Sistemas de Detección de Intrusiones y Gestión de Eventos e Información de Seguridad Basados en Nuevas Tecnologías de Código Abierto.

UNIVERSIDAD DE QUINTANA ROO

DIVISIÓN DE CIENCIAS, INGENIERÍA Y TECNOLOGÍA

Sistemas de Detección de Intrusiones y Gestión de Eventos e Información de Seguridad Basados en Nuevas Tecnologías de Código Abierto

MONOGRAFÍA
Para obtener el grado de
Ingeniero en Redes

PRESENTA
Joel Asunción Rangel Méndez

División de Ciencias, Ingeniería y Tecnología
UROO
Universidad de Quintana Roo

Supervisores propietarios
Dr. Homero Toral Cruz
Dr. José Antonio León Borges
M.M. Jesús Orifiel Álvarez Ruiz
Supervisores suplentes
Dr. Julio Cesar Ramírez Pacheco
M.M. José Raúl García Segura

Chetumal, Quintana Roo, México, 24 de noviembre de 2021

UNIVERSIDAD DE QUINTANA ROO

DIVISIÓN DE CIENCIAS, INGENIERÍA Y TECNOLOGÍA

Monografía elaborada bajo la supervisión del Comité del programa de licenciatura y aprobada como requisito para obtener el grado de:

INGENIERO EN REDES

COMITÉ DE MONOGRAFÍA

Supervisor propietario: _____
Dr. Homero Toral Cruz

Supervisor propietario: _____
Dr. José Antonio León Borges

Supervisor propietario: _____
M.M. Jesús Orifiel Álvarez Ruiz

Supervisor suplente: _____
Dr. Julio César Ramírez Pacheco

Supervisor suplente: _____
M.M. José Raúl García Segura

UNIVERSIDAD DE QUINTANA ROO
AREA DE TITULACION

Chetumal, Quintana Roo, México, 24 de noviembre de 2021

A mi madre Elizabeth Méndez Beltrán…… A mi padre Álvaro Rangel Aquino…….

A Dios…….

Mis hermanas Sagrario Rangel Méndez y Cinthia Rangel Méndez……

Mis abuelos……. Mi sobrina Pamela Gatica Rangel ……. A todos mis primos y tíos ….

A mis compañeros de carrera ……. A mis amigos …. A mi amigo José Juan Pech Kú…….

A mi amigo Ali Yassin Hussein …

A mis profesores……

A la universidad de quintana Roo……

A la vida……

Agradecimientos

A mi madre Elizabeth Méndez Beltrán, que en vida me dio todo lo posible a su alcance, a ti madre te dedico este logro en mi vida, que, sin tu apoyo, tus regaños, tus consejos, no sería nada de lo que soy, a ti, aunque ya no estas a nuestro lado, gracias a ti, he cumplido una meta más en mi vida, algo tarde para que tú lo puedas compartir conmigo, pero sé que donde estas, estar orgullosa de mí, aunque me gustaría estuvieras aquí a mi lado celebrando esta felicidad de a ver logrado titularme.

A mis abuelos que ya no están con nosotros, los amo y gracias por cuidarme desde el cielo, les dedico mis logros.

Gracias a dios por permitirme tener la dicha de estudiar, que a lo largo de mi vida en altas y bajas, al borde de abandonar mis estudios, me diste fuerza y voluntad de seguir paso a paso hasta terminar una profesión.

A mi padre Álvaro Rangel Aquino, por a verme dado lo mejor de él y apoyarme en mis estudios hasta donde pudo su economía, gracias, que, aunque tarde más tiempo. Logre la meta.

Gracias a mis hermanas Sagrario Rangel Méndez y Cinthia Rangel Méndez, por escucharme siempre y estar a mi lado, aunque no seamos los mejores hermanos, siempre estamos juntos como mama nos enseñó. A mi sobrina Pamela Elizabeth Gatica Rangel, te dedico este logro, para que veas lo difícil pero bonito esfuerzo y espero cuando tu llegues a tu carrera lo logres y enorgullezcas a tus padres.

