UNIVERSIDAD POLITECNICA DE VALENCIA

DEPARTAMENTO DE COMUNICACIONES

Generación de ciberinteligencia con Splunk

Autor: Osmany González de Juana

Director: Manuel Esteve Domingo

Fecha de comienzo: 01/01/2021

Lugar de trabajo: Valencia, España

Objetivos

Analizar y evaluar las capacidades de la plataforma Splunk Enterprise para importar e indexar información referente a ciberseguridad, procesarla y generar ciberinteligencia, además se evalúan las alternativas de exportarla a otras plataformas.

Metodología

Para la realización de este proyecto, inicialmente se efectuó una revisión del estado del arte de la materia tratada, se consultaron diferentes fuentes bibliográficas de referencia y se realizaron entrenamientos para adquirir habilidades imprescindibles. Luego se diseñó y desplegó un entorno virtual donde originar los ciber eventos. Posteriormente se llevaron a cabo varios ciber ataques controlados hacia un objetivo puntual, y los registros de esta actividad fue importada desde Splunk para generar ciberinteligencia, la cual se exportó exitosamente mediante diversas vías.

Desarrollos teóricos realizados

Se estudiaron los conceptos fundamentales sobre ciberseguridad y ciberinteligencia a través de consultas a fuentes de elevado prestigio internacional. También se examinó el estado del arte, adquiriendo un amplio conocimiento sobre el panorama actual: tipos de ciber amenazas y sus efectos, estándares y protocolos de ciberseguridad, softwares de seguridad, etc. Además, se realizó una investigación acerca de técnicas de *ethical hacking* y *penetration testing* con la finalidad de llevar a cabo la explotación de vulnerabilidades y ciber ataques para posteriormente generar ciberinteligencia con Splunk.

Desarrollo de prototipos y trabajo de laboratorio

Se desarrolló un entorno virtual en VMware Workstation Pro en el cual efectuar ciber ataques y pruebas de penetración desde Kali Linux hacia un objetivo puntual, Metasploitable 2, y empleando un *firewall* de nueva generación que intercepte esta actividad maliciosa, para luego enviarla a la instancia de Splunk Enterprise y generar ciberinteligencia.

Resultados

Se generó ciberinteligencia con Splunk Enterprise a través de búsquedas específicas, alarmas, reportes y *dashboards* con información detallada sobre los sucesos ocurridos en el entorno virtual empleado. También, se comprobaron las capacidades de Splunk Enterprise para exportar esta ciberinteligencia por distintas vías, permitiendo así la integración con otras plataformas.

Líneas futuras

Se propone realizar un proyecto con una perspectiva similar, donde se utilicen otras herramientas más específicas para el contexto de la ciberseguridad ofrecidas por Splunk bajo licencias de pago. Además, se sugiere el uso de otras soluciones más potentes, tanto para el *firewall* de nueva generación como para las pruebas de penetración y ejecución de los ciber ataques. También se recomienda al empleo del entorno virtual aquí desarrollado, o alguno equivalente, para actividades docentes de carácter práctico.

Resumen

El creciente desarrollo de las TICs y de los servicios sobre internet ha provocado el incremento de los usuarios y dispositivos conectados a nivel global, estrechándose el vínculo entre el ciberespacio y la vida cotidiana. Paralelamente el cibercrimen, el ciberterrorismo y otras actividades maliciosas han aumentado drásticamente, y las técnicas empleadas por sus autores son cada vez más eficaces y sofisticadas. Para hacer frente a esta situación es imprescindible mantener una Ciber Conciencia Situacional. En este sentido la ciberinteligencia juega un papel fundamental, por lo que en el presente trabajo se propone la generación de la misma mediante la plataforma Splunk Enterprise. Primeramente se efectúa un estudio de los principales conceptos sobre el asunto de la ciberseguridad, se analizan las características de la plataforma y se realiza una breve comparación con otras herramientas empleadas para fines similares. Luego se implementa un entorno virtualizado en el cual se realizan ciber ataques controlados, y los registros de esta actividad son enviados a Splunk para su indexado y procesamiento. Una vez hecho esto, se genera ciberinteligencia mediante búsquedas (*Search Processing Language*) que son guardadas como alertas, reportes y dashboards y se exporta esta información a través de diferentes vías, brindando la posibilidad de integración con otras plataformas.

Abstract

The growing development of ICTs and Internet services has led to an increase in the number of users and devices connected globally, strengthening the relationship between cyberspace and everyday life. At the same time, cybercrime, cyberterrorism and other malicious activities have increased dramatically, and the techniques used by their perpetrators are becoming more effective and sophisticated. To deal with this situation, it is essential to maintain a cyber situational awareness. In this sense, cyber-intelligence plays a fundamental role, therefore, in this paper we propose the generation of cyber-intelligence using the Splunk Enterprise platform. First, a study of the main concepts on the issue of cybersecurity is carried out, the characteristics of the platform are analyzed and a brief comparison is made with other tools used for similar purposes. Next, a virtualized environment is implemented where controlled cyber-attacks are performed, and the logs of this activity are sent to Splunk for indexing and processing. Once this is accomplished, cyber intelligence is generated through searches (*Search Processing Language*) that are saved as alerts, reports and dashboards and this information is exported through different channels providing the possibility of integration with other platforms.

Resum

El creixent desenvolupament de les TICs i dels serveis sobre internet ha provocat l'increment dels usuaris i dispositius connectats a nivell global, estrenyent-se el vincle entre el ciberespai i la vida quotidiana. Paral·lelament el cibercrim, el ciberterrorisme i altres activitats malèvoles han augmentat dràsticament, i les tècniques empleades pels seus autors són cada vegada més eficàcies i sofisticades. Per fer front a aquesta situació és imprescindible mantindré una ciber consciència situacional. En aquest sentit la ciberintel·ligència juga un paper fonamental. Per tant, al present treball es proposa la generació d'aquesta mitjançant la plataforma Splunk Enterprise. En primer lloc, s'efectua un estudi dels principals conceptes sobre l'assumpte de la ciberseguretat, s'analitzen les característiques de la plataforma i es realitzen una breu comparació amb altres ferramentes empleades per a fins similars. Després, s'implementa un entorn virtualitzat en el qual es realitzen ciber atacs controlats, i registres d'aquesta activitat son enviats a Splunk per al seu processament. Una vegada fet açò, es genera ciberintel·ligència mitjançant recerques (*Search Processing Language*) que són guardades com alertes, reports i dashboards i s'exporta aquesta informació a través de diferents vies brindant la possibilitat d'integració amb altres plataformes.

Autor: Osmany González de Juana …, email: osgonde@teleco.upv.es

Director: Manuel Esteve Domingo …, email: mesteve@dcom.upv.es

Fecha de entrega: 02-07-21

ÍNDICE

I. Introducción

El creciente desarrollo de las Tecnologías de la Información y las Comunicaciones (TICs) permite que cada vez más personas accedan a internet para hacer uso de múltiples servicios. En el informe Measuring Digital Development 2019 [1] se recogen varias estadísticas proporcionadas por la Unión Internacional de Telecomunicaciones (UIT) que reflejan lo antes expuesto (Fig.1).

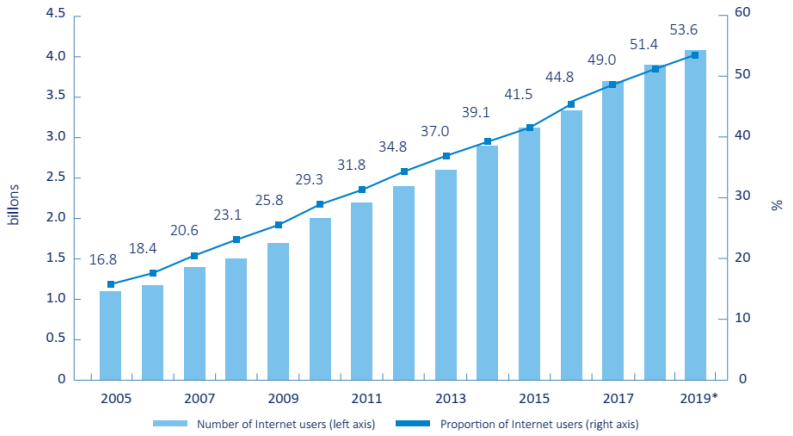

Fig. 1. Cantidad de personas (miles de millones) con acceso Internet en por año en el período 2005-2019

Como se puede observar en la Fig.1, y según refiere la fuente, hasta 2019 tenían acceso a internet alrededor de 4.1 miles de millones de personas, aproximadamente un 36.8% más que en el año 2005. En el Measuring Digital Development 2020 [2] se exhiben resultados semejantes. El informe anual de Internet de Cisco publicado el pasado año (2020) [3] recoge estadísticas desde el 2018 hasta la fecha actual y además realiza pronósticos hasta 2023 (Fig.3), en el mismo se expone el crecimiento de los dispositivos conectados a internet en el período en cuestión,

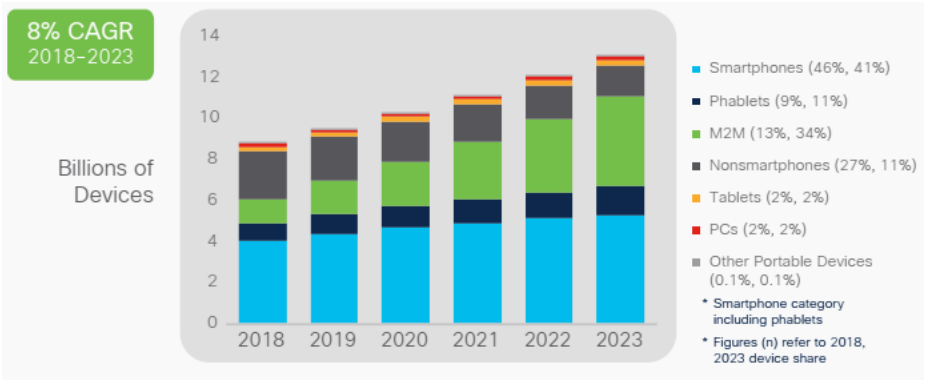

Fig. 2. Crecimiento global de conexiones por dispositivos.

Este incremento del uso de la internet, las TICs y de los servicios que brindan, más los nuevos paradigmas como la computación en la nube y el Internet de las Cosas (IoT), provocan que continuamente se origine y maneje con mayor facilidad un cuantioso volumen de información a través del ciberespacio. Estos factores han provocado el auge y la evolución de actividades maliciosas, perpetradas por actores motivados por distintos intereses sociales, políticos y económicos. Piratas informáticos o "*black hat hackers*", como los define Palo Alto Networks, Inc. en [4], que pueden ser categorizados como: ciberdelincuentes, ciberterroristas, hacktivistas o grupos de *hackers* afiliados a un gobierno. Todos ellos con su acción afectan severamente la ciberseguridad en torno a individuos, empresas, entidades gubernamentales, e incluso naciones.

El National Institute of Standards and Technology (NIST) define a la ciberseguridad [5] como: *"Prevención de daños, protección y restauración de computadoras, sistemas de comunicaciones electrónicas, servicios de comunicaciones electrónicas, comunicaciones por cable y comunicaciones electrónicas, incluida la información contenida en ellas, para garantizar su disponibilidad, integridad, autenticación, confidencialidad y no repudio."*

El NIST también en [6] refiere como el mantenimiento de la ciberseguridad está estrechamente relacionado con el de la seguridad de la información, la que en manos de actores malintencionados, puede ser empleada para provocar numerosos daños económicos, sociales, reputacionales e incluso poner en peligro la vida de personas.

El pasado año (2020), el mundo entero sufrió los embates de la pandemia del SARS-CoV-2. Para enfrentarla, la mayoría de los gobiernos adoptaron medidas de confinamiento y distanciamiento social, provocando que muchos trabajadores realizaran sus tareas mediante el teletrabajo, y que diversos sectores de la economía dependieran aún más de la internet y las TICs. Esta situación ofreció una gran oportunidad para los ciberdelincuentes como se refiere en el 2020 Internet Crime Report proporcionado por el FBI [7]. También en [8] la Secretaría General de INTERPOL describe como ha aumentado la ciberdelincuencia de manera global debido a la pandemia del coronavirus.

I.1. Motivación

Progresivamente los ciberdelincuentes perfeccionan sus métodos, herramientas y estrategias de ataque para que este sea mucho más efectivo y sigiloso. Ejemplo de esto es el empleo de técnicas como el *phishing*, el *spamming*; las *botnets*; los *malwares* avanzados con características como: arquitectura distribuida y tolerante a fallos, multifuncionalidad, polimorfismo y metamorfismo, ofuscación; las Amenazas Persistentes Avanzadas (siglas en inglés APT) que son mucho más deliberadas y potencialmente devastadoras, frecuentemente explotando Vulnerabilidades de Día

Zero. Lo antes mencionado, sumado a la creciente complejidad de las redes actuales, la computación en la nube, los dispositivos IoT y políticas empresariales como *Bring Your Own Device* (BYOD) hacen que las metodologías tradicionales de ciberseguridad perimetral sean insuficientes para detectar y erradicar un ciberataque. En consecuencia, Palo Alto Networks, Inc. y el NIST proponen el empleo de la arquitectura de seguridad de Zero-Trust [4][9]. En este nuevo paradigma, uno de los pilares fundamentales es el registro, inspección y análisis de tráfico de toda la red.

Dicho esto, y en consecuencia con lo estudiado en la asignatura de Ciberinteligencia, impartida en el Máster en Tecnologías, Sistemas y Redes de Comunicaciones, en este trabajo se propone realizar un estudio sobre la Generación de ciberinteligencia mediante el uso de la plataforma Splunk Enterprise.

I.2. Objetivos

Para la realización de este proyecto se definieron los siguientes objetivos:

1. Explicar las características y potencialidades que presenta Splunk, en especial para tareas de ciberseguridad y ciberinteligencia. También realizar un breve análisis y comparación con otras herramientas similares presentes en el mercado.
2. Diseñar y desplegar un entorno donde poner a prueba las funcionalidades de Splunk como herramienta para la generación de ciberinteligencia.
3. Analizar y comprobar las potencialidades de adquisición de datos que ofrece Splunk.
4. Analizar y comprobar las potencialidades de procesamiento de datos y generación de ciberinteligencia que brinda Splunk.
5. Analizar y comprobar las potencialidades de exportación de los datos procesados por Splunk.

