridad

El impacto del compromiso demostrado durante el ejercicio se analiza a continuación según las tres dimensiones fundamentales de la seguridad de la información: la confidencialidad, la integridad y la disponibilidad. Este análisis dimensional permite comprender de manera estructurada las consecuencias potenciales de la materialización de las amenazas identificadas y orientar la priorización de los controles de protección.

En la dimensión de la confidencialidad, el compromiso demostró un impacto severo. El acceso a la base de datos de clientes mediante la inyección SQL, sumado a la recuperación de credenciales y al compromiso del directorio activo, habría permitido a un adversario acceder a la totalidad de la información sensible almacenada por la organización, incluyendo datos personales de clientes, credenciales de acceso y documentación técnica de la infraestructura. La exfiltración de esta información representaría no solo una pérdida de confidencialidad, sino también un incumplimiento de las obligaciones legales de protección de datos personales, con las consecuentes responsabilidades administrativas y civiles.

En la dimensión de la integridad, el control administrativo obtenido sobre los sistemas habría permitido a un adversario modificar la información almacenada, alterar la configuración de los sistemas e introducir código malicioso en los repositorios de código fuente. Esta última posibilidad resulta de particular gravedad en el contexto de SecureNova Labs, dado que la organización desarrolla software para el sector financiero y gubernamental; la introducción de código malicioso en los productos de la organización podría propagar el compromiso a sus clientes, configurando un ataque a la cadena de suministro de software de consecuencias potencialmente catastróficas.

En la dimensión de la disponibilidad, si bien el ejercicio excluyó explícitamente las acciones que pudieran afectar la continuidad de los servicios, el nivel de acceso obtenido habría permitido a un adversario real interrumpir la operación de los sistemas críticos, cifrar la información mediante código de secuestro o destruir los datos almacenados. El impacto sobre la disponibilidad resultaría especialmente grave en ausencia de copias de seguridad adecuadamente protegidas, condición que no pudo verificarse positivamente durante el ejercicio y que constituye un aspecto adicional de preocupación.

El análisis dimensional del impacto confirma que el compromiso demostrado afectaría de manera severa a las tres dimensiones de la seguridad de la información, lo que justifica la clasificación de los escenarios de riesgo asociados en los niveles más altos de la matriz de riesgos. Esta valoración refuerza la urgencia de la implementación de las recomendaciones formuladas, particularmente aquellas orientadas a la protección de la confidencialidad de los datos de clientes y a la prevención de la introducción de código malicioso en los productos de la organización.

Análisis y Respuesta del Blue Team

El equipo defensivo estructuró su respuesta a las actividades del Red Team siguiendo el marco del NIST para la respuesta a incidentes, que define las fases de preparación, detección y análisis, contención, erradicación y recuperación, y actividades posteriores al incidente (NIST, 2012). La preparación incluyó la revisión de los playbooks existentes, la verificación de la operatividad de los controles disponibles y la definición de umbrales de alerta. Esta fase evidenció la falta de documentación actualizada de los procedimientos de respuesta y la inexistencia de criterios formales de clasificación de incidentes.

Detección de Amenazas

La detección de las actividades del Red Team representó el mayor desafío operacional para el Blue Team, en gran medida por las limitaciones de los controles disponibles. La ausencia de un sistema de gestión de información y eventos de seguridad centralizado obligó a revisar manualmente los registros de múltiples fuentes heterogéneas, lo que incrementó el tiempo de detección y la tasa de falsos negativos en las etapas iniciales.

La detección del escaneo de reconocimiento activo se logró mediante la revisión de los registros del firewall perimetral, que registró un patrón de solicitudes a múltiples puertos desde una misma dirección de origen en un intervalo reducido. Este patrón fue identificado aproximadamente cuarenta y cinco minutos después del inicio del escaneo, tiempo superior al umbral deseable. La detección del compromiso inicial del servidor web fue considerablemente más tardía: el acceso a través de la inyección SQL y la carga de la webshell no se detectaron en tiempo real debido a la ausencia de inspección profunda del tráfico y a la insuficiencia de los registros, que no capturaban el cuerpo de las solicitudes. La detección se produjo durante la revisión programada de los registros de acceso de la jornada siguiente.

La detección del movimiento lateral y del acceso al directorio activo se realizó a través de los registros de eventos, que documentaron múltiples intentos de autenticación fallidos seguidos de un inicio de sesión exitoso desde una dirección interna inusual. El sistema de detección de intrusiones de red generó alertas sobre la comunicación entre el servidor web y el controlador de dominio mediante protocolos ajenos al perfil de comportamiento normal del servidor.

Análisis Forense Digital

Fundamentos del análisis forense digital

El análisis forense digital constituye una disciplina especializada que aplica métodos científicos a la identificación, preservación, análisis y presentación de evidencias digitales. En el contexto de la respuesta a incidentes, el análisis forense persigue la reconstrucción de la secuencia de eventos de un ataque, la identificación de los artefactos dejados por el adversario y la determinación del alcance del compromiso. El rigor metodológico del proceso forense resulta esencial, dado que sus resultados pueden tener implicaciones legales y deben ser defendibles ante terceros.

El principio del orden de volatilidad guía la secuencia de adquisición de evidencias, estableciendo que las evidencias más volátiles deben capturarse antes que las menos volátiles, dado que las primeras pueden perderse con mayor facilidad. La memoria de acceso aleatorio de un sistema activo, que contiene información sobre los procesos en ejecución y las conexiones de red establecidas, constituye una de las fuentes más volátiles y debe capturarse de manera prioritaria. Los registros almacenados en medios no volátiles, en cambio, presentan una mayor persistencia y pueden adquirirse en etapas posteriores del proceso.

La preservación de la integridad de las evidencias mediante el cálculo y la verificación de valores hash constituye otro principio fundamental del análisis forense. El cálculo de un valor

hash de cada evidencia en el momento de su adquisición permite verificar posteriormente que la evidencia no ha sido alterada durante el proceso de análisis, lo que garantiza su validez. En el análisis forense de SecureNova Labs, la aplicación de estos principios permitió reconstruir con precisión la secuencia de eventos del ataque y documentar los artefactos de manera defendible.

El análisis forense se ejecutó sobre los sistemas comprometidos con el objetivo de reconstruir la secuencia de eventos, identificar los artefactos dejados por el Red Team y evaluar el alcance del compromiso. El proceso siguió el orden de volatilidad, comenzando por la memoria de los sistemas activos y concluyendo con los registros en medios no volátiles. El análisis de memoria del servidor web permitió identificar los procesos activos en el momento de la adquisición, incluyendo el proceso de la webshell y las conexiones de red establecidas. El análisis del sistema de archivos evidenció los archivos maliciosos cargados, las modificaciones en los scripts aprovechados para la escalación y los registros de comandos ejecutados.

El análisis forense del controlador de dominio se centró en los registros de eventos de seguridad, que documentaron la secuencia de autenticaciones, cambios de privilegios y creación de cuentas realizados durante el movimiento lateral y la post-explotación. El informe forense resultante incluyó una línea de tiempo del ataque, la lista de indicadores de compromiso, el inventario de artefactos y una evaluación del alcance del daño potencial. La Tabla 5 resume los indicadores de compromiso identificados.

Tabla 5

Indicadores de compromiso identificados

Tipo de indicador	Descripción
Archivo malicioso	Webshell cargada en el servidor web
Cuenta de usuario	Cuenta de dominio creada con privilegios administrativos
Tarea programada	Cron Job modificado para escalación de privilegios
Conexión de red	Tráfico hacia servidor de comando y control simulado
Acceso anómalo	Sesión remota desde dirección interna inusual

Nota. Indicadores empleados como referencia durante la erradicación y la verificación post-remediación.

Contención, Erradicación y Recuperación

Las medidas de contención buscaron limitar la propagación del compromiso y aislar los sistemas afectados sin interrumpir innecesariamente los servicios productivos. La contención a corto plazo incluyó el bloqueo de la dirección de origen en el firewall, la desactivación de los puertos de red del servidor comprometido, el restablecimiento forzado de las sesiones activas en el directorio activo y el bloqueo de la cuenta identificada como comprometida.