Gracias a mis tíos (a), primos, que, en algún momento de este viaje, apoyaron económica y moralmente para seguir con mis estudios.

Le dedico unas sencillas palabras, pero con gran afecto, a mis amigos que vivieron y estuvieron conmigo a lo largo de esta travesía y en especial a un gran amigo que sin ti no podría a ver logrado este último paso, a ti amigo gracias por tu apoyo moral y económica, gracias, Ali Yassin Hussein.

Gracias a mis profesores que estuvieron en cada día, cada año enseñando y dando su tiempo para lograr terminar mi carrera universitaria

Agradezco al Dr. Homero Toral Cruz, que me asesoro en este proceso de titulación, días de desvelos, papeleos, verificación de documentos de protocolo, clases y más, gracias por la catedra.

A todos mis maestros de la academia de redes, Rubén González, Vladimir Cabañas, Melisa Blanqueto, Laura Dávalos, Jaime Ortegón, Víctor Huerta, Fredy Chan.

Gracias a esta casa de estudio por a verme formado con los más altos estándares de calidad a lo largo de mi estancia de recorrido por cada aula, cada pasillo y cada departamento, gracias a todos por el apoyo brindado.

CONTENIDO

Resumen

La rápida evolución de las redes de datos ha provocado la aparición de nuevas tecnologías para compartir, transferir y distribuir cualquier tipo de información. Esto ha provocado una mayor conciencia sobre la importancia y necesidad de salvaguardar la integridad de los datos y protegerlos contra las amenazas basadas en red (introducción de código dañino en sistemas, ataques a páginas web, filtración de información, fraude electrónico, etc.).

En temas de seguridad informática, una de las soluciones más aplicadas para detectar comportamientos anómalos o maliciosos y registrar tales eventos son los sistemas de detección de intrusiones (Intrusion Detection System - IDS). Esta solución se puede entender como la evolución del concepto “antivirus”, y permite detectar más tipos de ataques como la denegación de servicio (Denial of Service - DoS) o su variante, la denegación de servicio distribuido (Destributed Denial of Service – DDoS), robo de información, ataque por encuentro a medio camino (Meet-in-the-Middle - MITM), [1], entre otros.

Para analizar la información generada por el IDS, se propone la implementación de un sistema de Gestión de Eventos e Información de Seguridad (Security Information and Event Management - SIEM), el cual combina las funciones de un sistema de Gestión de Información de Seguridad (Security Information Management - SIM) y un sistema de Gestión de Eventos de Seguridad (Security Event Management - SEM). El sistema SIM se encarga del almacenamiento, el análisis y la comunicación de los datos de seguridad. El sistema SEM se encarga del monitoreo en tiempo real, correlación de eventos, notificaciones y vistas de la consola de la información de seguridad. De esta manera, el sistema SIEM, centraliza el almacenamiento y el análisis de la información relevante de seguridad con la finalidad de permitir a los equipos de seguridad controlar todo lo que esta pasando en la red en tiempo real y reaccionar rápidamente ante posibles ataques y vulnerabilidades.

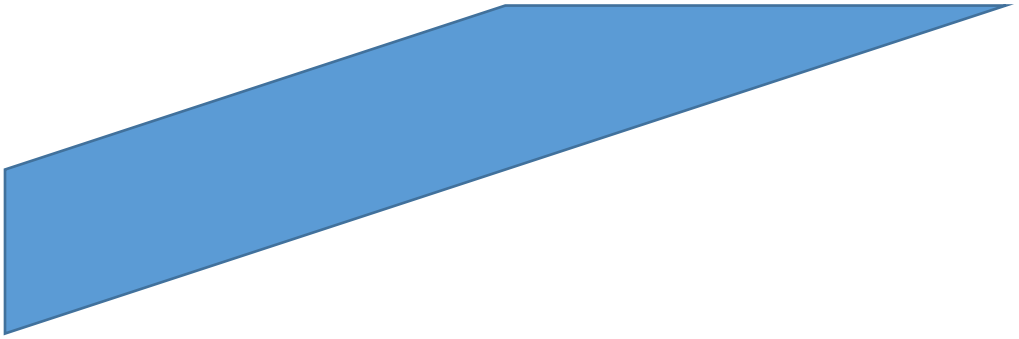
Existen herramientas comerciales que integran sistemas SIEM, tales como QRadar de IBM, RSA enVision, Security MARS, empow o Alien Vault USM; las cuales, tienen un alto costo. Sin embargo, una solución viable es mediante software de código abierto como ELK Stack (Elasticsearch-Logstash-Kibana).