I.3. Metodología y planificación del trabajo

Dados los objetivos previamente planteados, el enfoque adoptado es mayoritariamente práctico, y en el mismo fue consecuente la siguiente metodología:

1. Documentación y revisión del estado del arte, primer acercamiento a Splunk:
    - Ampliación de los conocimientos de ciberseguridad y ciberinteligencia mediante consultas a bibliografía de prestigio en el sector (NIST, MITRE, etc.) y realización de cursos teóricos de entrenamiento fundamental (Palo Alto Networks Cybersecurity Foundation, Fortinet NSE1 y NSE2).

- Investigación sobre Splunk, sus características y funcionamiento. Realización del curso “Splunk Fundamentals Part 1”.
- Instalación de Splunk Enterprise bajo licencia de pruebas gratuita por 60 días. Configuración de la instancia y familiarización con la misma.

2. Preparación y puesta en marcha del entorno de simulación:
   - Búsqueda y recopilación de información necesaria para desplegar el entorno en donde llevar a cabo las simulaciones y generar ciberinteligencia con Splunk. Evaluación de recursos de hardware disponibles y de las mejores soluciones adaptables a los mismos.
   - Instalación y configuración del software de virtualización elegido para alojar el entorno del laboratorio: VMware Workstation Pro.
   - Instalación y configuración inicial de la solución de *firewall* de nueva generación seleccionada: Pfsense Community Edition.
   - Aprendizaje y familiarización con Pfsense Community Edition y sus características.
   - Instalación y configuración de las máquinas virtuales (VMs) del entorno de simulación. Se instaló una VM con sistema operativo Kali Linux, y otra con Metasploitable 2 de Rapid7 (Ubuntu).
   - Aprendizaje y familiarización con Kali Linux y con las herramientas de “*penetration testing*” incorporadas en esta distribución de Linux.

3. Simulación en el entorno recreado y evaluación de las capacidades de Splunk:
   - Realización de ciberataques controlados en el entorno simulado para generar registros (*logs*) de actividad maliciosa.
   - Importación y manipulación de *logs* en Splunk para generar ciberinteligencia.
   - Exportación de información de ciberinteligencia desde Splunk.
   - Conclusiones sobre los resultados obtenidos.

II. Fundamentos teóricos y visión general de Splunk

II.1. Fundamentos Teóricos

El progresivo desarrollo y la constante evolución de las amenazas en el ciberespacio hacen que el manejo de la ciberseguridad sea un asunto sumamente complicado. Para hacer frente a esto es necesario la adopción de una postura proactiva, donde se tomen las medidas pertinentes y se asuman buenas prácticas que impidan o minimicen los efectos de una posible agresión.

El NIST propone el *"Marco para mejorar la ciberseguridad de la infraestructura crítica"* [10]. Su adopción permite a las organizaciones (independientemente de su tamaño, grado de riesgo o la sofisticación de la ciberseguridad) aplicar los principios y las mejores prácticas de gestión de riesgos para mejorar la seguridad y la resiliencia.

Fig. 3. Funciones del Marco para mejorar la ciberseguridad de la infraestructura crítica del NIST.

En la Fig.3 se muestran las funciones del núcleo de este marco: Identificar, Proteger, Detectar, Responder y Recuperar; las cuales están orientadas a ayudar a la organización en la gestión de la ciberseguridad. En este sentido se recomienda la realización de estas funciones de forma simultánea y continua, fomentando una *"cultura operativa"* que aborde el riesgo dinámico de la ciberseguridad. Un concepto muy importante para este contexto es el de "Conciencia Situacional" (*Situational Awareness*), definido en el glosario de términos del Comité de Sistemas de Seguridad Nacional de los EE.UU. (CNSS) [11] como:

"Dentro de un volumen de tiempo y espacio, la percepción de la postura de seguridad de una empresa y su entorno de amenazas; la comprensión / significado de ambos tomados en conjunto (riesgo); y la proyección de su estado en un futuro próximo."

Vale la pena destacar el hecho de que el progreso tecnológico ha creado un nuevo escenario, el ciberespacio, en el cual existen numerosas amenazas que evolucionan constantemente y acarrean innumerables peligrosl CNSS[11] lo define como:

"La red interdependiente de infraestructuras de tecnología de la información, incluyendo la Internet, las redes de telecomunicaciones, los sistemas informáticos y los procesadores y controladores integrados en industrias críticas."

El NIST también se refiere al ciberespacio como [12]:

"El entorno complejo resultante de la interacción de personas, software y servicios en Internet mediante dispositivos tecnológicos y redes conectadas a esta, que no existe en una forma física."

Partiendo de las definiciones anteriores, la Conciencia Situacional aplicada al ciberespacio (*Cyber Situational Awareness*) es planteada en [13] de la siguiente manera:

"(…) comprender el estado actual y la postura de seguridad con respecto a la disponibilidad, confidencialidad e integridad de redes, sistemas, usuarios y datos, así como proyectar estados futuros de estos."

En [14] se refiere que la Ciber Conciencia Situacional consta de, al menos, siete aspectos claves:

1. Conciencia de la Situación actual (percepción de la situación).
2. Conciencia del impacto del ataque (evaluación de impacto).
3. Conciencia de cómo evolucionan las situaciones. El seguimiento de las mismas es un componente importante de este aspecto.
4. Conciencia del comportamiento del actor (adversario).
5. Conciencia de por qué y cómo se produce la situación actual.
6. Conciencia de la calidad y confiabilidad de los elementos recopilados sobre la información de la situación, y de las decisiones de conocimiento-inteligencia derivadas de estos.
7. Evaluar los futuros plausibles de la situación actual.

En la Fig.4 se ilustra la propuesta de los autores en [15], un ciclo de vida de cinco fases para la Ciber Conciencia Situacional: análisis y comprensión de la situación ciberespacial, adopción de

políticas de seguridad y control, monitorización del ciberespacio, respuesta ante potenciales amenazas, recopilación e integración de inteligencia.

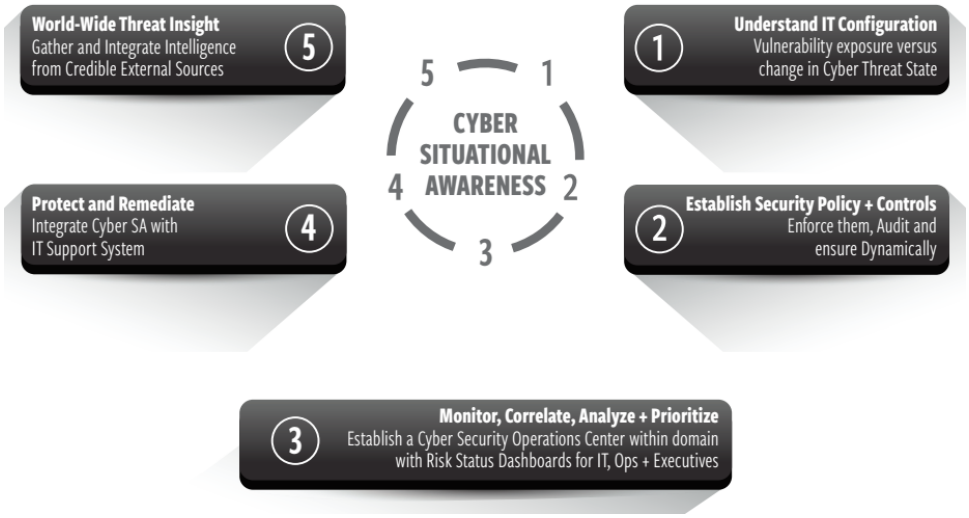

Fig. 4. Ciclo de vida de la Ciber Conciencia Situacional

Según los conceptos previamente planteados, es posible derivar que la generación de inteligencia aplicada al plano cibernético es un factor clave para alcanzar y mantener la Ciber Conciencia Situacional, más aún en el contexto convulso que hoy en día presenta el ciberespacio. El Departamento de Defensa de los EE.UU. define el término “inteligencia” en [16] de la siguiente manera:

“Es el producto resultante de la recopilación, procesamiento, integración, evaluación, análisis e interpretación de la información disponible sobre naciones extranjeras, fuerzas o elementos hostiles o potencialmente hostiles, o áreas de operaciones reales o potenciales.”

De manera específica en [17] se define la ciberinteligencia de amenazas como:

"Conjunto de datos recopilados, evaluados y aplicados con respecto a amenazas de seguridad, los actores de amenazas, exploits, malware, vulnerabilidades e indicadores de compromiso"

Actualmente existen herramientas de software que permiten recolectar información de diferentes fuentes con el objetivo de analizarla, correlacionarla y obtener un producto de inteligencia valioso para hacer frente a las ciber amenazas. En este proyecto se propone efectuar y evaluar el proceso de generación de ciberinteligencia en un entorno simulado y controlado, con la presencia de elementos hostiles para la ciberseguridad.

II.2. Splunk como solución para generar ciberinteligencia

Para materializar este proyecto se propone la utilización de Splunk, en este sentido resulta coherente explicar en qué consiste esta herramienta de software.

Splunk es una plataforma de *Big Data* que simplifica la tarea de recopilar y administrar volúmenes masivos de datos generados por máquinas y buscar información dentro de ellos, siendo muy utilizado para análisis empresarial y web, gestión de aplicaciones, cumplimiento y seguridad. Es una forma avanzada y escalable de software que permite capturar, indexar y correlacionar datos provenientes de diversas fuentes en tiempo real. A partir de estos es posible crear alertas, paneles, gráficos, informes y visualizaciones legibles por el personal del departamento o empresa que lo despliega. Las funcionalidades que ofrece ayudan a las organizaciones a reconocer patrones de datos, diagnosticar problemas potenciales, aplicar inteligencia a las operaciones y generar métricas[18].

También en [18] se refieren algunos de los puntos positivos de Splunk:

- Crea informes analíticos a través de cuadros y gráficos interactivos que luego se pueden compartir.
- Un sistema de registro de Splunk es altamente escalable y fácil de implementar para las organizaciones.
- Puede encontrar información útil dentro de los datos de manera automática.
- Guarda búsquedas y etiquetas que reconoce como información importante, lo que ayuda a las organizaciones a hacer que sus sistemas sean más inteligentes.
- Ofrece una interfaz gráfica de usuario (GUI) mejorada y una visibilidad en tiempo real.
- Los resultados instantáneos garantizan que los usuarios dediquen menos tiempo a solucionar problemas.
- Permite a las organizaciones incorporar inteligencia artificial (IA) en sus estrategias de datos y obtener inteligencia operativa a partir de los datos de sus máquinas.
- Puede recopilar cualquier forma de datos, incluidos CSV, JSON y formatos de registro.
- Las organizaciones pueden crear un repositorio central que les permita buscar datos de Splunk de múltiples fuentes.

La Fig.5 muestra un esquema con la idea general del producto ofrecido por Splunk Enterprise.

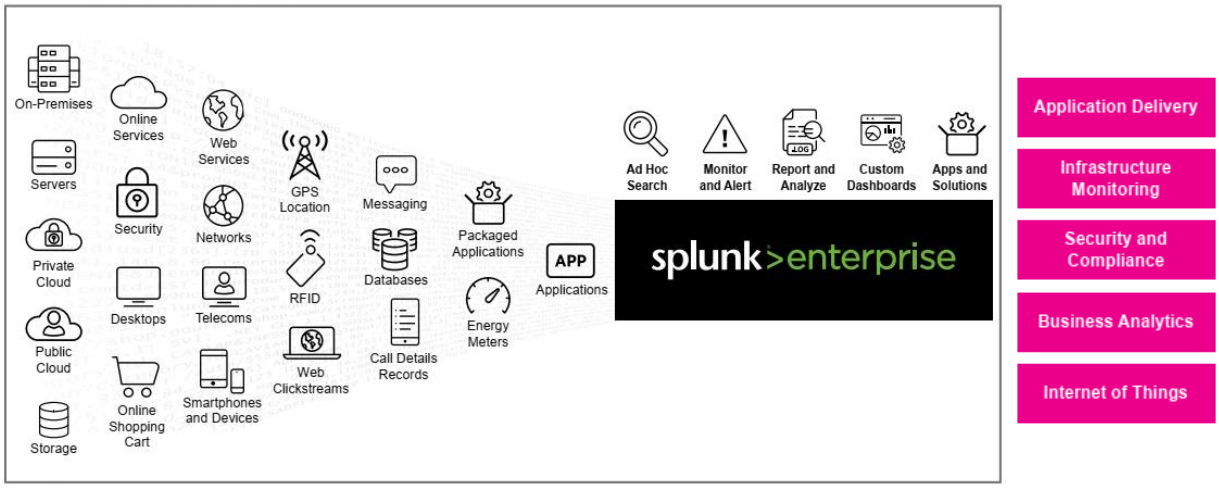

Fig. 5. Idea general de Splunk Enterprise

Esta plataforma brinda una amplia gama de soluciones adaptables y escalables para empresas de variados sectores. Además, permite a los usuarios la creación de sus propias aplicaciones mediante el uso de kits de desarrollo de software (SDK).

En [19] se analiza una de las interrogantes más comunes encontradas en internet alrededor del contexto de esta herramienta en la ciberseguridad: ¿Es Splunk un sistema de Gestión de Información y Eventos de Seguridad (SIEM)? En dicha fuente se afirma que va más allá de un SIEM lidiando con la detección avanzada de amenazas, el monitoreo de seguridad, la administración de incidentes y el análisis forense en tiempo real; permitiendo así una perspectiva de sondeo completo y profundo del entorno y de los peligros existentes en este, y realizando un seguimiento de amenazas específicas en el contexto.