La erradicación se realizó sobre la base del inventario de artefactos identificado en el análisis forense. Las acciones incluyeron la eliminación de los archivos maliciosos, la restauración de los scripts modificados a partir de copias verificadas, la eliminación de la cuenta maliciosa y la revocación de los tokens de sesión activos. Adicionalmente, se restablecieron las contraseñas de todas las cuentas cuyas credenciales podrían haber sido comprometidas. La recuperación se orientó a restaurar los sistemas a un estado operativo con garantías verificadas, mediante la revisión de los indicadores de compromiso, la comparación de hashes de archivos

críticos con valores de referencia y la ejecución de un escaneo post-remediación. La reintegración del servidor web se realizó tras la aplicación de los parches correspondientes, la corrección de la configuración deficiente y la actualización de las reglas del firewall de aplicaciones web.

La evaluación de la efectividad arrojó resultados mixtos. El tiempo medio de detección para los eventos críticos fue de aproximadamente dieciocho horas, valor superior al estándar recomendado para organizaciones de este perfil. La tasa de detección de los artefactos durante el análisis forense alcanzó el ochenta y siete por ciento, lo que implica que el trece por ciento de los indicadores no fue identificado durante la fase reactiva, hallazgo que subraya la necesidad de capacidades de búsqueda proactiva de amenazas.

Discusión sobre la efectividad defensiva

La evaluación de la efectividad del Blue Team durante el ejercicio ofrece lecciones valiosas sobre las brechas y fortalezas de la postura defensiva de SecureNova Labs. El tiempo de detección del compromiso inicial, superior a dieciocho horas, contrasta marcadamente con la detección del reconocimiento activo, lograda en aproximadamente cuarenta y cinco minutos.

Esta diferencia se explica por la naturaleza de los controles disponibles: el firewall perimetral registró el patrón de escaneo de manera relativamente evidente, mientras que el compromiso a través de la inyección SQL no generó alertas debido a la ausencia de inspección profunda del tráfico de aplicaciones y a la insuficiencia de la configuración de registro.

La dependencia de la revisión manual de registros constituye la limitación estructural más significativa de la capacidad de detección de la organización. En ausencia de un sistema de gestión de información y eventos de seguridad que correlacione automáticamente los eventos provenientes de múltiples fuentes, la detección de incidentes complejos que involucran múltiples

sistemas depende de la capacidad del analista para identificar manualmente patrones dispersos en grandes volúmenes de registros. Este enfoque resulta insostenible frente a adversarios sofisticados y explica directamente los tiempos de detección observados durante el ejercicio.

No obstante, el ejercicio también evidenció fortalezas relevantes en la capacidad de respuesta del Blue Team. Una vez detectado el compromiso, las acciones de contención y erradicación se ejecutaron de manera ordenada y efectiva, siguiendo un proceso estructurado que limitó la propagación del incidente y restauró la integridad de los sistemas afectados. Esta observación sugiere que, si bien las capacidades de detección requieren un fortalecimiento significativo, los fundamentos de los procesos de respuesta se encuentran presentes y constituyen una base sobre la cual desarrollar capacidades más avanzadas.

La tasa de detección de artefactos del ochenta y siete por ciento durante el análisis forense, si bien representa un resultado aceptable, indica que un porcentaje no despreciable de los indicadores de compromiso permaneció sin identificar durante la fase reactiva. En un escenario de ataque real, los artefactos no detectados podrían constituir mecanismos de persistencia que permitirían al adversario recuperar el acceso tras la remediación, lo que subraya la necesidad de complementar la respuesta reactiva con capacidades de búsqueda proactiva de amenazas que permitan identificar comprometimientos no detectados por los controles automatizados.

Hacia un modelo de detección por capas

Los resultados del ejercicio defensivo sugieren la conveniencia de adoptar un modelo de detección estructurado en capas, que combine controles de distinta naturaleza para maximizar la probabilidad de identificar la actividad maliciosa en cualquier fase del ciclo de ataque. Este enfoque, conocido como defensa en profundidad aplicada a la detección, reconoce que ningún

control de detección resulta infalible y que la combinación de múltiples mecanismos complementarios incrementa significativamente la cobertura defensiva.

La primera capa, correspondiente a la detección perimetral, comprende los controles que monitorean el tráfico entre la organización y el exterior, como el firewall de próxima generación y los sistemas de detección de intrusiones de red. Durante el ejercicio, esta capa demostró su efectividad en la detección del reconocimiento activo, aunque resultó insuficiente frente al compromiso a través de la aplicación web, lo que evidencia la necesidad de complementarla con capas adicionales orientadas a la detección de actividad en las capas de aplicación y de sistema.

La segunda capa, correspondiente a la detección a nivel de aplicación, comprende los controles que inspeccionan el tráfico y el comportamiento de las aplicaciones expuestas, como el firewall de aplicaciones web y el registro detallado de las solicitudes. La ausencia de inspección profunda del tráfico de aplicaciones constituyó la brecha determinante que permitió que el vector de acceso inicial pasara desapercibido, lo que subraya la prioridad de fortalecer esta capa en el contexto de SecureNova Labs.

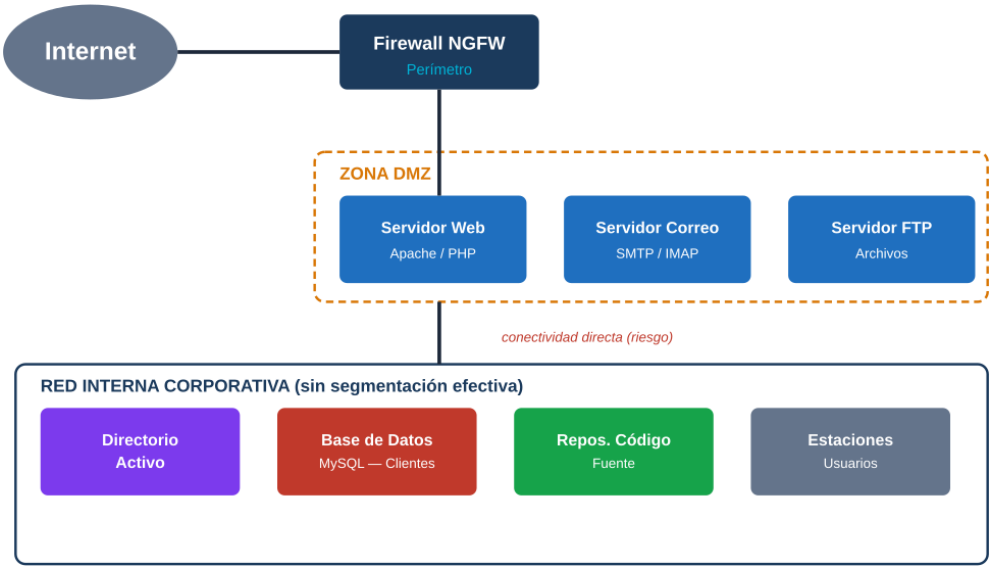
La tercera capa, correspondiente a la detección a nivel de sistema y de identidad, comprende los controles que monitorean la actividad en los sistemas internos y el comportamiento de las cuentas de usuario, como las soluciones de detección y respuesta en los puntos finales y el análisis del comportamiento de las entidades. El fortalecimiento de esta capa permitiría detectar las fases más avanzadas del ciclo de ataque, como la escalación de privilegios, el movimiento lateral y el establecimiento de la persistencia, que durante el ejercicio resultaron especialmente difíciles de identificar. La integración de las tres capas en una plataforma de correlación centralizada constituye el objetivo hacia el cual debe orientarse la evolución de las capacidades de detección de la organización.

Trazabilidad entre Hallazgos y Controles

Con el fin de fortalecer la relación entre los hallazgos ofensivos y las acciones defensivas, se construyó una matriz de trazabilidad que vincula cada vulnerabilidad explotada con el impacto observado durante el ejercicio y el control de mitigación propuesto para el Blue Team. Esta trazabilidad permite evidenciar de manera explícita cómo cada debilidad identificada se traduce en una acción concreta de fortalecimiento, evitando que el análisis quede en un plano exclusivamente descriptivo. La Figura 4 presenta esta relación.