Con base a los puntos mencionados anteriormente, en este trabajo se propone la implementación de una solución de seguridad que permita detectar comportamientos anómalos o maliciosos, alertar al administrador cada vez que se detecte algún evento sospechoso y analizar el estado de la red y el origen de los eventos.

Tabla de Ilustraciones

Índice de Tablas

CAPITULO 1



1 INTRODUCCIÓN

Hoy en día, la palabra hacker es un término que en ocasiones puede provocar temor entre la comunidad que han sido víctimas de sus acciones negativas o maliciosas, algunas organizaciones invierten muchos recursos para contrarrestar la presencia de los hackers mal intencionados (black hat hackers), mientras otras lo hacen para contratar a los hackers éticos (white hat hacker o crackers).

La diferencia entre los buenos hackers, también llamados hackers de sombrero blanco (white hat hacker) o hackers éticos y los hackers mal intencionados (black hat hackers) o crackers [2], es que los primeros tienen autorización expresa de revisar, probar, descubrir y modificar los sistemas informáticos con la finalidad de detectar vulnerabilidades y posteriormente desarrollar y aplicar medidas de seguridad, actualizaciones o mejoras; en cambio, los segundos irrumpen en dichos sistemas con la intención de robar o destruir información, sabotear, cometer fraudes y generar caos, actuando de manera ilegal e irresponsable. La gran ventaja de los hackers de sombrero blanco es que al tener habilidades y capacidades muy parecidas a las de sus contrapartes, toman acciones preventivas eficientes que les permiten proteger adecuadamente los sistemas, evitando posibles ataques que pudiesen debilitar la seguridad de la infraestructura informática de una organización.

Sin embargo, también existen los hackers de sombrero gris (gray hat hackers). Como su color lo indica, son una mezcla entre los sombreros blancos y negros. Se dedican a identificar vulnerabilidades y en ocasiones, comprometer la seguridad de los sistemas; una vez encontradas, establecen contacto con los propietarios para informar al respecto, solicitando eventualmente algún tipo de pago o recompensa económica por sus servicios. En principio, sus propósitos no son malignos, pero al hacerlo sin permiso caen en la ilegalidad. Adicionalmente a ellos, están los scripts kiddies, personajes que, con poca experiencia en el área del cómputo, hacen uso de herramientas empleadas o desarrolladas por los hackers y realizan ataques informáticos.

Los peligros que enfrentan los sistemas informáticos (Information Technology - IT) con impacto en las tecnologías de información y comunicación (TIC) han incremento exponencial durante los últimos años, debido al aumento de la actividad mundial de los hackers. Los atacantes se han vuelto más sofisticados y peligrosos y su detección adecuada y oportuna se ha convertido

en un verdadero reto. Los principales ataques que afectan a las IT y las TIC son [3], [4]: ataques de ransomware; malware que tiene un impacto negativo en la capacidad de la empresa sobre el uso de servicios públicos para realizar negocios y operaciones; campañas de phishing dirigidas a ejecutivos, asistentes ejecutivos, administradores de IT u otros usuarios privilegiados; incidentes de compromiso de correo electrónico empresarial, incluida la toma de control de cuentas o la suplantación de ejecutivos; fuga y robo de datos; ingeniería social para recopilar información sensible del personal.