En la Fig.6 se ilustra el Cuadrante Mágico para SIEMs en el año 2020, ofrecido por la consultora Gartner Inc. [20], donde se puede observar que Splunk es uno de los líderes en este sector. En dicho reporte se analizan las tres soluciones específicas de seguridad ofrecidas: Splunk Enterprise Security (ES), Splunk User Behavior Analytics (UBA) y Splunk Phantom. También se abordan las soluciones Splunk Enterprise (*on premise*) y Splunk Cloud, que brindan opciones de recopilación, búsqueda y visualización de datos y eventos para diversos usos en operaciones de TI y seguridad.

En la Fig.7 Gartner Inc. ofrece una comparación, basada en las opiniones de los usuarios, de algunos de los SIEMs con mayor reputación del mercado. En la imagen se puede apreciar la buena posición que ostenta Splunk Enterprise en este ámbito

Fig. 6. Cuadrante mágico de los mejores SIEMs 2020

Fig. 7. Comparativa de Splunk con otros SIEMs en base a opiniones de los usuarios

Con el fin de materializar los objetivos planteados para este proyecto, se podría utilizar la solución Splunk Cloud, pero por motivos de licencia resulta más factible la alternativa de Splunk Enterprise. Esta última ofrece una versión de prueba gratuita por 60 días y cuenta con todas las funcionalidades básicas activas para poder experimentar con casi todo su potencial.

III. Diseño y configuración del entorno para las simulaciones

Para generar ciberinteligencia con Splunk es necesario primero importar información referente a ciber eventos y posteriormente procesarla. Esta puede provenir de *logs* de terminales (ordenadores, servidores, etc.) o de dispositivos de red (*switches*, *routers*, *firewalls*, etc.). En el presente capítulo se propone desarrollar un entorno donde generar y procesar dicha información de manera segura y controlada. Es importante que el escenario sea aislado y que no intervengan otros dispositivos ajenos a las simulaciones, evitando así que su actividad afecte los resultados obtenidos y/o que se vea comprometida su seguridad. Por esta razón, y por motivos de disponibilidad de hardware, se empleará la herramienta de virtualización VMware Workstation Pro, donde se alojarán los siguientes dispositivos:

- Un *firewall* de nueva generación que monitorice la actividad del entorno de simulación y envíe los datos a Splunk.
- Una máquina virtual que genere ciber ataques.
- Una máquina virtual que sea objetivo de los ciber ataques.

Todo este despliegue se realizará en un ordenador personal con un procesador Intel® Core™ i5-10400 @2.4GHz, 16GB DDR4 de memoria RAM, y Windows 10 Education como sistema operativo. Para obtener el máximo rendimiento con este hardware es necesario que la opción de virtualización del microprocesador esté activada (Intel VT-x), configuración que se puede comprobar en el apartado “rendimiento/CPU” del administrador de tareas del sistema. Una vez instalado el VMware se procede a buscar soluciones de software apropiadas y asequibles para recrear el escenario antes descrito. Se comienza por el *firewall* de nueva generación (NGFW).

Numerosas fuentes manifiestan el concepto de NGFW. Las definiciones brindadas por Cloudflare Inc. y Cisco [21][22] son suficientemente precisas y describen algunas sus principales características:

- Capacidades de *firewall* estándar.
- Sistema de prevención (detección) de intrusiones (IPS/IDS).
- Inspección profunda de paquetes (DPI) y de tráfico encriptado.
- Control de aplicaciones.
- Integración de directorios.
- Fuentes de inteligencia de amenazas.
- Técnicas para hacer frente a la evolución de las amenazas de seguridad.

Hoy en día existen diversas soluciones de NGFW, en su mayoría de pago, y ofrecidas por proveedores de altísimo prestigio a nivel internacional (Cisco, Palo Alto Networks, Fortinet, etc.) y que son utilizadas por empresas de diferentes sectores para garantizar su ciberseguridad. Debido la índole investigativa de este trabajo se decidió enfocar esta búsqueda en una alternativa gratuita. Además, al no contarse con una plataforma de hardware individual, se hace necesario que el *firewall* pueda ser virtualizado y requiera de pocos recursos de cómputo para su funcionamiento. Consultando algunas fuentes en internet [23][24][25] se hallan varias soluciones gratuitas de NGFW, con o sin limitantes de funcionalidades y tiempo de uso. Sin duda alguna las mejores alternativas encontradas son las de Pfsense Community Edition y Sophos XG Firewall Home Edition. Dado que esta última requiere de mayores capacidades de cómputo se elige Pfsense, el cual además es mejor valorado por la comunidad de usuarios y, como según refieren los autores en [26]: *"...Pfsense es el firewall de código abierto más completo disponible en el mercado."*

Accediendo al sitio oficial se procede a la descarga del archivo adecuado (Fig.8).

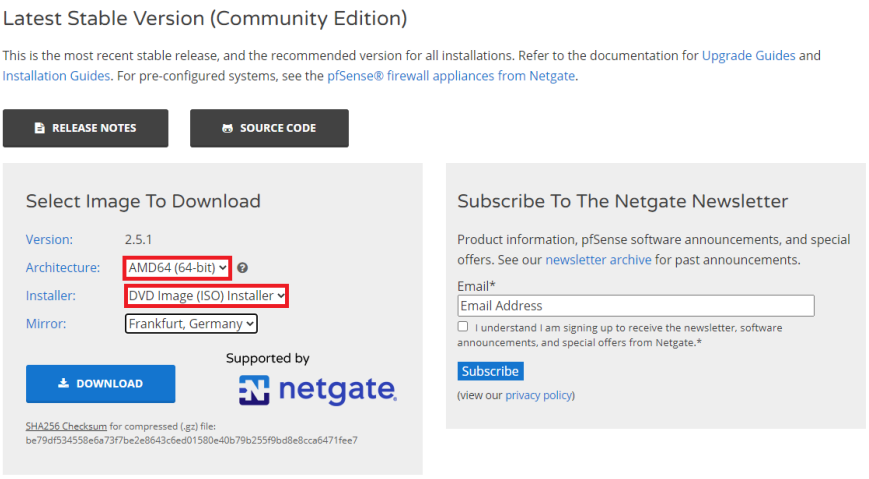

Fig. 8. Descarga del firewall Pfsense

Posteriormente se crea una máquina virtual (VM) de Pfsense con tres adaptadores de red (Fig.9), dos para las redes locales internas y uno para la salida a internet (WAN).

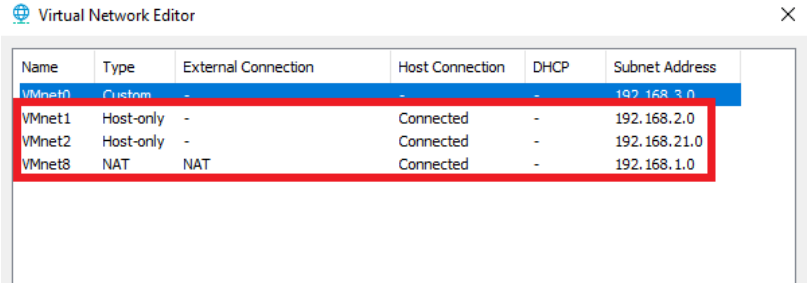

Fig. 9. Configuración de los adaptadores de red del *firewall* Pfsense

El adaptador VMnet8, que se configura con *Network Address Translation* (NAT), será el dedicado a la interfaz WAN del NGFW. Los adaptadores VMnet1 y VMnet2 son configurados como *Host-Only* con sus respectivas subredes, luego el *firewall* se encargará de enrutar sus conexiones hacia internet mediante la interfaz WAN. Una vez instalada la VM, se inicia y se accede al menú de consola, donde se configuran las interfaces (asignación de direcciones físicas y de IP):

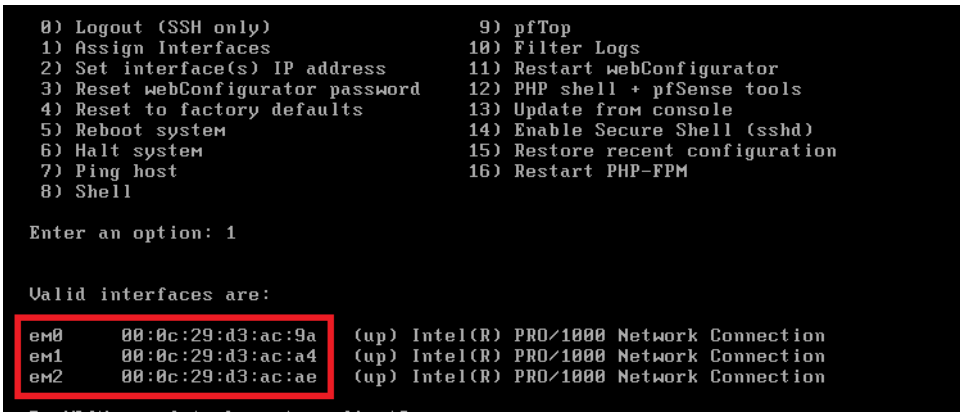

Fig. 10. Asignación de interfaces físicas del *firewall*

La asignación de direcciones físicas (MAC) debe coincidir con las de cada adaptador de red virtual configurado a la VM antes creada (Fig.10 y Fig.11).

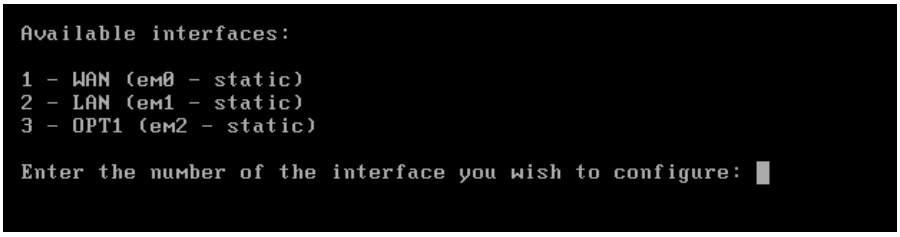

Fig. 11. Asignación de las direcciones IP a cada interfaz del *firewall*

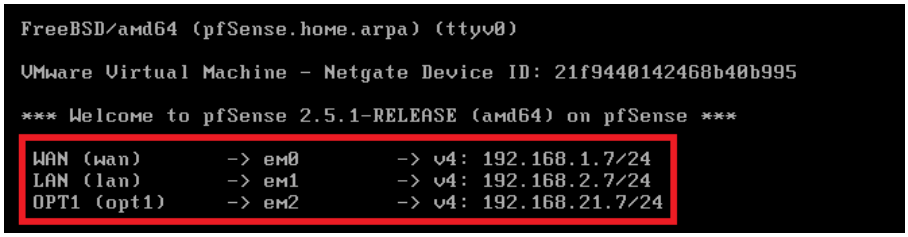

Fig. 12. *Firewall* con sus interfaces configuradas

Una vez hecho esto ya está disponible el acceso al configurador web del *firewall* desde las direcciones IP asignadas a cada interfaz. Por motivos de seguridad, Pfsense tiene restringido el acceso al configurador web desde la WAN, lo cual es una buena práctica y bastante recomendable, puesto que de lo contrario el sistema estaría expuesto a un acceso no autorizado desde el exterior

que podría resultar muy peligroso. Existen diversas alternativas para garantizar el acceso remoto al configurador web, sin embargo este asunto se desvincula del propósito de esta investigación, por lo cual se administrará el *firewall* desde la LAN a través de la dirección IP 192.168.2.7/24. Con las credenciales por defecto se inicia sesión y se configura lo necesario para el correcto funcionamiento del entorno.

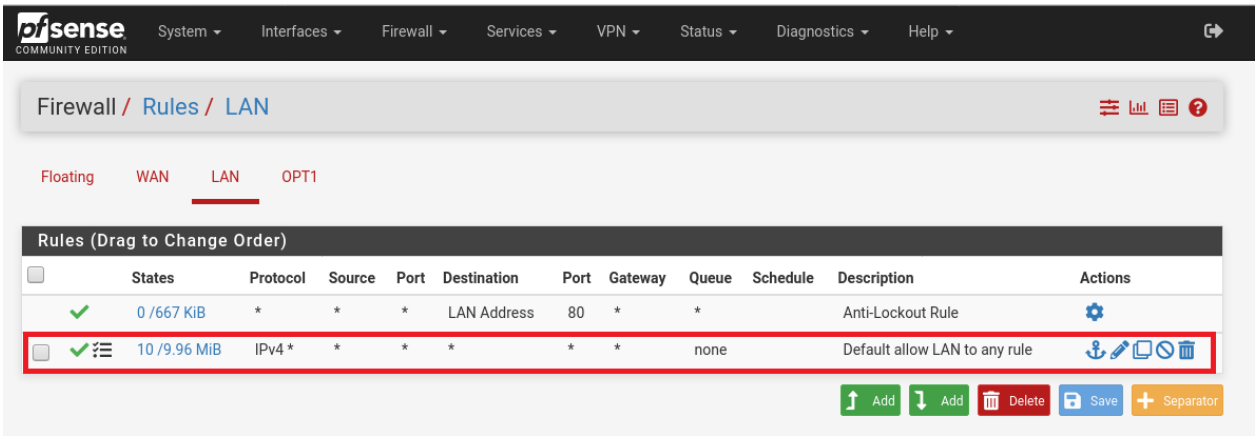

Fig. 13. Reglas para la interfaz LAN

En la Fig.13 se muestra la regla que es agregada en la interfaz LAN para autorizar el tráfico cursado a través de esta. También se habilita el registro de todos los paquetes manejados (Fig.14), con el objetivo de enviar posteriormente estos *logs* a Splunk para su análisis. A la interfaz OPT1 se le realiza la misma configuración. Nótese además la presencia de una regla por defecto en la interfaz LAN (Fig.13), la regla (*Anti-lockout*), la cual está orientada a impedir que se bloquee el acceso del administrador al *firewall* desde esta red.