Figura 4

Trazabilidad vulnerabilidad-impacto-control



Nota. Relación entre cada vulnerabilidad explotada por el Red Team, su impacto observado y el control de mitigación implementado por el Blue Team.

La trazabilidad establecida demuestra que las acciones del Blue Team no respondieron de manera genérica, sino que se orientaron específicamente a neutralizar los vectores explotados.

Así, la inyección SQL se mitiga mediante consultas parametrizadas y reglas del firewall de aplicaciones web; la ejecución remota en el servidor de correo, mediante parcheo urgente y segmentación; y el compromiso del directorio activo, mediante la implementación de

autenticación multifactor y el monitoreo continuo. Este enfoque garantiza la coherencia entre el diagnóstico ofensivo y la respuesta defensiva.

Marco Legal Aplicable

El ejercicio de evaluación se desarrolla en un contexto jurídico definido por normas nacionales e internacionales que regulan el tratamiento de los sistemas de información, la protección de datos personales y las actividades de seguridad informática. El conocimiento de este marco resulta esencial tanto para la legitimidad de las actividades de evaluación como para la definición de las obligaciones de la organización.

Marco Normativo en Colombia

La Ley 1273 de 2009 constituye el referente legal fundamental en Colombia para la tipificación de conductas que afectan la seguridad de los sistemas de información. Esta ley incorporó al Código Penal un título dedicado a la protección de la información y de los datos, estableciendo sanciones para conductas como el acceso abusivo a sistemas, la obstaculización ilegítima, la interceptación de datos y la instalación de software malicioso (Congreso de Colombia, 2009). En el ejercicio de Red Team y Blue Team, las actividades de explotación se ampararon en la autorización expresa del escenario, lo que confiere legitimidad a las acciones realizadas dentro del alcance definido.

La Ley 1581 de 2012 establece el régimen de protección de datos personales en Colombia. Esta norma es de especial relevancia para SecureNova Labs, dado que la organización procesa y almacena datos personales de sus clientes. Las vulnerabilidades identificadas, particularmente aquellas que habilitan la exfiltración de bases de datos, representan riesgos directos de incumplimiento de las obligaciones de seguridad del tratamiento de datos personales (Congreso de Colombia, 2012). El documento de política nacional de seguridad digital, por su parte, define los lineamientos para la gestión de riesgos en el entorno digital y proporciona

orientaciones sobre las prácticas mínimas de ciberseguridad esperadas (Consejo Nacional de Política Económica y Social [CONPES], 2016).

Análisis detallado de la Ley 1273 de 2009

La Ley 1273 de 2009 representa un hito en la legislación colombiana en materia de delitos informáticos, al crear un bien jurídico tutelado específico para la protección de la información y de los datos. La norma tipifica una serie de conductas que resultan directamente relevantes para el ejercicio de seguridad realizado. El acceso abusivo a un sistema informático, tipificado en la norma, describe precisamente la conducta que un adversario real ejecutaría al explotar las vulnerabilidades identificadas en SecureNova Labs. La distinción jurídica fundamental entre esta conducta delictiva y el ejercicio legítimo de Red Team radica en la autorización: las actividades de evaluación se ejecutaron en el marco de una autorización expresa que delimitó su alcance, lo que excluye la antijuridicidad de las acciones.

La interceptación de datos informáticos y el daño informático, igualmente tipificados, describen otras conductas que un adversario podría ejecutar tras comprometer la infraestructura. La relevancia de esta norma para el contexto de SecureNova Labs es doble: por una parte, fundamenta la legitimidad del ejercicio de evaluación autorizado; por otra, establece el marco de las consecuencias jurídicas que enfrentaría un adversario real que explotara las vulnerabilidades identificadas, lo que refuerza la importancia de su remediación.

Análisis detallado de la Ley 1581 de 2012

La Ley 1581 de 2012 establece el régimen general de protección de datos personales en Colombia, imponiendo a los responsables y encargados del tratamiento una serie de obligaciones orientadas a garantizar la seguridad de la información personal. El principio de seguridad, consagrado en la norma, exige la implementación de las medidas técnicas, humanas y

administrativas necesarias para otorgar seguridad a los registros, evitando su adulteración, pérdida, consulta, uso o acceso no autorizado o fraudulento.

Las vulnerabilidades identificadas en SecureNova Labs comprometen directamente la capacidad de la organización para cumplir con este principio. La inyección SQL que permite la exfiltración de la base de datos de clientes, la ausencia de cifrado robusto en las comunicaciones y la gestión inadecuada de credenciales constituyen deficiencias en las medidas técnicas de seguridad que la norma exige. La materialización de cualquiera de estos riesgos podría configurar un incumplimiento de las obligaciones legales de la organización, con las consecuentes responsabilidades administrativas ante la autoridad de protección de datos y la potencial responsabilidad civil frente a los titulares de los datos afectados.

Referentes Internacionales

A nivel internacional, el estándar ISO/IEC 27001 proporciona el marco de referencia más adoptado para la gestión de la seguridad de la información, definiendo los requisitos para establecer, implementar y mejorar un sistema de gestión de seguridad de la información (International Organization for Standardization [ISO], 2022). La evaluación reveló que la organización no cuenta con un sistema de gestión formalmente implementado, lo que implica la ausencia de múltiples controles del anexo de la norma. El NIST Cybersecurity Framework, con su estructura en cinco funciones (identificar, proteger, detectar, responder y recuperar), ofrece un lenguaje común para comunicar el nivel de madurez en ciberseguridad y fue utilizado como referente para la evaluación del estado de SecureNova Labs (NIST, 2018).

Implicaciones Legales de los Hallazgos

Los hallazgos tienen implicaciones legales directas en al menos tres dimensiones. En primer lugar, la existencia de vulnerabilidades que permiten la exfiltración de datos personales

configura un riesgo de responsabilidad bajo el régimen de protección de datos, que exige medidas técnicas y organizativas adecuadas. En segundo lugar, la ausencia de controles básicos podría considerarse negligencia en el cumplimiento de las obligaciones de seguridad establecidas en los contratos con clientes del sector financiero y gubernamental. En tercer lugar, la gestión inadecuada de credenciales privilegiadas incrementa el riesgo de que un compromiso de SecureNova Labs sea utilizado como vector para ataques a sus clientes, lo que podría derivar en responsabilidades contractuales adicionales.

Análisis de Riesgos

El análisis de riesgos adoptó el enfoque de evaluación cualitativa y semicuantitativa, que permite identificar los activos de mayor valor, determinar las amenazas y vulnerabilidades asociadas y calcular el nivel de riesgo. Para cada riesgo se calculó un valor inherente mediante la relación entre la probabilidad de ocurrencia y el impacto potencial, valorados en una escala del uno al cinco, lo que genera una matriz de cinco por cinco con valores entre uno y veinticinco.

Posteriormente se estimó el riesgo residual considerando la efectividad de los controles existentes.