En las tecnologías de información y comunicación, la mayoría de los dispositivos pueden generar, almacenar y enviar información. Sus registros pueden provenir de diferentes y numerosas fuentes, incluidos firewalls, sistemas de detección de intrusiones (Intrusion Detection System IDS), sistemas de prevención de intrusiones (Intrusion Prevention System - IPS) y redes privadas virtuales (Virtual Private Network - VPN). Los sistemas recopilan y analizan esta enorme cantidad de información. Estas soluciones de sistema tienen muchos acrónimos diferentes, como gestión de seguridad empresarial (Enterprise Security Management - ESM), Gestión de eventos empresariales (Enterprise Event Gestión - EEM), gestión de la información de seguridad (SIM), Gestión de eventos de seguridad (SEM) y Gestión de eventos de información de seguridad (SIEM). Nos referimos a todos ellos con el término SIEM y generalmente está diseñado para proporcionar los siguientes servicios [5], [6]:

- Gestión de registros: recolecta, almacena y analiza todos los registros.
- Cumplimiento normativo de IT: audita y valida el comportamiento, identifica violaciones, intrusiones, intentos de ataque o accesos no autorizados.
- Correlación de eventos: analiza y relaciona automáticamente los datos para poder identificar riesgos y vulnerabilidades.
- Respuesta activa: implementa contramedidas actuando directamente desde el sistema SIEM.

Existen diferentes proveedores que pueden desarrollar los dispositivos (IDS, IPS, firewall, computadoras), que generan los datos de entrada (logs, Json, etc.). Para el sistema SIEM, sus datos generalmente se guardan en diferentes formatos propietarios. Incluso la forma en que los eventos se informan a las funciones del servidor de registro ascendente puede no ser universal [5], [7]. Esto puede crear incompatibilidad al analizar datos de diferentes fuentes. Algunas

normas abordan este problema, un campo de investigación interesante se refiere a las estrategias de correlación utilizadas por el motor de reglas. Si bien algunos patrones de ataque pueden ser detectados mediante el uso de reglas simples, los ataques más complejos pueden requerir enfoques más avanzados, como los algoritmos de aprendizaje automático (ML - Machine Learning) para ser detectados.

Cuando el sistema haya recolectado los datos, deberá priorizarlos y alertar sobre los posibles problemas. Estos sistemas pueden ser utilizados para el monitoreo en tiempo real, lo cual permite dar una respuesta inmediata ante posibles amenazas. Así también, se puede usar para realizar un análisis posterior, brindando el apoyo a investigaciones. Con base en esto, nuevos paradigmas tecnológicos como Big Data ofrecen grandes posibilidades para el análisis de sistemas complejos. Por esta razón, los motores de correlación en los sistemas IDS y SIEM, pueden hacer uso de la tecnología Machine Learning. El proceso de seguridad que implementa es disuadir-detectar-retrasar-responder.

En este trabajo monográfico se presenta la documentación sobre la implementación de una solución de seguridad, basada principalmente en un Sistema de Detección de Intrusiones y un Sistema de Gestión de Eventos e Información de Seguridad. Dicha solución permitirá detectar comportamientos anómalos o maliciosos, alertar al administrador cada vez que se detecte algún evento sospechoso y analizar el estado de la red y el origen de los eventos.

1.1 Justificación

Uno de los objetivos principales de los hackers es obtener información sensible de una o varias personas, las cuales pueden ser como: números de tarjetas de crédito, número de seguridad social, contraseñas, o toda clase de información sensible, etc.

Para lograr este objetivo los hackerss pueden utilizar una o varias técnicas de ataque y pueden demorar pocos o varios días, meses, etc. Entre los ataques más conocidos tenemos: phishing, keyloggers, virus, gusanos, troyanos, etc. Una de las principales ventajas que aprovechan los ciberdelincuentes es el error humano; por ejemplo, en el caso del phishing aprovechan este error para extraer información personal de la víctima.