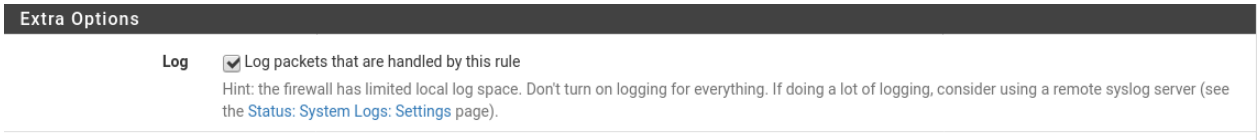

Fig. 14. Registro de paquetes manejados por la regla añadida a la interfaz LAN

Pfsense además brinda la opción de que se le instalen aplicaciones que amplíen sus funcionalidades, tal es el caso del Sistema de Detección/Prevención de Intrusiones (IDS/IPS) *open source* Suricata. Su inclusión en este entorno dará un valor añadido al proceso de monitorización e identificación de la actividad maliciosa. Una vez instalado Suricata se procede a su configuración y se eligen las fuentes de las reglas para la detección de amenazas: ETOpen y Snort Community Ruleset GPLv2. Luego se inicia el servicio para las interfaces LAN y OPT1, activando en ambas la opción de envío de alertas al registro del *firewall* e incluyendo información

adicional en formato EVE (Extensible Event Format). El resto de las configuraciones que vienen por defecto se conservan.

Pfsense trae incorporado un servicio de *syslog* sobre UDP para el envío de *logs* a un servidor externo, el cual se utilizará para transmitir hacia Splunk. Mediante este estándar la información se envía como texto plano, opción no recomendable para un sistema real, puesto que los datos podrían ser interceptados y utilizados por un agente malicioso. Por esta razón, es aconsejable la encriptación de los mensajes, pero dado el enfoque de este proyecto y a que el entorno se despliega en una red local y segura, no se tendrán en cuenta esos detalles. Para exportar los datos a Splunk es suficiente especificar la IP del ordenador donde este se aloja (Fig.15). En este caso se empleó una de las direcciones IP asignadas al ordenador físico mediante los adaptadores virtuales configurados en *Host-only*, que como se explica en [27] dicha configuración crea además una red local virtual entre el *host* (PC física) y la VM.

Fig. 15. Configuración del servidor remoto de *syslog* para Pfsense

Una vez configurado el *firewall* se procede a la instalación de las VMs restantes. Para realizar los ciber ataques se utilizará Kali Linux, una distribución de Linux de código abierto basada en Debian y orientada a diversas tareas de seguridad de la información, como pruebas de penetración (*penetration testing*), investigación de seguridad, informática forense e ingeniería inversa [28]. En este caso está disponible una VM para VMware en el sitio oficial de Offensive-Security, por lo que solo es necesario descargarla e importarla. Varias de las aplicaciones preinstaladas en Kali resultan de gran utilidad para este trabajo, en especial Metasploit Framework (Fig.16). La misma es plataforma modular que contiene una colección de herramientas que proporcionan un entorno completo para pruebas de penetración y desarrollo de *exploits* [29].

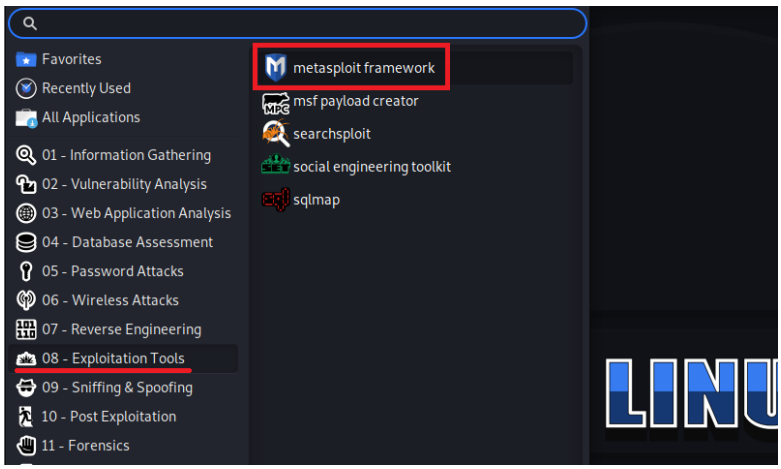

Fig. 16. Herramientas disponibles en Kali Linux: Metasploit Framework

Para facilitar y hacer más interesante el trabajo de la detección de ciber ataques resulta idóneo contar con un objetivo con múltiples vulnerabilidades, al cual se le puedan aplicar diversas técnicas para explotarlas. Por ese motivo para el elemento restante del entorno se elige Metasploitable 2 [30], una versión de Ubuntu intencionalmente vulnerable diseñada para pruebas de penetración y ofrecida por Rapid7 como una VM. Luego se accede al enlace para descargarla y se importa hacia la plataforma de virtualización, ubicándola en una subred diferente a la VM de Kali.

En este punto solo falta configurar la entrada de datos a la instancia de Splunk Enterprise. Existen diversas maneras de hacerlo, ya sea de manualmente cargando un archivo con la información a indexar, de forma automática mediante el uso Splunk Universal Forwarder o a través de *syslog* (TCP o UDP). La última opción mencionada es la más conveniente para el entorno desplegado, adaptándose perfectamente a las características de este.

Para configurar esta entrada de datos, se accede a la plataforma como administrador, se ingresa al menú "*Settings*", luego a *"Data Inputs"*, posteriormente se selecciona en el apartado *"Local inputs"* la opción UDP y se añade una entrada nueva. Una vez dentro, se abre un menú guiado donde se configura el puerto de escucha: 514/UDP; y el indexado: *sourcetype = pfsense*, *index = tfm_security_logs*.

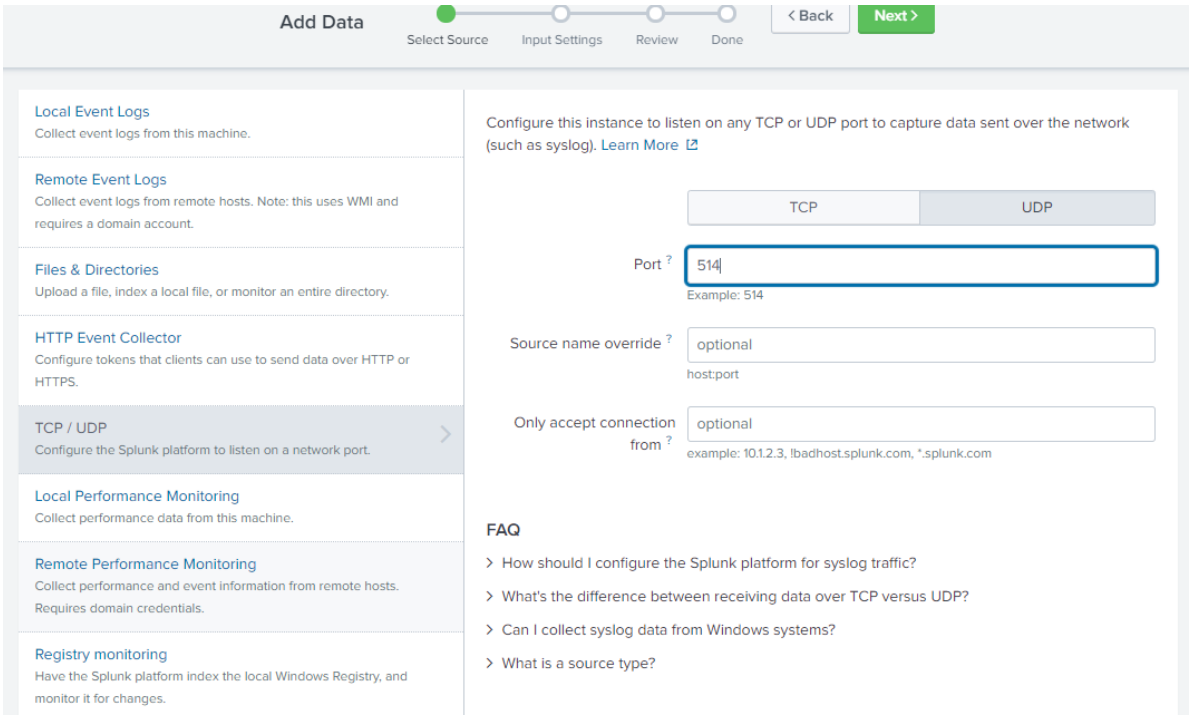

Fig. 17. Configuración de entrada de datos vía *syslog* a Splunk

Terminadas las configuraciones pertinentes, el entorno previamente propuesto quedaría representado como ilustra la Fig.18.

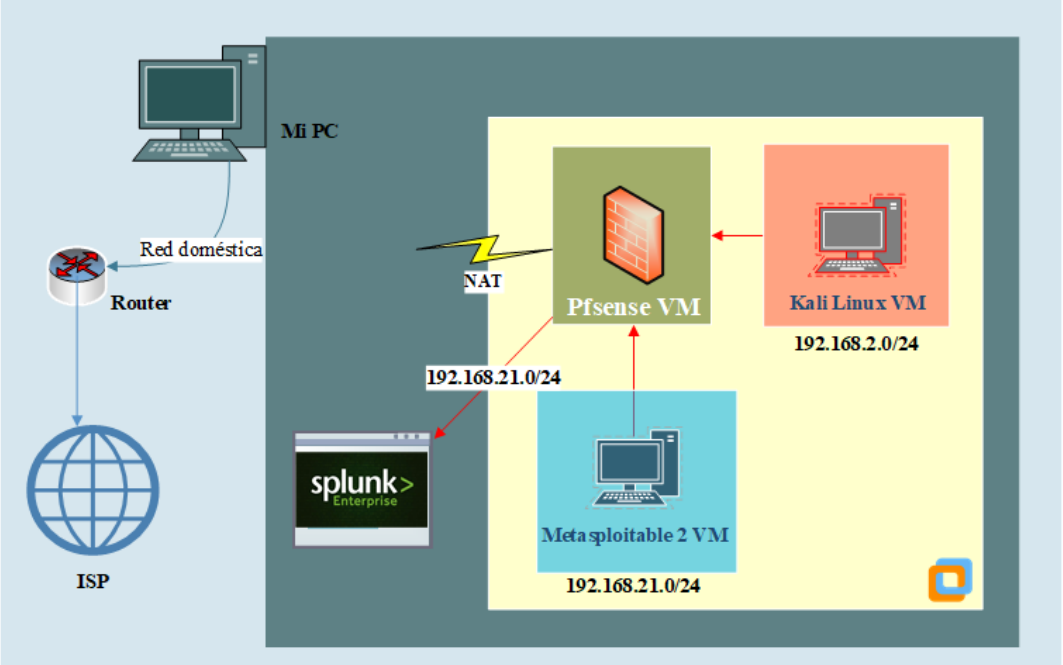

Fig. 18. Entorno recreado para las pruebas

IV. Puesta en marcha del entorno y generación de ciberinteligencia

IV.1. Descripción de los ciber ataques controlados

En este apartado se explican brevemente los ciber ataques y las técnicas para efectuarlos en el entorno antes descrito. Primeramente, es necesario realizar una recopilación de información sobre la víctima, detalles como la dirección IP, los puertos abiertos, los servicios presentes y las vulnerabilidades existentes. Esto es posible mediante el empleo de la herramienta *nmap* (Fig.19), la cual ofrece numerosas opciones para el sondeo de la red, incluso técnicas de furtividad (no empleadas en este caso) para evitar *firewalls* e IDS/IPS.

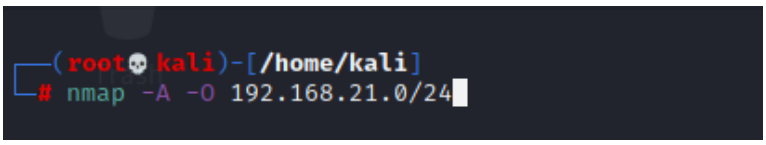

Fig. 19. Sondeo de la red empleando la herramienta *nmap* desde la consola de comandos

De antemano es conocido que el objetivo está en la subred 192.168.21.0/24. La opción *-O* habilita la detección del sistema operativo, mientras que la *-A* además permite la detección de versiones, el escaneo de scripts y el rastreo de ruta. La Fig.20 ilustra parte del resultado de la ejecución del comando anterior.