Tabla 6

Matriz de riesgos de los escenarios críticos

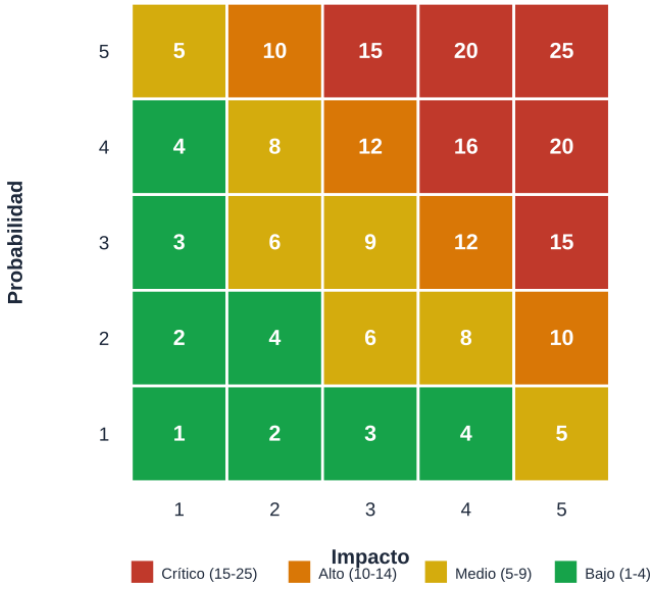
Escenario de riesgo	Prob.	Impacto	Nivel	Prioridad
Explotación SQL y exfiltración	4	5	20	Crítico
Compromiso de directorio activo	4	5	20	Crítico
Detección tardía de incidentes	5	4	20	Crítico
Exposición de credenciales	4	4	16	Alto
Persistencia de acceso	3	4	12	Alto

Nota. Valoración de los escenarios de riesgo según probabilidad e impacto. Adaptado de ISO/IEC 27005.

La representación visual de la matriz facilita la comprensión de la distribución de los escenarios según su probabilidad e impacto. La Figura 5 ubica gráficamente los niveles de riesgo en la escala de cinco por cinco.

Figura 5

Matriz de riesgos



Nota. Distribución de los niveles de riesgo según probabilidad e impacto. Los escenarios críticos se concentran en la región superior derecha.

La comparación del riesgo inherente con el riesgo residual evidencia que la mayoría de los controles implementados en SecureNova Labs tienen una efectividad reducida, por lo que el riesgo residual se mantiene en niveles inaceptables para la mayoría de los escenarios identificados.

Discusión sobre el perfil de riesgo

El perfil de riesgo resultante del análisis evidencia una concentración de escenarios críticos que demanda una atención prioritaria por parte de la dirección de SecureNova Labs. Los tres escenarios clasificados con el máximo nivel de riesgo (explotación de la inyección SQL, compromiso del directorio activo y detección tardía de incidentes) comparten una característica

común: su materialización tendría un impacto severo sobre la confidencialidad de los datos de clientes, activo cuya protección constituye una obligación legal y contractual de la organización.

La comparación entre el riesgo inherente y el riesgo residual ofrece una medida de la efectividad de los controles existentes. En la mayoría de los escenarios evaluados, la diferencia entre ambos valores resultó reducida, lo que indica que los controles actualmente implementados aportan una mitigación limitada del riesgo. Esta observación refuerza la conclusión de que las inversiones en seguridad de la organización deben orientarse no solo a la corrección de vulnerabilidades específicas, sino al desarrollo de capacidades de control transversales que reduzcan el riesgo de manera sistemática.

La metodología de análisis de riesgos empleada, basada en la valoración de la probabilidad y el impacto, presenta la ventaja de su simplicidad y comunicabilidad, lo que facilita la comprensión de los resultados por parte de audiencias no técnicas como la alta dirección. No obstante, su naturaleza cualitativa introduce un grado de subjetividad en la valoración que debe ser gestionado mediante la documentación de los criterios empleados y la participación de múltiples evaluadores. En etapas de mayor madurez, la organización podría complementar este enfoque con metodologías cuantitativas que expresen el riesgo en términos monetarios y faciliten el análisis de costo-beneficio de las inversiones en seguridad.

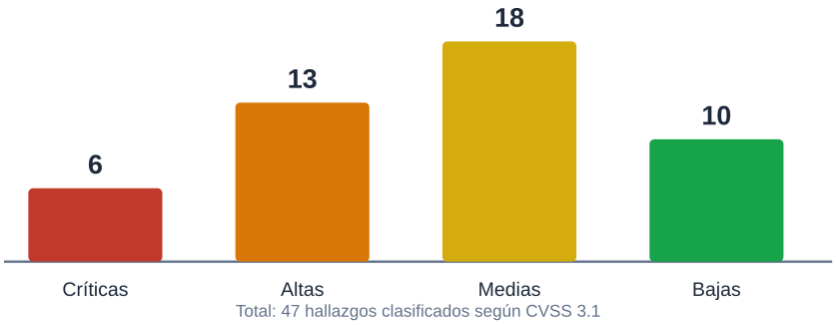
Evaluación del Nivel de Madurez

La evaluación del nivel de madurez se realizó utilizando el modelo del NIST Cybersecurity Framework, que clasifica el nivel de implementación de los controles en cuatro niveles, desde parcial hasta adaptable. La organización fue evaluada en cada una de las cinco funciones del marco. En las funciones identificar y detectar fue clasificada en el nivel parcial, reflejando la ausencia de inventarios completos, la inexistencia de un proceso formal de gestión

de riesgos y las limitaciones en las capacidades de monitoreo. En las funciones proteger, responder y recuperar fue clasificada en el nivel informado por riesgos, dado que existen controles y procedimientos parcialmente implementados, aunque su efectividad demostrada fue inferior a la esperada. La Figura 6 sintetiza esta evaluación.

Figura 6

Nivel de madurez NIST CSF



Nota. Clasificación de la organización en cada función del marco. Las funciones identificar y detectar presentan el menor nivel de madurez.

Los resultados evidencian que las principales brechas se concentran en la capacidad de detección temprana, la correlación de eventos entre múltiples fuentes y la respuesta coordinada ante incidentes multisistema. La implementación de un sistema de gestión de información y eventos de seguridad, el fortalecimiento del monitoreo continuo y el entrenamiento regular del equipo de respuesta emergen como las prioridades de mejora más críticas.

Lecciones Aprendidas y Discusión Integral

El ejercicio de seguridad sobre SecureNova Labs generó un conjunto de lecciones aprendidas que trascienden los hallazgos técnicos específicos y que resultan aplicables tanto a la organización evaluada como a entornos de características similares. Estas lecciones constituyen uno de los productos de mayor valor del ejercicio, dado que orientan la mejora continua de las prácticas de seguridad y de la metodología de evaluación.

Sobre la Postura de Seguridad Organizacional

La primera lección, y quizás la más significativa, se relaciona con la naturaleza de las vulnerabilidades identificadas. El hecho de que la mayoría de las debilidades críticas correspondiera a problemas conocidos y bien documentados, con soluciones de remediación disponibles, evidencia que el principal desafío de SecureNova Labs no reside en la sofisticación de las amenazas, sino en la ausencia de procesos básicos de gestión de la seguridad. Esta observación sugiere que las inversiones de mayor impacto no son necesariamente las de mayor costo tecnológico, sino aquellas orientadas a establecer procesos sistemáticos de gestión de vulnerabilidades, endurecimiento de configuraciones y gestión de parches.

La segunda lección se relaciona con la importancia de la arquitectura de la red como factor determinante del impacto de un compromiso. La ausencia de segmentación efectiva transformó un compromiso inicial limitado en un compromiso total del entorno, al permitir el movimiento lateral sin restricciones. Esta observación refuerza la relevancia de los principios de diseño seguro y de la arquitectura de confianza cero, que limitan el radio de alcance de un compromiso mediante la segmentación y la verificación continua de la confianza.

Sobre las Capacidades de Detección y Respuesta

La tercera lección se relaciona con la criticidad de las capacidades de detección. El contraste entre la detección relativamente rápida del reconocimiento activo y la detección tardía del compromiso a través de la aplicación web evidencia que la efectividad de la detección depende directamente de la cobertura y la profundidad de los controles de monitoreo. La ausencia de inspección profunda del tráfico de aplicaciones constituyó una brecha determinante que permitió que el vector de acceso inicial pasara desapercibido durante un período prolongado.

La cuarta lección se relaciona con la complementariedad entre las capacidades reactivas y proactivas. La tasa de detección de artefactos inferior al cien por ciento durante el análisis forense evidencia que la respuesta reactiva, por rigurosa que sea, puede resultar insuficiente para identificar la totalidad de los artefactos dejados por un adversario. Esta observación subraya la necesidad de complementar la respuesta reactiva con capacidades de búsqueda proactiva de amenazas, que permitan identificar los comprometimientos que los controles automatizados no detectan.