Otro método utilizado, consiste en aprovechar la vulnerabilidad existente en el software que utilizan comúnmente los usuarios, dichas vulnerabilidades son aprovechadas por los atacantes que consiguen acceder al sistema vulnerado e implantar su backdoor3 (software y/o script) para obtener información sensible de su víctima. En ciertos casos el atacante intenta infectar otras máquinas del entorno, en este momento, el análisis del tráfico de red juego un papel importante en el monitoreo y captura de información del software malicioso [8].

Entre los métodos comúnmente utilizados para el análisis del tráfico de red en el ámbito de seguridad, destacan los sistemas de detección de intrusiones (Intrusión Detection System - IDS). Los IDSs permiten identificar las amenazas de las máquinas vulneradas por medio de sus sensores y las alertas de seguridad que estos generan, basados en reglas de seguridad previamente configuradas de acuerdo con el entorno en el que se encuentren [9]. Por otro lado, para analizar la información generada por los IDSs, se puede utilizar un sistema de Gestión de Eventos e Información de Seguridad. De esta manera, el sistema SIEM, centraliza el almacenamiento y el análisis de la información relevante de seguridad con la finalidad de permitir a los equipos de seguridad controlar todo lo que esta pasando en la red en tiempo real y reaccionar rápidamente ante posibles ataques y vulnerabilidades.

Con base a los puntos mencionados anteriormente, en este trabajo trabajo monográfico se presenta la documentación referente a la implementación de una solución de seguridad que permita detectar comportamientos anómalos o maliciosos, alertar al administrador cada vez que se detecte algún evento sospechoso y analizar el estado de la red y el origen de los eventos.

1.2 Objetivos

1.2.1 Objetivos Generales

Documentar la implementación de un IDS y un SIEM, combinando nuevas tecnologías de código abierto para detectar comportamientos maliciosos o anómalos y analizar el estado de una red en un ambiente controlado.

1.2.2 Objetivos Específicos

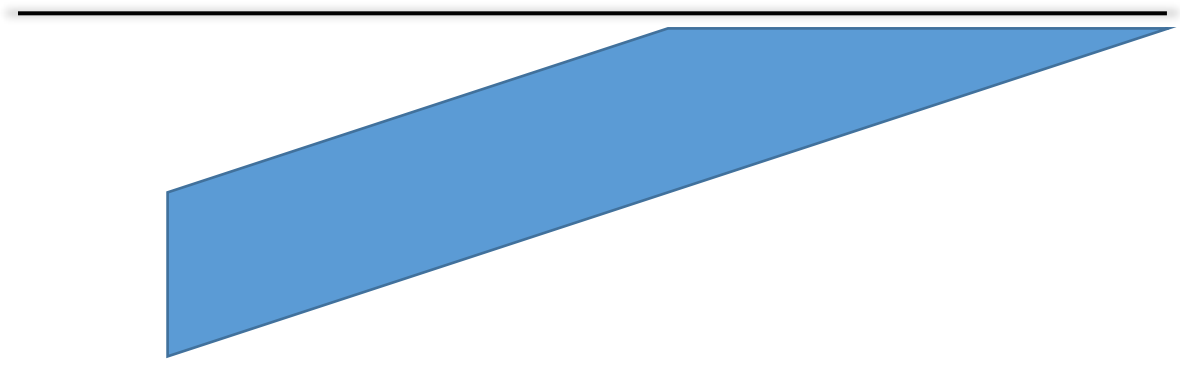
- Realizar un estudio del estado del arte de los Sistemas de Detección de Intrusiones y Gestión de Eventos e Información de Seguridad.

- Elegir el IDS de código abierto a implementar.
- Configurar el IDS para la detección de posibles intrusiones en la red.
- Implementar y configurar el sistema SIEM.
- Realizar pruebas del sistema integrado y documentarlas.