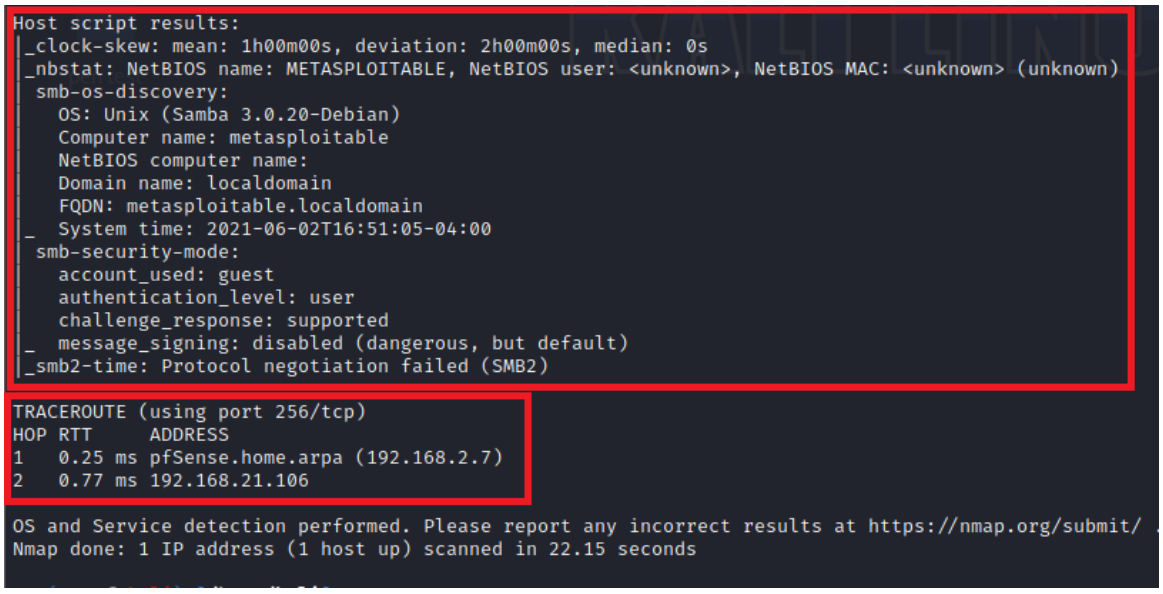

Fig. 20. Fragmento de los resultados de la ejecución del comando *nmap*

Es posible observar que en la IP 192.168.21.106 se encuentra alojado el *host* al que se pretende explotar. Ahora se procede a emplear una de las herramientas más versátiles que ofrece Kali, Metasploit Framework (Msf), la cual es ofrecida por Rapid7 al igual que Metasploitable2, siendo

ambas soluciones parte de un conjunto para pruebas de penetración. Desde Msf el primer paso será realizar un escaneo más profundo del *host* de interés, y esta vez con un análisis de vulnerabilidades incluido, la Fig.21 aporta los detalles:

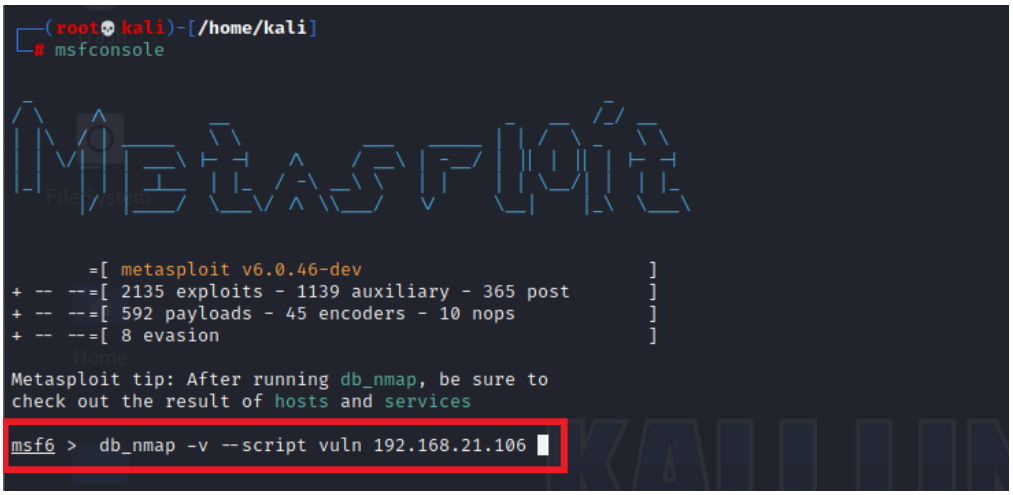

Fig. 21. Búsqueda de vulnerabilidades en el *host* seleccionado

El comando previamente ilustrado retorna una extensa lista de todos los servicios en activo por cada puerto y las vulnerabilidades detectadas, junto con posibles *exploits* para sacarles provecho, y en algunos casos enlaces con documentación al respecto. En base a la información obtenida aquí, se deciden realizar los siguientes ataques:

- Ataque de denegación de servicio por inundación de paquetes SYN (*DOS synflood*).
- Aprovechamiento de la vulnerabilidad de puerta trasera (*Backdoor*) presente en el servidor vsftpd 2.3.4 (CVE-2011-2523).
- Inyección de argumentos mediante PHP_cgi (CVE-2012-1823).
- Ataque de fuerza bruta contra los servicios SSH y MySQL.

IV.1.1. Denegación de servicio por inundación de paquetes SYN (DOS Synflood)

El ataque de inundación de paquetes SYN funciona aprovechando el proceso de protocolo de enlace de una conexión TCP [31], constituido por tres pasos conocidos como *Three-way hanshake*:

1. El cliente envía un paquete SYN al servidor para iniciar la conexión.
2. El servidor responde al paquete SYN inicial con un SYN/ACK reconociendo la comunicación.

3. El cliente devuelve un paquete ACK para confirmar la recepción del paquete del servidor. Tras finalizar esta secuencia la conexión TCP se establece y es posible enviar y recibir información.

En la Fig.22 [31] se ilustra la denegación de servicio mediante este proceso, lo cual se explica a continuación:

1. El cliente (atacante) envía una gran cantidad de paquetes SYN al servidor (objetivo), incluso empleando una dirección IP falsificada.
2. El servidor responde a cada una de las solicitudes de conexión y mantiene abierto un puerto listo para recibir la réplica.
3. Mientras el servidor espera el último paquete ACK, el cual nunca llega, el atacante persiste en el envío paquetes SYN. El arribo de cada nuevo SYN origina que el servidor conserve temporalmente una nueva conexión de puerto abierto. Una vez utilizados todos los puertos disponibles el servidor ya no puede operar con normalidad.

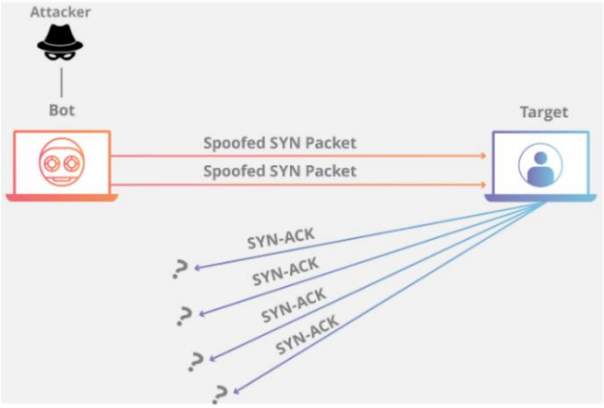

Fig. 22. Funcionamiento del ataque *DOS synflood*

Este ataque de denegación de servicio (DOS) se efectuará a través de las facilidades que ofrece Metasploit, configurándose los parámetros necesarios como muestra la Fig.23 y posteriormente ejecutando el comando *exploit*.

```
msf6 auxiliary(dos/tcp/synflood) > show options

Module options (auxiliary/dos/tcp/synflood):

   Name       Current Setting  Required  Description
   ----       ---------------  --------  -----------
   INTERFACE                   no        The name of the interface
   NUM        0                no        Number of SYNs to send (else unlimited)
   RHOSTS     192.168.21.106   yes       The target host(s), range CIDR identifier, or hosts file with syntax 'file:<path>'
   RPORT      80               yes       The target port
   SHOST                       no        The spoofable source address (else randomizes)
   SNAPLEN    65535            yes       The number of bytes to capture
   SPORT                       no        The source port (else randomizes)
   TIMEOUT    500              yes       The number of seconds to wait for new data

msf6 auxiliary(dos/tcp/synflood) >
```

Fig. 23 Configuración para realizar ataque *DOS synflood*

IV.1.2. Backdoor vsftpd 2.3.4 (CVE-2011-2523)

Este ataque se ejecuta explotando una vulnerabilidad que fue introducida en el código fuente del servidor vsftpd 2.3.4. en el año 2011. Se realiza intentando acceder al puerto 21 del *host* que aloja el vsftpd, y usando la cadena ":)" en el nombre de usuario que solicita. De manera automática se ejecuta un intérprete de comandos (*shell*) a través del puerto 6200/TCP por el cual acceder a los directorios de la víctima. La explotación de esta vulnerabilidad representa una situación de impacto crítico en la seguridad del *host* comprometido. La Fig.24 ilustra la configuración previa para explotar esta vulnerabilidad con Metasploit:

Fig. 24. Configuración para vulnerar el servicio vsftp2.3.4

IV.1.3. Inyección de argumentos mediante PHP_cgi (CVE-2012-1823)

Este ataque utiliza una vulnerabilidad presente en las versiones de PHP anteriores a la 5.3.12 y 5.4.2, que consiste en un manejo inadecuado de las consultas, permitiendo la inyección de argumentos con contenido malicioso que proporcione acceso al código fuente del sitio web, o incluso establecer una conexión de *shell* inverso que brinda al atacante acceso a los directorios del servidor. Es deducible que el éxito de esta ofensiva compromete severamente la información contenida en el objetivo. La configuración de Metasploit para ejecutarlo se ilustra en Fig.25.

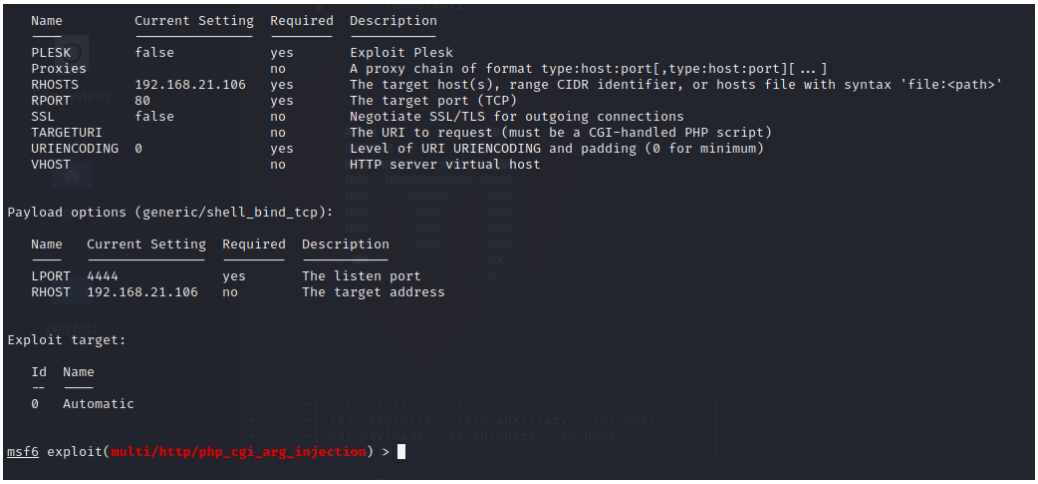

Fig. 25. Configuración del Msf para efectuar el ataque de inyección de argumentos mediante *PHP_cgi*

IV.1.4. Ataque de fuerza bruta

Se basa en la técnica exhaustiva para descifrar las credenciales de acceso a un servicio o aplicación determinada, en el cual se suelen utilizar listas con posibles candidatos para dichas credenciales. Para esto el atacante realiza varios intentos con combinaciones de estos candidatos hasta encontrar los correctos. El éxito de este ataque depende en mayor medida de la robustez de la contraseña empleada para regular el acceso al servicio; por otro lado, y con menor efecto, dependerá de lo preciso que sea el diccionario de candidatos empleado, jugando un papel importante el conocimiento previo de la mayor información posible acerca de la víctima.

Con la intención de efectuar este ciber ataque se empleará la herramienta xHydra ofrecida por Kali, específicamente fijando como objetivo el servicio SSH alojado en el puerto 22/TCP y el servicio de base de datos MySQL ubicado en el puerto 3306/TCP. Se utilizará además una lista en texto plano con algunas posibles contraseñas. La Fig.26 muestra la configuración para ambos servicios.

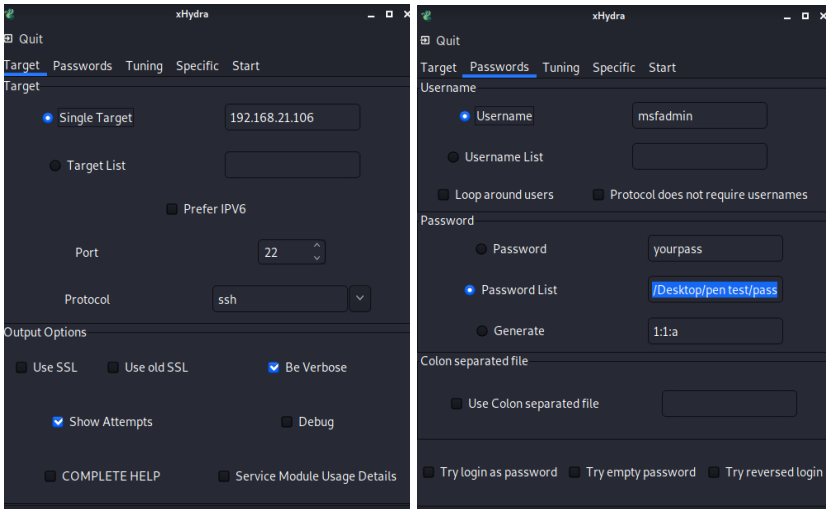

Fig. 26. Configuración de xHydra para efectuar ataques de fuerza bruta

IV.2. Importación de datos a Splunk

En el capítulo anterior, se realizó la configuración de la entrada de datos a Splunk Enterprise, definiéndose como fuente de *logs* el puerto 514/UDP, como *sourcetype = "pfsense"* e *index = "tfm_security_logs"*. Estos parámetros son de gran importancia para acceder a los eventos procedentes del *firewall*. Splunk Enterprise funciona como una aplicación web base sobre la que se ejecutan otras aplicaciones, basando todas sus operaciones en búsquedas de datos previamente indexados. Para esto se utiliza el *Search Processing Language (SPL)* en la aplicación que *Search and Reporting*, desde donde se procederá a buscar los elementos de interés para generar la ciberinteligencia.

La Fig.27 muestra la interfaz de dicha aplicación con las siguientes zonas de interés numeradas y marcadas en rojo:

1. *SPL Editor*, donde se introducen las búsquedas en *SPL*.
2. Selector de rango de tiempo, el cual permite que las búsquedas se ejecuten en tiempo real o en diferentes y flexibles intervalos de tiempo.
3. Línea de tiempo, donde se muestra en forma de histogramas el número de eventos en función del tiempo y con la que es posible interactuar para refinar la búsqueda.
4. Campos identificados por Splunk al indexar los datos.
5. Eventos indexados, pero sin ser tratados aún. En la literatura suelen ser llamados *raw data* (datos en bruto).