Sobre la Metodología de Evaluación

La quinta lección se relaciona con el valor de la colaboración entre los equipos ofensivo y defensivo. La comunicación de las técnicas empleadas por el Red Team al Blue Team, una vez concluidas las fases de explotación, permitió a este último desarrollar y validar las capacidades de detección correspondientes, generando un aprendizaje que difícilmente se habría obtenido mediante una evaluación puramente adversaria. Esta observación valida el enfoque Purple Team como un modelo de mayor valor formativo y operacional para organizaciones en proceso de desarrollo de sus capacidades de seguridad.

La sexta lección se relaciona con la importancia de la trazabilidad entre los hallazgos y las acciones de remediación. El establecimiento de una matriz de trazabilidad que vincula cada vulnerabilidad con su impacto y su control de mitigación correspondiente permitió garantizar que las acciones defensivas respondieran de manera específica a los vectores explotados, evitando la dispersión de los esfuerzos de remediación y facilitando la priorización de las inversiones en seguridad.

En conjunto, estas lecciones configuran un marco de referencia para la mejora continua de la postura de seguridad de SecureNova Labs y para la evolución de la metodología de evaluación. Su incorporación en los procesos de la organización permitirá no solo remediar las vulnerabilidades identificadas, sino también desarrollar las capacidades necesarias para prevenir, detectar y responder eficazmente ante futuras amenazas.

Consideraciones Éticas y Limitaciones del Estudio

La ejecución de un ejercicio de seguridad ofensiva, aun en un contexto autorizado y controlado, plantea consideraciones éticas que deben ser explícitamente abordadas para garantizar la legitimidad y la responsabilidad de las actividades realizadas. Así mismo, la naturaleza del ejercicio impone una serie de limitaciones que condicionan el alcance de los hallazgos y que deben ser reconocidas para una interpretación adecuada de los resultados.

Consideraciones Éticas

El principio fundamental que rige la ética de las actividades de seguridad ofensiva es el de la autorización expresa. Todas las actividades de reconocimiento, análisis y explotación se ejecutaron dentro del alcance autorizado del ejercicio, definido en el escenario planteado, sin exceder en ningún momento los límites establecidos. Esta autorización constituye el elemento que distingue las actividades legítimas de evaluación de seguridad de las conductas ilícitas tipificadas en la legislación, y su documentación rigurosa resulta esencial para la defensa de la legitimidad de las acciones realizadas.

El principio de minimización del daño orientó la ejecución de las actividades de explotación, que se diseñaron para demostrar el impacto potencial de las vulnerabilidades sin causar afectaciones reales a la disponibilidad de los servicios ni a la integridad de la información productiva. Los mecanismos de persistencia se implementaron en modalidad simulada y se documentaron exhaustivamente para su posterior eliminación, evitando dejar el entorno en un estado comprometido tras la conclusión del ejercicio. Esta práctica refleja el compromiso ético de no generar perjuicios que excedan los necesarios para la consecución de los objetivos de la evaluación.

El principio de confidencialidad rige el tratamiento de la información obtenida durante el ejercicio. Los hallazgos, las vulnerabilidades identificadas y la información sensible a la que se accedió durante las actividades de explotación constituyen información de carácter reservado, cuya divulgación podría facilitar su aprovechamiento por parte de adversarios reales. El presente informe, en consecuencia, describe los hallazgos en un nivel de detalle suficiente para orientar la remediación, evitando la inclusión de información técnica específica que pudiera facilitar la reproducción de los ataques por parte de terceros no autorizados.

Limitaciones del Estudio

La primera limitación del estudio se relaciona con el carácter puntual de la evaluación. El ejercicio proporciona una fotografía del estado de seguridad de la organización en un momento determinado, que no captura necesariamente las variaciones que pudieran producirse a lo largo del tiempo como consecuencia de cambios en la infraestructura, la incorporación de nuevos sistemas o la evolución de las amenazas. Esta limitación refuerza la recomendación de establecer un programa regular de evaluaciones que permita mantener una comprensión actualizada de la postura de seguridad.

La segunda limitación se relaciona con el alcance definido para el ejercicio. La exclusión de las pruebas de denegación de servicio y de las acciones que pudieran afectar la disponibilidad de los servicios productivos, si bien necesaria para preservar la continuidad operativa, implica que ciertos aspectos de la resiliencia de la organización no fueron evaluados. De manera similar, el alcance no incluyó la evaluación de la seguridad física ni de los aspectos relacionados con la ingeniería social dirigida al personal, dimensiones que un adversario real podría explotar y que ameritarían una evaluación complementaria.

La tercera limitación se relaciona con la naturaleza del escenario, que reproduce las condiciones de una organización real, pero en un contexto controlado. Si bien el escenario se diseñó para reflejar condiciones realistas, ciertos aspectos del comportamiento de una organización real ante un incidente, como la presión temporal, la coordinación entre múltiples áreas y la comunicación con las partes interesadas, no pueden reproducirse plenamente en un contexto de evaluación. Esta limitación sugiere la conveniencia de complementar los ejercicios técnicos con ejercicios de simulación de crisis que evalúen las dimensiones organizacionales de la respuesta a incidentes.

El reconocimiento de estas limitaciones no reduce el valor de los hallazgos obtenidos, sino que contextualiza su interpretación y orienta la planificación de evaluaciones futuras que aborden las dimensiones no cubiertas por el presente ejercicio. La combinación de evaluaciones técnicas periódicas, ejercicios de simulación de crisis y evaluaciones de las dimensiones física y humana de la seguridad proporcionaría una comprensión integral de la postura de seguridad de la organización.

Evidencias de Sustentación

En cumplimiento de los requisitos de la Etapa 5 del Seminario Especializado, se presenta el video de sustentación del informe final, en el cual se expone el proceso general del seminario, las estrategias aplicadas por el Red Team y el Blue Team, y las conclusiones y recomendaciones derivadas del ejercicio realizado sobre el escenario de SecureNova Labs. El video cumple con las especificaciones de duración establecidas y con la presencia en cámara del estudiante. El enlace de acceso se presenta a continuación:

Video de sustentación del informe final: https://go.screenpal.com/watch/cO1Zr4nu9yF

El enlace fue verificado para su visualización sin restricciones de acceso desde una sesión externa, de manera que pueda ser consultado por los evaluadores durante el proceso de revisión.

Conclusiones

El ejercicio de seguridad ofensiva evidenció vulnerabilidades críticas en múltiples capas de la infraestructura de SecureNova Labs. La cadena de compromiso, que avanzó desde una inyección SQL hasta el control total del directorio activo, demostró que las debilidades identificadas habrían permitido el compromiso total del entorno por parte de un adversario real en un tiempo reducido, lo que confirma la importancia de la estrategia del Red Team para revelar riesgos que los controles convencionales no detectan.

Las limitaciones en las capacidades de detección representan la principal brecha defensiva de la organización. El tiempo de detección superior a dieciocho horas en un escenario controlado con un único atacante sugiere que, frente a un ataque real con múltiples vectores, la capacidad de respuesta del Blue Team sería significativamente menor. La estrategia defensiva demostró su valor en la contención y erradicación, pero requiere fortalecer la detección temprana.

El análisis del marco legal confirmó que las vulnerabilidades generan riesgos de responsabilidad significativos, particularmente en relación con las obligaciones de protección de datos personales. La ausencia de un sistema de gestión de seguridad formalmente implementado implica que la organización no dispone de los controles ni de las evidencias documentales necesarias para demostrar el cumplimiento de sus obligaciones.

La confrontación entre las estrategias ofensivas del Red Team y las capacidades defensivas del Blue Team generó un conocimiento contextualizado sobre las brechas de seguridad que difícilmente se habría obtenido mediante revisiones técnicas convencionales. La trazabilidad establecida entre vulnerabilidades, impacto y controles confirma la coherencia entre el diagnóstico y la respuesta, y refuerza el valor de la colaboración entre ambos equipos.