1.3 Metodología

Se realizará un trabajo de investigación documental sobre los conceptos básicos de seguridad informática, la implementación de un sistema de detección de intrusiones y un sistema de gestión de eventos e información de seguridad; para proporcionar a los usuarios de red, información sobre las posibles amenazas a las cuales se encuentran expuestos los sistemas de información y algunas técnicas para detectar y prevenir ataques e intrusiones a nuestros sistemas.

CAPITULO 2



2 CONCEPTOS DE SEGURIDAD INFORMÁTICA

2.1 Definición de Seguridad Informática

La seguridad informática se puede definir como la protección otorgada a un sistema de información automatizada, con el fin de preservar la integridad, disponibilidad y confidencialidad de los recursos del sistema de información (incluye hardware, software, firmware, información / datos y telecomunicaciones, etc.) [10], [11].

Esta definición introduce tres objetivos clave de la seguridad informática:

Confidencialidad: Este término cubre dos conceptos relacionados:

- **Confidencialidad de los datos**: Asegura que la información privada o confidencial no sea puesta a disposición o liberada a personas no autorizadas.
- **Privacidad:** Asegura que las personas controlen o influyan en que información relacionada con ellos pueden recopilar y almacenar y quién y para quién, esa información puede ser divulgada.

Integridad: Este término contine dos conceptos relacionados:

- **La Integridad de los datos**: Asegura que la información y los programas solo se modifiquen de una manera especificada y autorizada.
- **Integridad del sistema**: Asegura que un sistema realice su función prevista de una manera intachable, libre de intención no autorizada deliberada o inadvertida del sistema.

Disponibilidad: Asegura que los sistemas funcionen con prontitud y que no se niegue el servicio a usuarios autorizados.

Estos tres conceptos forman lo que a menudo se conoce como la tríada de la CIA (Confidentiality, Integrity, Availability). Los tres conceptos incorporan los objetivos fundamentales de seguridad tanto para los datos como para los servicios informáticos y de información. El estándar NIST FIPS 199 (Standards for Security Categorization of Federal

Information and Information Systems) considera la confidencialidad, integridad y disponibilidad como los tres objetivos de seguridad de la información y de los sistemas de información [10]:

Confidencialidad: Preservar las restricciones autorizadas sobre el acceso a la información y divulgación, incluidos los medios para proteger la privacidad personal y la información de propiedad. Una pérdida de confidencialidad es la divulgación no autorizada de información.

Integridad: Proteger contra la modificación o destrucción indebida de la información, incluyendo asegurar el no repudio y la autenticidad de la información. Una pérdida de integridad es la modificación o destrucción no autorizada de información.

Disponibilidad: Garantizar el acceso y el uso oportuno y confiable de la información. Una pérdida de disponibilidad es la interrupción del acceso o uso de información o un sistema de información.

Aunque la triada de la CIA define bien los objetivos de seguridad, algunos en el campo de la seguridad sienten que se necesitan conceptos adicionales para presentar una imagen completa, como los que se muestran a continuación:

Autenticidad: Propiedad de ser genuino y poder ser verificado y confiable, confianza en la validez de una transmisión, un mensaje o creador del mensaje. Esto significa verificar que los usuarios sean quienes dicen ser y que cada entrada que llega al sistema proviene de una fuente confiable.

Responsabilidad: El objetivo de la seguridad que genera la exigencia de acciones de una entidad para ser rastreado exclusivamente a esa entidad. Esto apoya al no repudio, disuasión, aislamiento de fallas, detección y prevención de intrusiones, acción de recuperación posterior y acción legal. Porque los sistemas verdaderamente seguros aún no son alcanzables, debemos poder rastrear una brecha de seguridad hasta una parte responsable. Los sistemas deben mantener registros de sus actividades para permitir un análisis forense posterior para identificar las violaciones de seguridad o para ayudar en disputas de transacciones [10], [11].