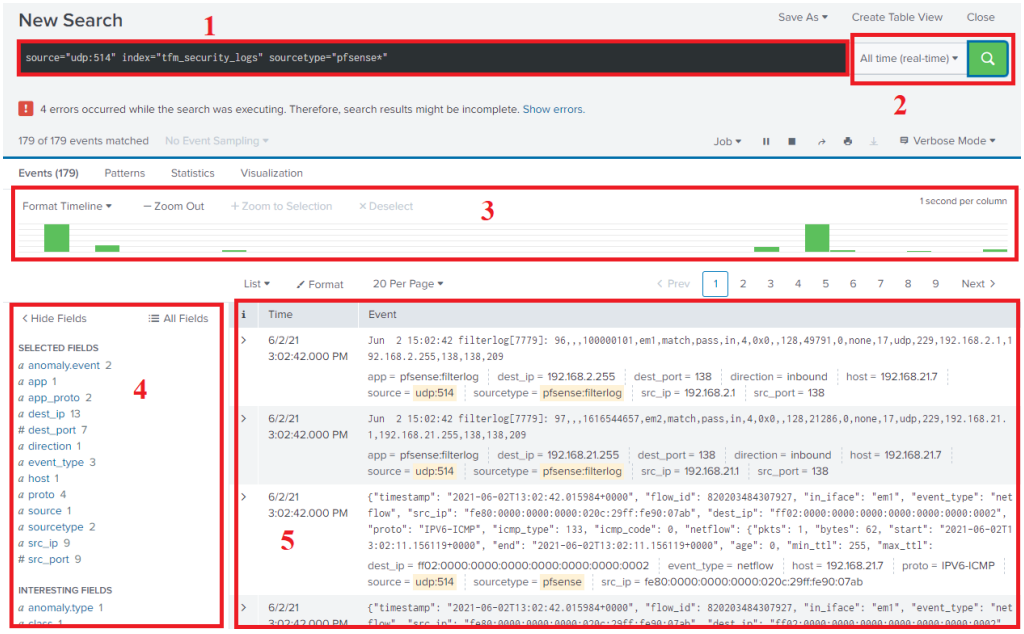

Fig. 27. Interfaz de la aplicación *Search and Reporting*

Con los diversos comandos y funciones de *SPL* combinados adecuadamente se pueden manipular estos datos en bruto para obtener una información legible y detallada.

IV.3. Procesamiento de los datos y generación de ciberinteligencia

Al efectuar los ciber ataques controlados previamente descritos, se procede a generar ciberinteligencia. Filtrando mediante búsquedas *SPL* los eventos por campos de interés como: *alert.signature*, *src_ip*, *dest_ip*, *src_port*, *dest_port*; se obtienen tablas con información muy útil y legible. La Fig.28 ilustra una búsqueda relacionada con los eventos generados por el ataque *DOS synflood*.

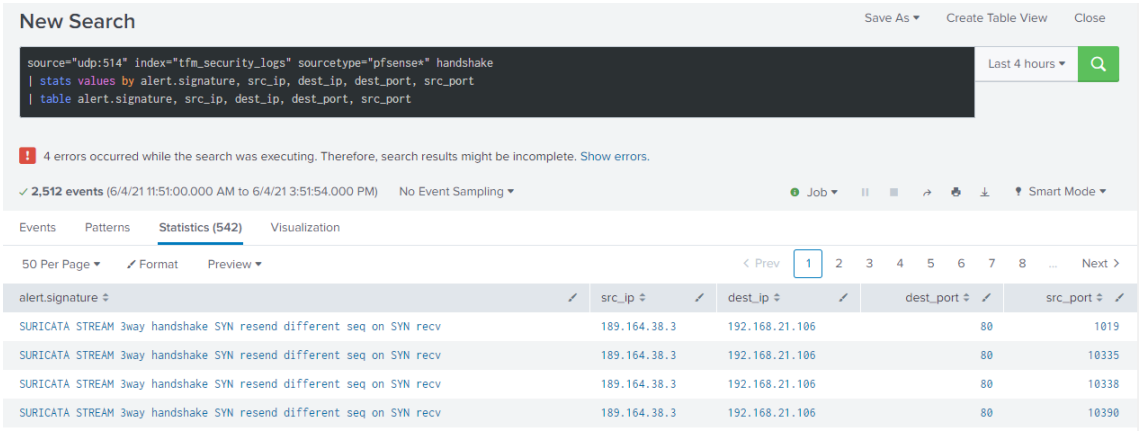

Fig. 28. Estadísticas sobre los eventos relacionados con el ataque *DOS synflood*

En la imagen muestra el uso del comando *stats*, que emplea funciones estadísticas para la manipulación de los datos. Seguidamente se utiliza la función *values* la cual devuelve los distintos valores por los que se invoca, condicionada por la cláusula *by* para separarlos por los campos enumerados a continuación: firma de alerta, IP de origen y destino, puerto fuente y destino (*alert.signature*, *src_ip*, *dest_ip*, *src_port*, *dest_port*). De manera similar se encuentran los incidentes relacionados con los otros ataques, esta vez empleando la función *count*, que devuelve la cuenta de los eventos por los campos especificados según la cláusula *by* (Fig.29).

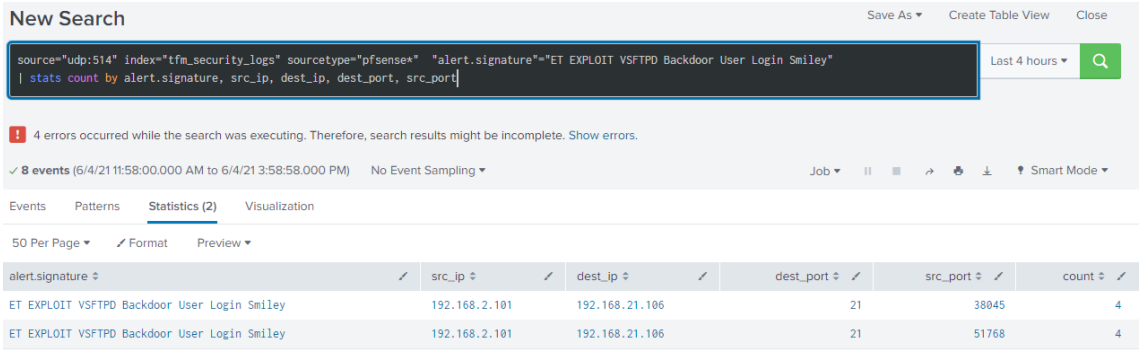

Fig. 29. Estadísticas sobre los eventos relacionados con el ataque al vsftpd

Los resultados arrojados por estas búsquedas y otras similares pueden ser guardados dentro de la plataforma como reportes, *dashboards* o alertas, aportando un valor añadido a la capacidad de análisis y a la Ciber Conciencia Situacional para enfrentar a las ciber amenazas. En consecuencia, se crea un *dashboard* titulado *INTERNAL SOC for TFM* (Fig.30), donde se exhiben los datos indexados y procesados durante un rango de tiempo determinado, siendo posible también su actualización en tiempo real. El mismo cuenta con los siguientes sub-paneles, creados específicamente para este entorno:

1. Número de alertas por ciberataques.
2. *Top* de alertas recientes lanzadas por Suricata, con información de direcciones IP de origen y destino, así como los puertos del servicio atacado.

3. Direcciones IP de los *hosts* víctimas de ataques DOS.
4. Las direcciones IP de los *hosts* más peligrosos en el entorno según la cantidad de alertas generadas.
5. Relación entre las direcciones IP involucradas en cada alerta disparada en el *firewall*.
6. Servicios afectados por los ciber ataques en el intervalo de tiempo seleccionado.

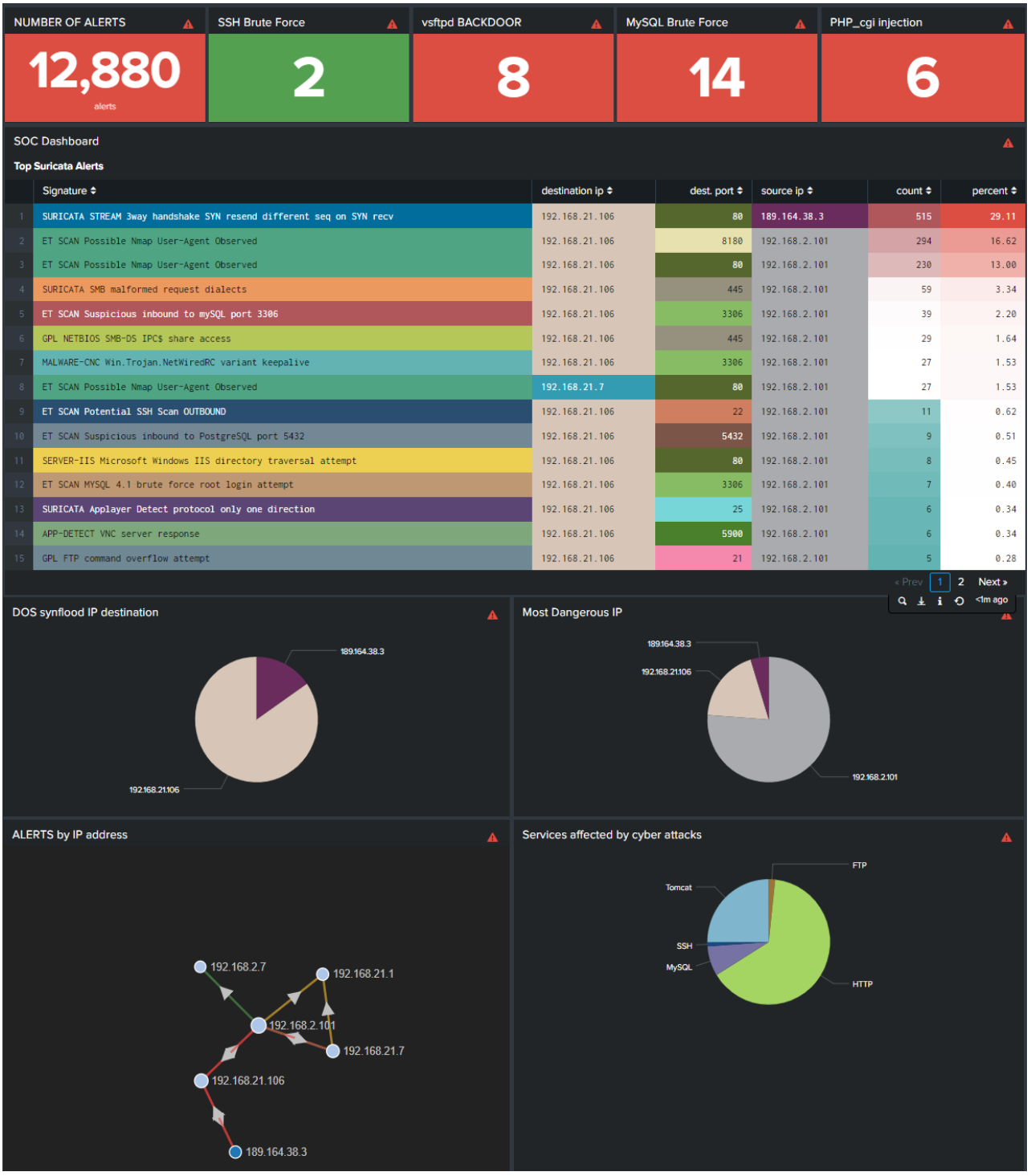

Fig. 30. *Dashboard INTERNAL SOC for TFM*

Se puede observar como las direcciones IP origen y destino de la mayoría las alertas corresponden con la de los *hosts* involucrados en el entorno simulado: la VM Kali (192.168.2.101) y la VM Metasploitable2 (192.168.21.106) respectivamente. Vale la pena señalar que en el caso de los ataques DOS el agresor emplea una falsa “IP de origen” para confundir al servidor.

Es coherente afirmar que este *dashboard* brinda una visión general y suficientemente detallada de la actividad maliciosa que sucede en la red, permitiendo arribar a conclusiones más precisas y tomar decisiones de manera más rápida en el tratamiento de ciber ataques. De manera similar, es posible crear otros *dashboards* que ofrezcan información relevante, como el expuesto en la Fig.31, que ilustra el comportamiento del tráfico (en bytes) durante un período de tiempo en las previamente subredes analizadas.

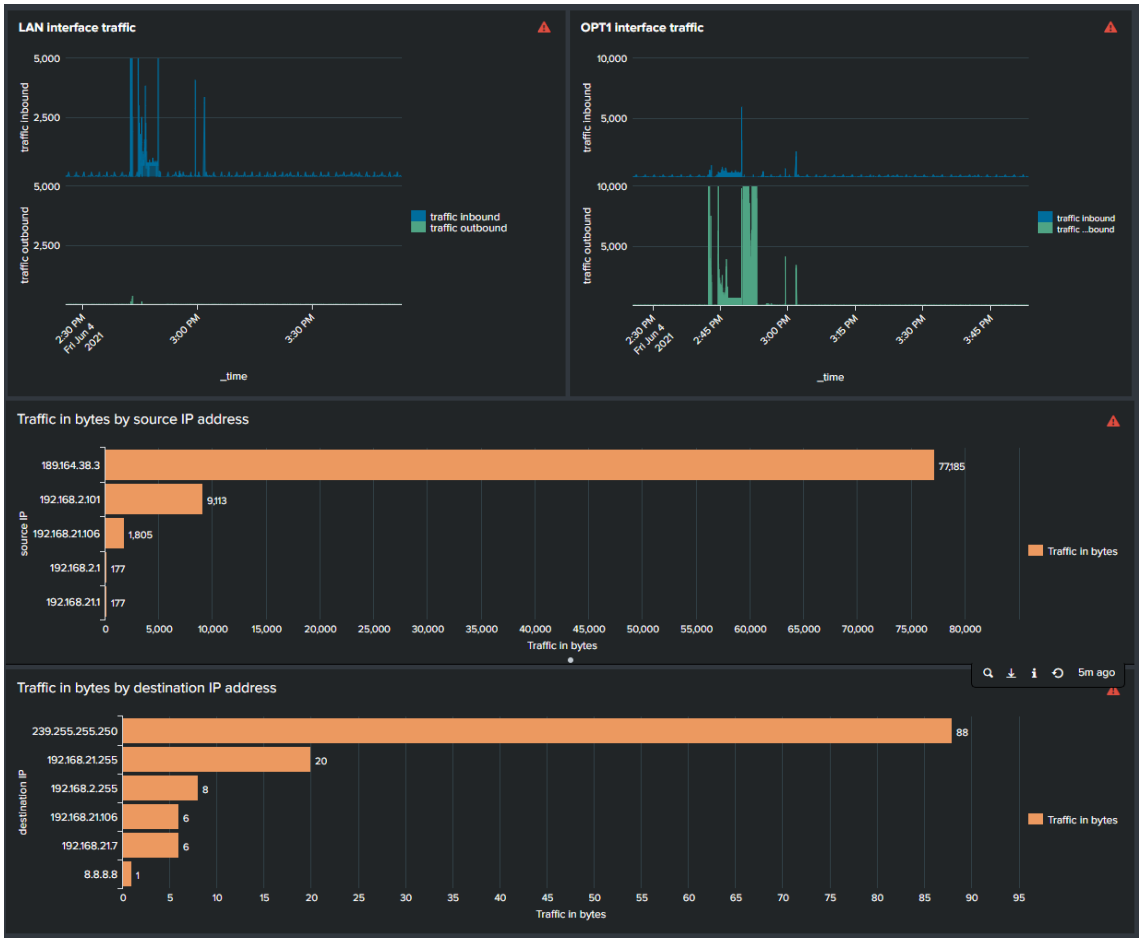

Fig. 31. *Dashboard* del comportamiento del tráfico de las subredes analizadas

La información del tráfico en la red es de vital importancia para el equipo de ciberseguridad de una institución. El análisis del comportamiento de este parámetro por rangos de tiempo, por direcciones IP, e incluso por servicio puede ser un factor clave para detectar un ciber ataque y responder ante el mismo.