Finalmente, el ejercicio confirma que el valor de una evaluación de seguridad no reside únicamente en la identificación de vulnerabilidades, sino en la capacidad de traducir los hallazgos en un plan de mejora accionable y medible. El conjunto de recomendaciones priorizadas, el plan de implementación por etapas y los indicadores de gestión propuestos constituyen una hoja de ruta que permite a SecureNova Labs avanzar de manera ordenada desde su nivel de madurez inicial hacia una postura de seguridad sostenible. La adopción de un enfoque de mejora continua, sustentado en la repetición periódica de ejercicios de evaluación y en la medición objetiva de los indicadores definidos, resulta determinante para que las mejoras implementadas se consoliden y para que la organización desarrolle la resiliencia necesaria frente a la evolución constante del panorama de amenazas.

Conclusión Integradora

De manera integral, SecureNova Labs presenta un nivel de madurez inicial en ciberseguridad, ubicado en el nivel parcial en las funciones de identificación y detección del NIST Cybersecurity Framework, y en el nivel informado por riesgos en las funciones de protección, respuesta y recuperación. Las principales brechas se concentran en la ausencia de un sistema de gestión de información y eventos de seguridad, la segmentación insuficiente de la red y la gestión inadecuada de credenciales. La implementación progresiva de las recomendaciones propuestas permitiría a la organización avanzar hacia un nivel repetible de madurez, reducir el riesgo residual a niveles aceptables y demostrar el cumplimiento de sus obligaciones legales y contractuales, fortaleciendo así su posición competitiva frente a los clientes del sector financiero y gubernamental.

Recomendaciones

Las recomendaciones se organizan según su horizonte temporal y su criterio de priorización, distinguiendo entre las acciones de ejecución inmediata, las que dependen de recursos adicionales y las que forman parte de una estrategia de madurez a largo plazo. La Tabla 7 sintetiza esta priorización.

Tabla 7

Priorización de recomendaciones

Recomendación	Horizonte	Criterio
Parchear CVEs críticos	0–30 días	Ejecución inmediata
Implementar autenticación multifactor	0–30 días	Ejecución inmediata
Restablecer credenciales comprometidas	0–30 días	Ejecución inmediata
Implementar SIEM centralizado	1–6 meses	Requiere recursos
Segmentar la red interna	1–6 meses	Requiere recursos
Implementar SGSI ISO/IEC 27001	6–18 meses	Madurez organizacional
Programa de concienciación	6–18 meses	Madurez organizacional

Nota. Clasificación de las recomendaciones por horizonte temporal y criterio de priorización.

Acciones de Ejecución Inmediata: En los primeros treinta días deben aplicarse los parches de seguridad pendientes en los sistemas con vulnerabilidades críticas y altas, priorizando el servidor de correo, el servidor web y las aplicaciones con inyección SQL. La corrección de la inyección SQL debe abordarse mediante la sustitución de las consultas dinámicas por consultas parametrizadas, acompañada de un proceso de revisión de código de seguridad. La implementación de autenticación multifactor para todos los accesos remotos y cuentas privilegiadas reduce significativamente el riesgo de compromiso de identidad, incluso en presencia de contraseñas débiles (Grassi et al., 2017).

Así mismo, deben restablecerse las credenciales comprometidas y corregirse la configuración del servidor DNS para deshabilitar las transferencias de zona no autorizadas.

Acciones que Requieren Recursos Adicionales: En un horizonte de uno a seis meses, la implementación de una plataforma de gestión de información y eventos de seguridad centralizada constituye una prioridad operacional, integrando los registros de todos los activos críticos y configurando reglas de correlación orientadas a las técnicas identificadas en el ejercicio. La segmentación efectiva de la red interna, mediante un modelo de zonas de seguridad, reduce el radio de alcance de un compromiso inicial y limita el movimiento lateral, en línea con los principios de la arquitectura de confianza cero (Rose et al., 2020). Adicionalmente, debe implementarse un proceso formal de gestión de vulnerabilidades y un programa de seguridad en el ciclo de desarrollo de software.

Estrategia de Madurez a Largo Plazo: En un horizonte de seis a dieciocho meses, la organización debe iniciar la implementación de un sistema de gestión de seguridad de la información alineado con el estándar ISO/IEC 27001, con un alcance inicial que cubra los activos de mayor criticidad. El establecimiento de ejercicios regulares de Red Team y Blue Team, complementados con ejercicios colaborativos de tipo Purple Team, permitirá evaluar de manera continua la efectividad de los controles. El factor humano continúa siendo el vector de ataque de mayor efectividad para los adversarios, por lo que la inversión en la educación y el entrenamiento de los usuarios representa uno de los controles preventivos de mayor relación costo-efectividad (Hadnagy, 2018). La adopción del modelo Purple Team maximiza la transferencia de conocimiento entre los equipos ofensivo y defensivo (Mathos & Jung, 2023).

Desarrollo de las recomendaciones prioritarias: Cada recomendación prioritaria se desarrolla a continuación incorporando su justificación técnica, los criterios de éxito que

permitirán evaluar su implementación y las consideraciones para su ejecución. Este nivel de detalle busca facilitar la traducción de las recomendaciones en acciones concretas y medibles por parte de la organización.

La implementación de un sistema de gestión de información y eventos de seguridad constituye la recomendación de mayor impacto sobre la capacidad de detección de la organización. Su justificación se fundamenta directamente en el principal hallazgo del ejercicio defensivo: el tiempo de detección superior a dieciocho horas resultó de la ausencia de correlación automatizada de eventos. La implementación debe contemplar la integración de los registros de todos los activos críticos, la definición de reglas de correlación orientadas a las técnicas identificadas en el ejercicio y el establecimiento de un proceso de monitoreo continuo. El criterio de éxito principal consiste en la reducción del tiempo medio de detección a un valor inferior a una hora para los eventos de severidad crítica, métrica que deberá validarse mediante la repetición de ejercicios ofensivos controlados.

La implementación de autenticación multifactor para los accesos remotos y las cuentas privilegiadas constituye una recomendación de alta relación costo-efectividad. Su justificación se fundamenta en que esta medida habría neutralizado el vector de reutilización de credenciales que habilitó el compromiso del directorio activo, dado que el conocimiento de una contraseña resulta insuficiente para la autenticación cuando se exige un segundo factor. La implementación debe priorizar las cuentas administrativas y los accesos desde redes externas. El criterio de éxito consiste en la cobertura del cien por ciento de las cuentas privilegiadas y los accesos remotos, verificable mediante una auditoría de configuración.

La segmentación efectiva de la red interna aborda una de las condiciones estructurales que amplificaron el impacto del ejercicio ofensivo. Su justificación se fundamenta en que la

conectividad directa entre las zonas de la red permitió el movimiento lateral desde el servidor web comprometido hacia los activos críticos internos. La implementación debe contemplar la definición de zonas de seguridad basadas en la criticidad y la función de los activos, así como la aplicación de controles de filtrado entre las zonas que restrinjan la comunicación a lo estrictamente necesario. El criterio de éxito consiste en la imposibilidad de establecer comunicación directa entre la zona desmilitarizada y las subredes de activos críticos, verificable mediante pruebas de conectividad controladas.

El establecimiento de un proceso formal de gestión de vulnerabilidades aborda la causa raíz de la acumulación de hallazgos identificada en el ejercicio. Su justificación se fundamenta en que la mayoría de las vulnerabilidades críticas correspondían a debilidades conocidas con soluciones de remediación disponibles, lo que evidencia la ausencia de un proceso sistemático de identificación y corrección. La implementación debe contemplar escaneos periódicos, la priorización de la remediación según el riesgo y la verificación de la efectividad de las correcciones. El criterio de éxito consiste en la reducción sostenida del número de vulnerabilidades críticas y altas no remediadas, medida mediante escaneos periódicos cuyos resultados deberán mostrar una tendencia decreciente.