2.2 Desafíos de la Seguridad Informática

Los desafíos de la seguridad informática son complejos, algunas de las razones son las siguientes:

- La seguridad informática no es tan simple como podría parecerle a un principiante. Los requisitos parecen ser sencillos; de hecho, la mayoría de los requisitos principales para los servicios de seguridad pueden recibir etiquetas auto explicativas de una palabra [11]: confidencialidad, autenticación, no repudio, integridad. Pero los mecanismos utilizados para cumplir con esos requisitos pueden ser bastante complejo, y la comprensión pueden implicar un razonamiento bastante sutil.
- Al desarrollar un algoritmo o mecanismo de seguridad particular, siempre se debe considerar posibles ataques a esas características de seguridad. En muchos casos el éxito de los ataques se diseña considerando el problema de una manera completamente diferente, aprovechando por tanto una debilidad inesperada en el mecanismo.

1. Debido al punto anterior, los procedimientos utilizados para proporcionar servicios particulares son a menudo contradictorios. Normalmente, un mecanismo de seguridad es complejo y no es obvio a partir de la declaración de un requisito particular que se necesiten medidas tan elaboradas. Solo cuando los diversos aspectos de la amenaza son considerados, los mecanismos de seguridad elaborados tienen sentido.
2. Habiendo diseñado varios mecanismos de seguridad, es necesario decidir dónde usarlos. Esto es cierto tanto en términos de ubicación física (en qué puntos de una red son necesarios ciertos mecanismos de seguridad) como en un sentido lógico (en qué capa o capas de una arquitectura TCP / IP) deben ser posicionados los mecanismos.
3. Los mecanismos de seguridad generalmente intervienen más que un algoritmo particular o protocolo. También requieren que los participantes estén en posesión de alguna información secreta (una clave de cifrado), lo que plantea preguntas sobre la creación, distribución y protección de esa información secreta. Puede también se una dependencia de los protocolos de comunicación cuyo comportamiento puede complicar la tarea de desarrollar el mecanismo de seguridad. Por ejemplo, si el correcto funcionamiento del mecanismo de seguridad requiere establecer límites de tiempo, el tiempo de tránsito de un mensaje del remitente al receptor, luego cualquier protocolo o

la red que introduce retrasos variables e impredecibles puede generar tal tiempo límite sin sentido.

4.La seguridad informática es esencialmente una batalla de ingenio entre un perpetrador que intenta encontrar agujeros y el administrador que intenta cerrarlos. La gran ventaja que tiene el atacante es que solo necesita encontrar una sola vulnerabilidad, mientras que el administrador o experto en ciberseguridad, debe encontrar y eliminar todas las fallas críticas y fortalecer las debilidades de los sistemas para lograr una seguridad lo más perfecta posible.

5.Existe una tendencia natural por parte de los usuarios y administradores de sistemas, perciben pocos beneficios en la inversión en seguridad, hasta que se produce una falla en el sistema de seguridad.

6.La seguridad requiere un monitoreo regular, incluso constante, y esto es difícil en el corto plazo actual y en ambiente sobrecargado.

7.La seguridad sigue siendo una idea tardía que no se puede incorporar a un sistema después de que el diseño esté completo en lugar de ser una parte integral del diseño de red implementado.

8.Muchos usuarios e incluso administradores de seguridad ven la seguridad sólida como un impedimento para el funcionamiento eficiente y fácil de usar de un sistema o uso de información.

Las dificultades que acabamos de enumerar se encontrarán de numerosas formas a medida que examine las diversas amenazas y mecanismos de seguridad.

2.3 Ataque de Red

Durante décadas, las tecnologías de las redes se han utilizado para mejorar la transferencia de información. Sus continuas mejoras han facilitado una amplia gama de nuevos servicios y tecnologías, como Internet de las cosas (Internet of Things - IoT). IoT es una herramienta poderosa utilizada para mejorar la comunicación, al conectar diferentes dispositivos a Internet y recopilar información. La información recopilada ayuda a las empresas en el análisis y la previsión del comportamiento del consumidor para mejorar la calidad de sus productos. Otras tecnologías como Machine Learnig y Deep Learning, se utilizan para construir sistemas de red que pueden realizar análisis avanzados y automatizados. Esta tecnología está transformando

las experiencias de red de los usuarios mediante la simulación del intelecto humano y la recopilación de datos con algoritmos incorporados [12], [13].