Además de los *dashboards*, es posible crear alertas para búsquedas personalizadas y sujetas a condiciones específicas. La Fig.32 muestra la configuración para la alerta *"SSH BRUTE FORCE ATTACK"*.

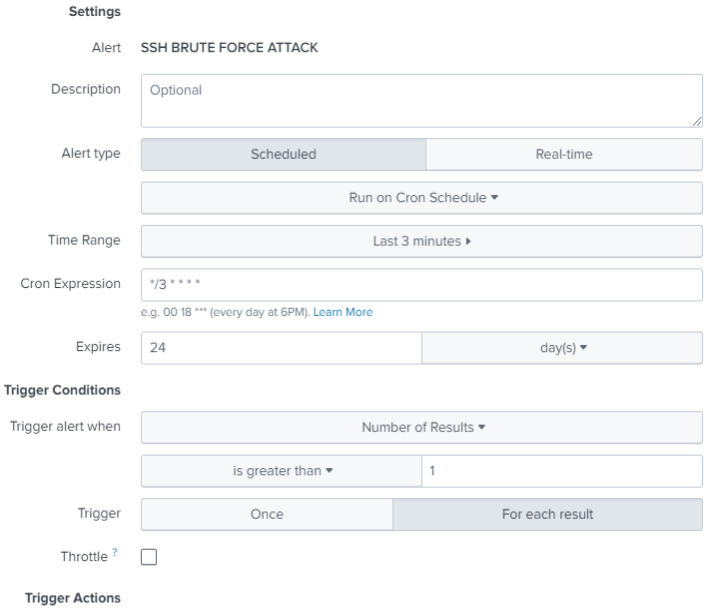

Fig. 32. Configuración para la alerta "*SSH BRUTE FORCE ATTACK"*

Se puede apreciar que las opciones que condicionan la emisión de la alerta son considerablemente flexibles. Si esta se configura para ejecutarse en tiempo real, se disparará al instante en que la búsqueda con la que está relacionada devuelva algún valor. Si se define como programada, se efectuará la búsqueda según el intervalo de tiempo definido y se disparará la alerta si en ese intervalo se devuelve algún valor. También es posible regular la activación de la alerta bajo otras condiciones como: número de resultados, número de *hosts* involucrados, o incluso mediante una búsqueda *SPL* personalizada; de forma que si en el período de tiempo programado cumple las condiciones fijadas entonces se active.

Splunk Enterprise ofrece un panel interno donde se pueden consultar las alertas lanzadas recientemente. Es posible, además, efectuar una búsqueda sobre los eventos relacionados con estas y posteriormente guardarlos en un reporte en el que se muestre la información deseada. La Fig.33 ilustra un ejemplo de esto:

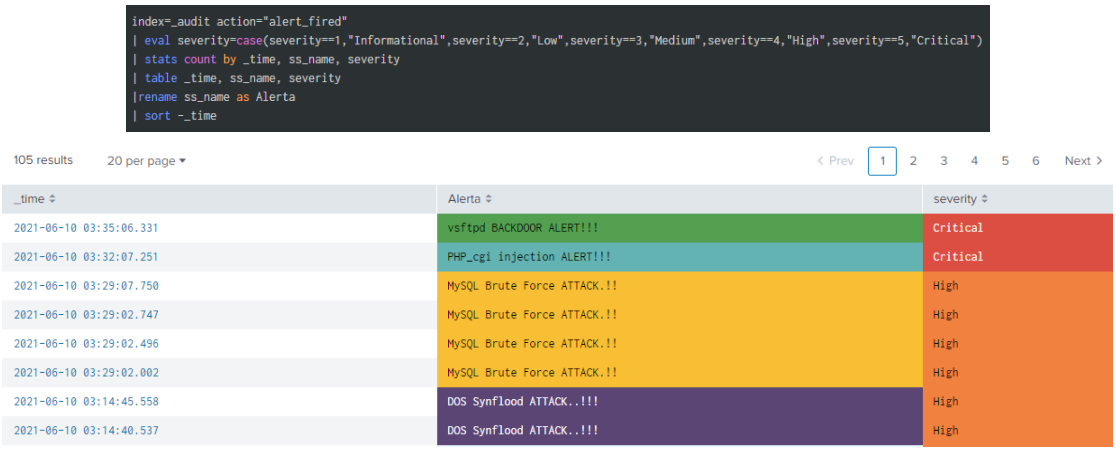

Fig. 33. Reporte sobre las alertas disparadas en un período de tiempo determinado

IV.4. Exportación de la ciberinteligencia desde Splunk

Hasta este punto se ha conseguido procesar los datos en bruto procedentes del entorno simulado para la creación de ciberinteligencia, pero resultaría de gran utilidad el poder exportar esta información al exterior de la plataforma. A continuación, se describen algunas de las opciones que ofrece Splunk Enterprise para dicha actividad.

Los resultados obtenidos mediante una búsqueda *SPL* se pueden exportar manualmente en un fichero (*.cvs*, *.xml*, *.json*) desde la misma interfaz de búsqueda (Fig.34). Esta opción quizás parezca la manera más sencilla, pero puede resultar de gran utilidad para un posterior análisis o envío de la información hacia otra plataforma.

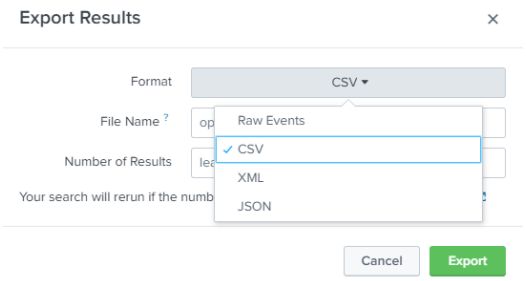

Fig. 34. Exportación en un fichero de los resultados de una búsqueda *SPL*

Una alternativa similar a la anterior es la que ofrece el comando *dump*, con la diferencia de que puede ser incluido en una búsqueda automática (alerta, reporte, etc.). Por esta vía se genera un archivo que contiene los resultados devueltos y se almacena en un directorio del servidor donde se encuentra instalado la instancia de Splunk.

Como ha sido explicado previamente, la creación de alertas a partir de búsquedas específicas es un método muy interesante para generar ciberinteligencia. Son diversas las acciones brindadas por la plataforma que permiten responder a las alertas, incluso se pueden agregar otras instalando la aplicación correspondiente. En especial vale la pena resaltar la acción de envío de datos vía *webhook*. Básicamente un *webhook* es una retrollamada HTTP, una notificación mediante la cual se transmite información en el cuerpo de la solicitud POST, posibilitando su empleo para eventos en tiempo real como las alertas. Para su implementación se hace uso del sitio de pruebas https://webhook.site/, el cual genera una URL única para la recepción de las notificaciones. La configuración para la alerta *"MySQL Brute Force ATTACK..!!"* se ilustra en la Fig.35.

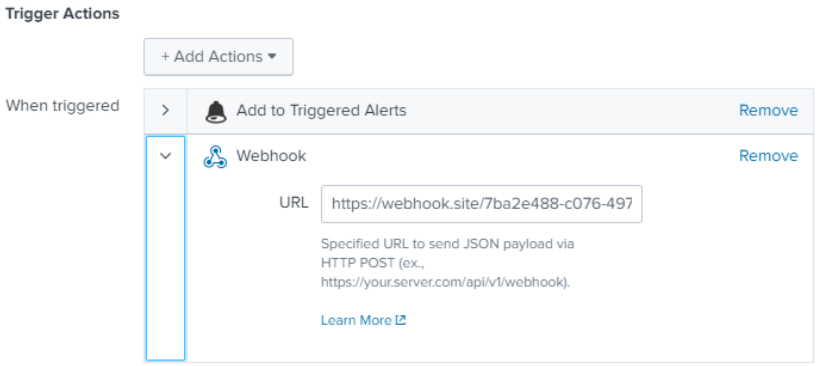

Fig. 35. Configuración de la acción de notificación mediante *webhook* de una alerta

Una vez disparada la alerta se envía la notificación en formato JSON, a la URL habilidata para este *webhook*, conteniendo la información sobre los parámetros devueltos por la búsqueda asociada (Fig.36).

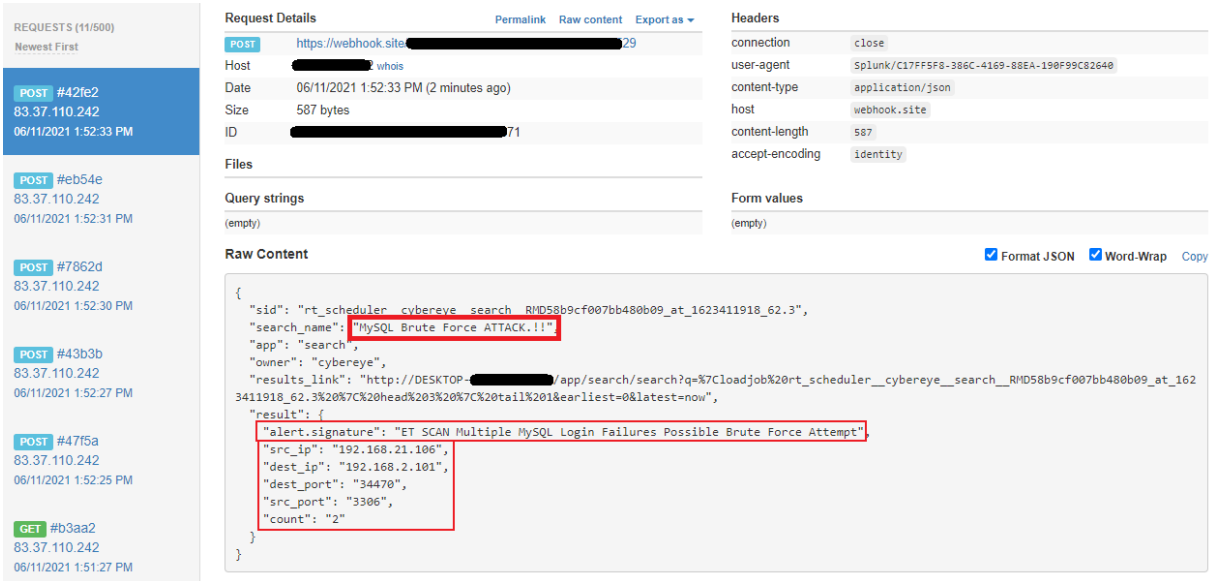

Fig. 36. Notificación recibida mediante *webhook*

Otra alternativa para exportar la ciberinteligencia es mediante e-mails, empleando el comando *sendemail* en búsquedas específicas, *dashboards*, alertas y/o reportes. Para ello es necesario configurarle a la instancia de Splunk Enterprise los parámetros necesarios para el envío de correos. En este caso se empleará un servidor SMTP externo proporcionado por Yahoo (*smtp.mail.yahoo.com:465*). En la Fig.37 se puede observar el comando *sendemail* usado en la búsqueda relacionada con la alerta creada para los ciber ataques de vsftp *Backdoor*. Por esta vía se puede notificar a uno o varios destinatarios, incluyéndole los resultados de la búsqueda tanto en el cuerpo del mensaje como en un archivo adjunto en formato *.csv* o *.pdf*. Además, el asunto y el cuerpo del mensaje pueden contener *tokens* que representen valores de campos devueltos en la búsqueda. En la Fig.37 se emplean los *tokens $result.src_ip$* y *$result.dest_ip$* para brindar una información más explícita.

```
source="udp:514" index="tfm_security_logs" sourcetype="pfsense*"  "alert.signature"="ET EXPLOIT VSFTPD Backdoor User Login Smiley"
| stats count by _time, alert.signature, src_ip, dest_ip, dest_port, src_port
| sendemail from="[redacted].com" to="[redacted].com" subject="ET EXPLOIT VSFTPD Backdoor" sendresults=true content_type=html
    inline=true message="vsftpd backdoor vulnerability exploited (CVE-2011-2523) by $result.src_ip$ ... target $result.dest_ip$" sendcsv=true
```

Fig. 37. Exportación de la información relacionada con una alerta mediante el comando *sendemail*

Cada vez que se dispare la alerta, será enviado un e-mail al destinatario especificado informando sobre el suceso, la Fig.38 recoge los detalles.