Consideraciones de Endurecimiento de Sistemas: El endurecimiento de los sistemas constituye una práctica transversal cuya ausencia se evidenció de manera recurrente en los hallazgos del ejercicio. Esta práctica consiste en la reducción de la superficie de ataque de cada sistema mediante la eliminación de los componentes innecesarios, la corrección de las configuraciones inseguras por defecto y la aplicación de las líneas base de seguridad reconocidas en la industria. La adopción sistemática del endurecimiento como parte del proceso de

despliegue de los sistemas habría prevenido varias de las vulnerabilidades identificadas, particularmente aquellas originadas en configuraciones deficientes.

En el ámbito de los servidores, el endurecimiento debe contemplar la deshabilitación de los servicios no utilizados, la modificación de las credenciales por defecto, la aplicación del principio de mínimo privilegio en la asignación de permisos y la configuración adecuada de los mecanismos de registro de eventos. La existencia de credenciales por defecto sin modificar en el servidor FTP y la configuración insegura de las tareas programadas en el servidor web evidencian la ausencia de estas prácticas en SecureNova Labs, lo que justifica la incorporación del endurecimiento como una recomendación de carácter estructural.

En el ámbito de las aplicaciones web, el endurecimiento debe contemplar la implementación de las cabeceras de seguridad estándar, la configuración de políticas de seguridad de contenido, la deshabilitación de la exposición de información sensible en los mensajes de error y la aplicación de las prácticas de desarrollo seguro a lo largo del ciclo de vida del software. La ausencia de cabeceras de seguridad y la exposición de directorios de trabajo identificadas durante el reconocimiento reflejan la necesidad de fortalecer estas prácticas, que constituyen controles preventivos de bajo costo y alta efectividad.

La verificación periódica del estado de endurecimiento de los sistemas, mediante auditorías de configuración automatizadas que comparen el estado de cada sistema con las líneas base definidas, permite identificar y corregir las desviaciones antes de que sean aprovechadas por un adversario. Esta práctica, integrada en el proceso de gestión de vulnerabilidades, cierra el ciclo de mejora continua de la configuración de seguridad y previene la reaparición de las deficiencias corregidas.

Plan de Implementación por Etapas: La traducción de las recomendaciones en acciones concretas requiere un plan de implementación que ordene las actividades según su prioridad, sus dependencias y los recursos necesarios. El plan propuesto se estructura en tres etapas que corresponden a los horizontes temporales definidos en la priorización de las recomendaciones, estableciendo para cada una los objetivos, las actividades principales y los criterios de verificación que permitirán evaluar su cumplimiento.

Etapa de estabilización inmediata: La primera etapa, con un horizonte de treinta días, persigue la neutralización de los riesgos críticos mediante la remediación de las vulnerabilidades que habilitaron la cadena de compromiso demostrada durante el ejercicio. Las actividades principales comprenden la aplicación de los parches de seguridad pendientes en los sistemas críticos, la corrección de la inyección SQL mediante la implementación de consultas parametrizadas, el restablecimiento de las credenciales comprometidas, la corrección de las configuraciones inseguras identificadas y el despliegue de autenticación multifactor para las cuentas privilegiadas y los accesos remotos. El criterio de verificación de esta etapa consiste en la confirmación, mediante un escaneo de validación, de que las vulnerabilidades críticas han sido remediadas y de que los vectores de acceso inicial demostrados durante el ejercicio ya no resultan explotables.

Esta etapa reviste un carácter de máxima urgencia, dado que las vulnerabilidades que aborda constituyen riesgos activos que un adversario real podría explotar en cualquier momento. La ejecución de estas acciones no requiere inversiones significativas ni la adquisición de nuevas tecnologías, sino la asignación de recursos del personal técnico existente a las tareas de remediación, lo que la convierte en una etapa de alta relación entre el impacto sobre la reducción del riesgo y el costo de su implementación.

Etapa de fortalecimiento estructural: La segunda etapa, con un horizonte de uno a seis meses, persigue el desarrollo de las capacidades estructurales de detección y contención cuya ausencia se evidenció durante el ejercicio. Las actividades principales comprenden la implementación de una plataforma de gestión de información y eventos de seguridad, la segmentación efectiva de la red interna, el establecimiento de un proceso formal de gestión de vulnerabilidades y la implementación de un programa de seguridad en el ciclo de desarrollo de software. El criterio de verificación de esta etapa consiste en la demostración, mediante la repetición de un ejercicio ofensivo controlado, de una reducción significativa del tiempo de detección y de una limitación efectiva de la capacidad de movimiento lateral.

Esta etapa requiere inversiones en tecnología y, posiblemente, en la incorporación de personal especializado o en la contratación de servicios gestionados de seguridad. La implementación de la plataforma de gestión de eventos de seguridad, en particular, constituye un proyecto de complejidad media que requiere la integración de las fuentes de registro, la configuración de las reglas de correlación y el establecimiento de los procesos operativos de monitoreo, por lo que su planificación debe contemplar un período de maduración hasta alcanzar la plena operatividad.

Etapa de consolidación de la madurez: La tercera etapa, con un horizonte de seis a dieciocho meses, persigue la consolidación de un nivel de madurez sostenible mediante la formalización de los procesos de gestión de la seguridad. Las actividades principales comprenden la implementación de un sistema de gestión de seguridad de la información alineado con el estándar ISO/IEC 27001, el establecimiento de un programa regular de ejercicios de Red Team y Blue Team, y el desarrollo de un programa de concienciación y formación en seguridad para el personal de la organización. El criterio de verificación de esta etapa consiste en la

obtención de evidencias documentales del funcionamiento de los procesos de gestión de la seguridad y en la demostración de una tendencia sostenida de mejora en los indicadores de seguridad.

Esta etapa transforma las mejoras puntuales de las etapas anteriores en un sistema de gestión sostenible que garantiza el mantenimiento y la evolución continua de la postura de seguridad. La implementación del sistema de gestión de seguridad de la información, además de su valor operativo, proporciona a la organización la capacidad de demostrar el cumplimiento de sus obligaciones ante sus clientes del sector financiero y gubernamental, lo que constituye un factor de diferenciación competitiva en estos mercados.

Indicadores de Gestión de la Seguridad: La sostenibilidad de las mejoras propuestas requiere el establecimiento de un conjunto de indicadores que permitan medir de manera objetiva la evolución de la postura de seguridad de la organización a lo largo del tiempo. La definición de indicadores cuantificables transforma la gestión de la seguridad de un ejercicio reactivo en un proceso medible y orientado a resultados, lo que facilita la rendición de cuentas ante la dirección y la demostración del retorno de las inversiones realizadas.

El tiempo medio de detección constituye el indicador de mayor relevancia para evaluar la capacidad defensiva de la organización. Este indicador, que durante el ejercicio se situó por encima de las dieciocho horas para los eventos críticos, debe reducirse progresivamente hasta alcanzar valores inferiores a una hora, meta que solo resulta alcanzable mediante la implementación de capacidades de correlación automatizada de eventos. La medición periódica de este indicador, mediante la repetición de ejercicios ofensivos controlados, permite verificar la efectividad de las inversiones en capacidades de detección.

El tiempo medio de respuesta complementa al indicador anterior, midiendo el intervalo transcurrido entre la detección de un incidente y su contención efectiva. Durante el ejercicio, las acciones de contención se ejecutaron de manera ordenada una vez detectado el compromiso, lo que sugiere una base adecuada sobre la cual desarrollar capacidades de respuesta más ágiles. La formalización de los procedimientos de respuesta y la realización de ejercicios de simulación periódicos permitirían reducir este indicador y mejorar la coordinación entre los equipos involucrados.

El porcentaje de vulnerabilidades críticas remediadas dentro de los plazos establecidos constituye un indicador de la madurez del proceso de gestión de vulnerabilidades. La definición de plazos máximos de remediación según la severidad de cada vulnerabilidad, junto con la medición del cumplimiento de dichos plazos, permite identificar las brechas en el proceso de remediación y orientar las acciones de mejora. La cobertura de los controles de seguridad, expresada como el porcentaje de activos críticos sujetos a monitoreo continuo y a evaluaciones periódicas de seguridad, completa el conjunto de indicadores fundamentales para la gestión de la postura de seguridad.