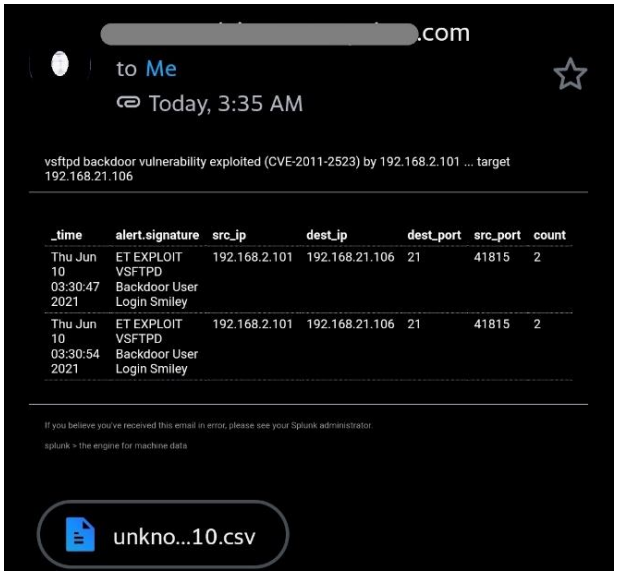

Fig. 38. Notificación recibida mediante-mail con información la alerta

El ejemplo previamente analizado también es aplicable a los reportes y a los *dashboards*. El comando *sendemail* se puede añadir en la búsqueda de uno de los paneles del *dashboard "INTERNAL SOC for TFM"* (Fig.39).

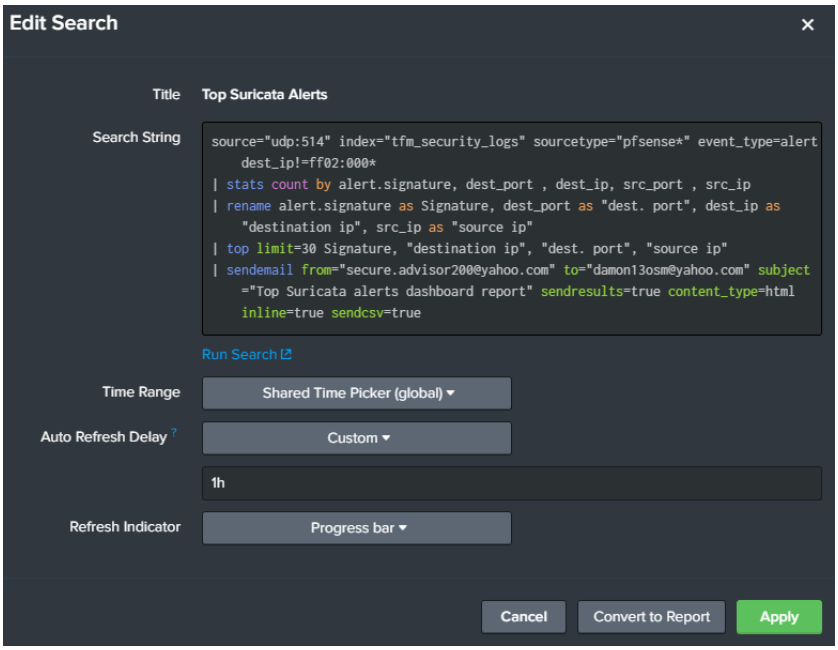

Fig. 39. Empleo del comando *sendemail* en un *dashboard* para exportar la información contenida

Definiendo el intervalo de actualización del panel se establece el período con que se enviará la notificación informando sobre su contenido, en este caso fue fijado cada una hora. El e-mail recibido, al igual que el antes observado para la alerta, contiene un archivo adjunto con formato *.csv* con los resultados de la búsqueda raíz en forma de tabla. Este fichero puede ser importado por otra plataforma para visualizar o utilizar la información incluida.

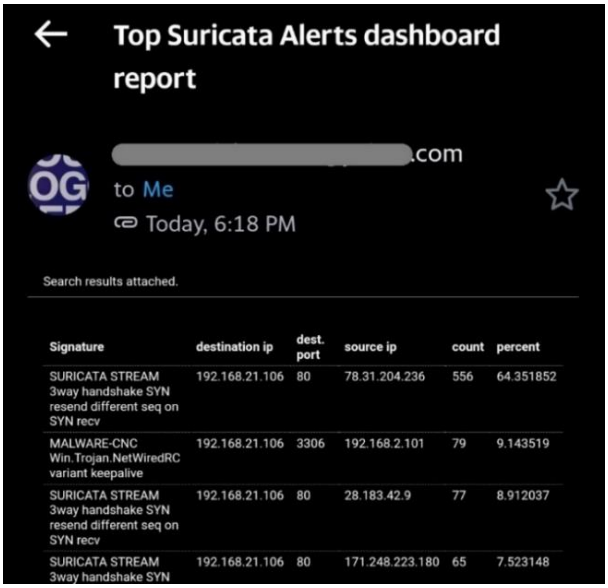

Fig. 40. Notificación recibida mediante-mail con la información del dashboard

V. Conclusiones

El creciente progreso de las TICs y de los servicios brindados a través de internet ha provocado el incremento de los usuarios y los dispositivos conectados. Paralelamente, la cibercriminalidad, el ciberterrorismo y otras variantes de actividades maliciosas en este entorno continúan aumentando. Para hacer frente a esto no basta solo con implementar los métodos y protocolos de seguridad existentes, sino que además es necesario mantener una Ciber Conciencia Situacional que garantice un enfoque proactivo sobre el asunto de la ciberseguridad. En este trabajo se propuso estudiar la generación de ciberinteligencia con la plataforma Splunk Enterprise, la cual ostenta con elevado prestigio en el campo de la ciberseguridad.

Primeramente, se llevó a cabo un estudio profundo de los conceptos fundamentales en torno al tema abordado, específicamente en el contexto actual, consultándose algunas de las fuentes bibliográficas de mayor reputación a nivel mundial. Además, se adquirieron y profundizaron conocimientos y habilidades sobre herramientas de seguridad de redes, de *hacking* ético y pruebas de penetración; que resultaron fundamentales para la realización de este proyecto.

Posteriormente se desplegó un entorno funcional y escalable donde se realizaron ciber ataques controlados a un objetivo puntual, y esto se monitorizó a través de un *firewall* de nueva generación que exporta los registros de actividad a la instancia de Splunk Enterprise. Luego se corroboró la capacidad de dicha plataforma de importar e indexar datos en bruto, incluso en tiempo real. Los mismos fueron procesados mediante búsquedas *SPL* con el fin de generar contenido de ciberinteligencia embebido en alertas, reportes y *dashboards* personalizados, permitiendo efectuar un análisis profundo y pragmático de lo sucedido en la red. También se consiguió exportar exitosamente esta ciberinteligencia por varias vías: en un fichero, vía e-mail y mediante uso de *webhook*; lo que posibilita la de Splunk integración con otras plataformas.

A través de las tareas desarrolladas y habiéndose cumplido con éxito los objetivos planteados en este trabajo, es posible concluir que Splunk Enterprise presenta un elevado potencial, flexibilidad y adaptabilidad como herramienta para la gestión de la ciberseguridad y generación de ciberinteligencia.

VI. Recomendaciones para trabajos futuros

Splunk ofrece bajo licencia de pago varias soluciones específicamente orientadas al tratamiento de la ciberseguridad y la generación de ciberinteligencia, como es el caso de Splunk Enterprise Security, User Behavior Analytics y Splunk Phantom. A las mismas no se pudo acceder para la realización de este proyecto de fin de máster, por lo que sería recomendable emplearlas en trabajos futuros con un enfoque similar al aquí propuesto.

También resultaría conveniente, aprovechar el esquema del entorno de simulación diseñado, pero empleando otras herramientas más potentes y accesibles bajo licencias de pago: *firewalls* como los proporcionados por Palo Alto Networks, Fortinet o Cisco; y aplicaciones de *penetration testing* como Metasploit Pro. Además, sería idóneo disponer de una infraestructura de red física y con mayor capacidad de cómputo.

A pesar de lo antes expuesto, cabe señalar que el escenario virtual implementado para la realización en este trabajo podría ser utilizado en múltiples investigaciones similares e incluso para actividades docentes de carácter práctico.

Bibliografía

[1] ITU, “Measuring Digital Development,” 2019. Accessed: Apr. 15, 2021. [Online]. Available: https://www.itu.int/en/mediacentre/Documents/MediaRelations/ITU Facts and Figures 2019 - Embargoed 5 November 1200 CET.pdf.

[2] “Measuring digital development. Facts and figures 2020,” 2020. Accessed: Apr. 15, 2021. [Online]. Available: https://www.itu.int/en/ITU-D/Statistics/Documents/facts/FactsFigures2020.pdf.

[3] C. Annual and I. Report, “White paper Cisco public,” 2018.

[4] L. C. Miller, J. Boardman, K. Cantillon, J. Dalton, M. Frohlich, and T. Trevethan, *Cybersecurity survival guide*, 5th ed., no. August. Palo Alto Networks, Inc., 2020.

[5] J. Task Force, “NIST Special Publication 800-37 Risk Management Framework for Information Systems and Organizations A System Life Cycle Approach for Security and Privacy JOINT TASK FORCE,” doi: 10.6028/NIST.SP.800-37r2.

[6] C. Paulsen and P. Toth, “NISTIR 7621 Revision 1 Small Business Information Security: The Fundamentals,” doi: 10.6028/NIST.IR.7621r1.

[7] Internet Crime Complaint Center, “2020 Internet Crime Report,” pp. 1–28, 2020, Accessed: Apr. 13, 2021. [Online]. Available: https://www.ic3.gov/Media/PDF/AnnualReport/2020_IC3Report.pdf.

[8] Interpol, “Ciberdelincuencia: Efectos de la COVID-19 (COVID-19 Cybercrime Analysis Report).,” *Secr. Gen. la Interpol*, pp. 1–20, 2020, [Online]. Available: https://www.interpol.int/es/Noticias-y-acontecimientos/Noticias/2020/Un-informe-de-INTERPOL-muestra-un-aumento-alarmante-de-los-ciberataques-durante-la-epidemia-de-COVID-19.

[9] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, “Zero Trust Architecture,” Gaithersburg, MD, Aug. 2020. doi: 10.6028/NIST.SP.800-207.

[10] “Framework for Improving Critical Infrastructure Cybersecurity, Version 1.1,” Gaithersburg, MD, Apr. 2018. doi: 10.6028/NIST.CSWP.04162018.

[11] “Committee on National Security Systems Glossary CNSSI No. 4009,” no. 4009, 2015, [Online]. Available: https://www.cnss.gov/CNSS/issuances/Instructions.cfm.

[12] “cyberspace - Glossary | CSRC.” https://csrc.nist.gov/glossary/term/cyberspace (accessed May 14, 2021).

[13] “Cybersecurity Glossary | National Initiative for Cybersecurity Careers and Studies.” https://niccs.cisa.gov/about-niccs/cybersecurity-glossary#S (accessed May 11, 2021).

[14] S. Jajodia, P. Liu, V. Swarup, and C. Wang, *Cyber situational awareness: Issues and research*, 1st ed., vol. 46. Boston, MA: Springer, 2010.

[15] E. D. Matthews, H. J. Arata, and B. L. Hale, “Cyber Situational Awareness,” *Cyber Def. Rev.*, vol. 1, no. 1, pp. 35–46, May 2016, [Online]. Available: http://www.jstor.org/stable/26267298.

[16] C. of the Joint Chiefs of Staff, “DOD Dictionary of Military and Associated Terms, January 2021.” Accessed: May 11, 2021. [Online]. Available: http://www.jcs.mil/Doctrine/DOD-Terminology/.

[17] A. Dehghantanha, M. Conti, and T. Dargahi, Eds., *Cyber Threat Intelligence*, vol. 70. Cham, 2018.

[18] “What Is Splunk? ‘Splunking’ of Data and More,” *Fortinet*. https://www.fortinet.com/resources/cyberglossary/what-is-splunk (accessed May 11, 2021).

[19] “Is Splunk a SIEM? | Security Information & Event Management.” https://www.comodo.com/is-splunk-a-siem.php (accessed May 12, 2021).

[20] G. Sadowski, K. Kavanagh, and T. Bussa, “Magic Quadrant for Security Information and Event

Management," 2020. Accessed: May 11, 2021. [Online]. Available: https://www.gartner.com/doc/reprints?id=1-1YEDHXVD&ct=200219&st=sb.

[21] "¿Qué es un firewall de nueva generación (NGFW)? | NGFW vs FWaaS | Cloudflare." https://www.cloudflare.com/es-es/learning/cloud/what-is-a-next-generation-firewall/ (accessed Apr. 29, 2021).

[22] "What Is a Next-Generation Firewall (NGFW)? - Cisco." https://www.cisco.com/c/en/us/products/security/firewalls/what-is-a-next-generation-firewall.html (accessed Apr. 29, 2021).

[23] "Best Firewalls 2021 | Top Enterprise Firewalls | Next-Generation Firewalls (NGFW) | IT Central Station." https://www.itcentralstation.com/categories/firewalls (accessed Apr. 29, 2021).

[24] "Best Free Firewalls for 2021 (9 for Windows and 1 for Mac)." https://www.comparitech.com/antivirus/best-free-firewalls/ (accessed Apr. 29, 2021).

[25] "10 Best Open Source Firewall for 2021 - Cyber Security News." https://cybersecuritynews.com/best-open-source-firewall/ (accessed Apr. 29, 2021).

[26] D. Sampaio and J. Bernardino, "Evaluation of Firewall Open Source Software," 2017, doi: 10.5220/0006361203560362.

[27] "Host-Only Networking." https://www.vmware.com/support/ws3/doc/ws32_network6.html (accessed May 03, 2021).

[28] "Kali Linux | Penetration Testing and Ethical Hacking Linux Distribution." https://kali.org/ (accessed May 04, 2021).

[29] "Metasploit Framework | Metasploit Documentation." https://docs.rapid7.com/metasploit/msf-overview/ (accessed May 04, 2021).

[30] "Metasploitable 2 Exploitability Guide | Metasploit Documentation." https://docs.rapid7.com/metasploit/metasploitable-2-exploitability-guide/ (accessed May 04, 2021).

[31] "Ataque DDoS de inundación SYN | Cloudflare." https://www.cloudflare.com/es-es/learning/ddos/syn-flood-ddos-attack/ (accessed Jul. 02, 2021).