Continuidad Operativa y Resiliencia: Si bien el ejercicio se centró en la evaluación de las capacidades de prevención, detección y respuesta, los hallazgos obtenidos tienen implicaciones directas sobre la capacidad de la organización para mantener la continuidad de sus operaciones ante un incidente de seguridad. La resiliencia, entendida como la capacidad de una organización para mantener sus funciones esenciales durante y después de un incidente, constituye una dimensión complementaria de la postura de seguridad que SecureNova Labs debe desarrollar.

La existencia de copias de seguridad adecuadamente protegidas y verificadas constituye el control fundamental para la recuperación ante incidentes que comprometan la disponibilidad o

la integridad de la información. Durante el ejercicio se identificaron copias de seguridad con contraseñas almacenadas en texto claro, lo que evidencia deficiencias en la protección de este control crítico. La implementación de un esquema de copias de seguridad que contemple el almacenamiento de copias aisladas de la red, la verificación periódica de su capacidad de restauración y la protección de su confidencialidad e integridad resulta indispensable para garantizar la recuperación ante un incidente.

La elaboración y el mantenimiento de un plan de continuidad del negocio y de un plan de recuperación ante desastres permiten a la organización responder de manera ordenada ante incidentes de gran magnitud, minimizando el tiempo de interrupción de las operaciones y las pérdidas asociadas. Estos planes deben definir los procedimientos de recuperación, los tiempos objetivo de recuperación, los responsables de cada actividad y los recursos necesarios, y deben someterse a pruebas periódicas que validen su efectividad. La realización de ejercicios de simulación de incidentes, que evalúen no solo las capacidades técnicas sino también las dimensiones organizacionales y de comunicación de la respuesta, completa el desarrollo de la capacidad de resiliencia de la organización.

Referencias Bibliográficas

Alhamed, M., Rash, M., & Al-Garadi, M. A. (2023). *A systematic literature review on penetration testing in network environments*. Applied Sciences, 13(12), 6986. https://doi.org/10.3390/app13126986

Congreso de Colombia. (2009). *Ley 1273 de 2009. Por medio de la cual se modifica el Código Penal y se crea un nuevo bien jurídico tutelado denominado de la protección de la información y de los datos*. Diario Oficial No. 47.223.

Congreso de Colombia. (2012). *Ley 1581 de 2012. Por la cual se dictan disposiciones generales para la protección de datos personales*. Diario Oficial No. 48.587.

Consejo Nacional de Política Económica y Social. (2016). *Política nacional de seguridad digital (Documento CONPES 3854)*. Departamento Nacional de Planeación. https://colaboracion.dnp.gov.co/CDT/Conpes/Economicos/3854.pdf

Grassi, P. A., García, M. E., & Fenton, J. L. (2017). *Digital identity guidelines (NIST SP 800-63-3).* National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-63-3

Hadnagy, C. (2018*). Social engineering: The science of human hacking* (2nd ed.). Wiley.

International Organization for Standardization. (2022). Information security, cybersecurity and privacy protection: Information security management systems. Requirements (ISO/IEC 27001:2022).

Mathos, M., & Jung, C. (2023). *Purple team strategies: Enhancing global security posture*. Packt Publishing.

MITRE Corporation. (2023). MITRE ATT&CK: *Adversarial tactics, techniques, and common knowledge framework.* https://attack.mitre.org/

National Institute of Standards and Technology. (2012). *Computer security incident handling guide* (NIST SP 800-61 Rev. 2). U.S. Department of Commerce. https://doi.org/10.6028/NIST.SP.800-61r2

National Institute of Standards and Technology. (2018*). Framework for improving critical infrastructure cybersecurity* (Version 1.1). https://doi.org/10.6028/NIST.CSWP.04162018

Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). *Zero trust architecture (NIST SP 800-207).* National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-207

Scarfone, K., & Mell, P. (2022). *Guide to enterprise patch management planning (NIST SP 800-40 Rev. 4)*. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-40r4

Apéndices

Apéndice A

Comandos y Herramientas Utilizadas

Este apéndice presenta el respaldo técnico del proceso práctico desarrollado durante el ejercicio. Durante el ejercicio se emplearon herramientas de reconocimiento, análisis de vulnerabilidades y explotación de uso estándar en la industria. El reconocimiento pasivo se apoyó en herramientas de enumeración de subdominios y consulta de fuentes abiertas. El reconocimiento activo utilizó escáneres de red y de aplicaciones web. El análisis de vulnerabilidades se realizó con escáneres automatizados complementados con validación manual.

Tabla 8

Herramientas utilizadas por fase

Fase	Tipo de herramienta	Propósito
Reconocimiento pasivo	Enumeración de subdominios y OSINT	Recopilación de información pública
Reconocimiento activo	Escáneres de red y de aplicaciones web	Identificación de servicios y versiones
Análisis de vulnerabilidades	Escáneres automatizados	Detección y clasificación de hallazgos
Explotación controlada	Frameworks de pruebas de penetración	Validación del impacto de las vulnerabilidades
Análisis forense	Herramientas de análisis de memoria y disco	Reconstrucción de la secuencia de eventos

Nota. Categorías de herramientas empleadas en cada fase del ejercicio, sin afectar la disponibilidad de los servicios productivos.

La fase de explotación empleó frameworks de pruebas de penetración en entornos controlados. Las herramientas se ejecutaron exclusivamente dentro del alcance autorizado del

ejercicio, sin afectar la disponibilidad de los servicios en producción. La Tabla 8 resume las herramientas empleadas en cada fase.

La selección de las herramientas se orientó por el principio de utilizar instrumentos de uso estándar en la industria, ampliamente documentados y validados por la comunidad profesional, con el fin de garantizar la reproducibilidad de los resultados y la confiabilidad de los hallazgos. Durante todas las fases del ejercicio se mantuvo un registro detallado de los comandos ejecutados, los parámetros empleados y los resultados obtenidos, lo que constituye la base documental que respalda las conclusiones del informe y que permitiría la verificación independiente de los hallazgos por parte de los analistas de la organización.

En las actividades de análisis forense se aplicaron los principios de la cadena de custodia de las evidencias digitales, documentando la adquisición, el almacenamiento y el análisis de cada evidencia de manera que se preservara su integridad y su valor probatorio. El cálculo de valores hash en el momento de la adquisición y su verificación posterior garantizaron que las evidencias no fueran alteradas durante el proceso de análisis, lo que resulta esencial para la validez de las conclusiones forenses y para su eventual utilización en procesos de carácter legal o disciplinario.

A modo de evidencia técnica del proceso práctico, la Tabla 9 presenta una selección representativa de los comandos ejecutados durante las distintas fases del ejercicio. Los valores específicos de direcciones, dominios y credenciales se han sustituido por marcadores genéricos, en concordancia con los principios de confidencialidad y minimización del daño que rigieron la evaluación.

Tabla 9

Comandos representativos ejecutados por fase

Fase	Comando	Propósito
Reconocimiento	nmap -sV -O -p- <objetivo>	Identificación de servicios, versiones y sistema operativo
Enumeración DNS	dig axfr @<servidor_dns> <dominio>	Intento de transferencia de zona DNS
Enumeración web	gobuster dir -u <url> -w <wordlist>	Descubrimiento de directorios y rutas ocultas
Análisis de vulnerabilidades	nikto -h <url>	Detección de configuraciones inseguras en el servidor web
Explotación SQLi	sqlmap -u <url> --batch --dbs	Validación y explotación de la inyección SQL
Escalación de privilegios	find / -perm -4000 -type f2>/dev/null	Búsqueda de binarios con permisos especiales
Post-explotación	hashdump (módulo de post-explotación)	Extracción de hashes de credenciales
Análisis forense	sha256sum <evidencia>	Cálculo del valor hash para la cadena de custodia

Nota. Selección representativa de comandos con valores sensibles sustituidos por marcadores.

Esta documentación de los comandos, junto con las herramientas descritas y los resultados consignados a lo largo del informe, constituye el respaldo técnico del proceso práctico desarrollado y permite la verificación independiente de la metodología empleada por parte de los analistas de la organización.