Tecnologías emergentes de computación en la nube han traído consigo notables evoluciones en la tecnología de redes de comunicación, donde diferentes aplicaciones, servicios de computación y recursos de almacenamiento se ofrecen bajo demanda a un gran número de usuarios a través de Internet [14], [13].

Un nuevo estándar inalámbrico global, la red móvil de quinta generación (5G), que representa un tipo de red lógica que conecta esencialmente cualquier cosa, incluidas las máquinas, objetos y gadgets. 5G no solo ofrece velocidades más rápidas y una mayor cantidad de enlaces de dispositivos, también permite el corte de red. El corte de red es el proceso de dividir varias redes virtuales que operan en la misma infraestructura de red para crear subredes que satisfacen las demandas de diversas aplicaciones. Desde entretenimiento y juegos, hasta la escuela y seguridad comunitaria. La tecnología de red 5G tiene la capacidad de desarrollar cualquier cosa. 5G tiene el potencial para proporcionar velocidades de descarga más altas, respuestas en tiempo real y una conexión mejorada a lo largo del tiempo, lo que permite a las empresas y los consumidores explorar nuevas innovaciones [15], [13].

Un crecimiento tan exponencial en las tecnologías de red ha ofrecido muchas ventajas y ha mejorado enormemente las comunicaciones. Sin embargo, cada tecnología de red emergente presenta nuevos desafíos de seguridad y desencadena la necesidad de desarrollar sistemas de detección, herramientas y contramedidas para satisfacer las nuevas demandas.

2.4 Tipos de Ataques de Red

Vivimos en la revolución de la era digital, hoy en día la mayoría de las personas utilizan una computadora con Internet. Debido a la dependencia de las herramientas digitales**,** la actividad informática ilegal ha crecido grande y continuamente busca nuevas y más efectivas formas de vulnerar una red de datos, ya sea una red doméstica, empresaria, de servicios, etc.

Un ataque de red es un enfoque para dañar, revelar, cambiar, destruir, robar u obtener información ilegal, acceso a un recurso del sistema de red. El ataque puede provenir del interior (ataque interno) o desde afuera (ataque externo).

Existen muchos ejemplos de ataques de red que han aparecido con este cambio y revolución digital, estos podríamos clasificarlos como ataques del tipo phishing attacks, malware attacks, web attacks [16]. Cada una de estas clasificaciones contiene diferentes tipos y variantes, a continuación, mencionaremos algunos de estos en las siguientes tablas.

Tabla 1: Ataques DOS [13]

Nombre de Ataque	**Descripción**	**Ataque por: Paquete, Herramientas, etc.**
Ataques Activos **Ataques de denegación de Servicio**		
Jamming Attack	Al usar el canal en el que se están comunicando, prohíbe otros nodos accedan a él para conectarse.	Ruido en radiofrecuencias.
Flooding	Ataque DoS en el que un servidor recibe muchas solicitudes de conexión, pero no responde para completar las peticiones a este (inundación ICMP, SYN Flood, HTTP Flood).	Número de solicitudes sin consolidar, sin reconocimiento de paquetes después de recibirlo.
Smurf Attack	Ataque DDoS en la capa de red, causado por la red de instrumentos mal configuración.	IP de origen engañando a la IP de la víctima.
Teardrop Attack	Ataque DoS que bombardea una red con muchos fragmentos de datos del protocolo IP, entonces la red no puede recombinar los fragmentos de nuevo en sus paquetes originales.	Enviar paquetes fragmentados a la máquina destino.

Tabla 2: Ataque Man in the Middle [13]

Nombre de Ataque	**Descripción**	**Ataque por: Paquete, Herramientas, etc.**
Ataque Activos **Ataques Man in the Middle**		
Ransomware	Malware que se infiltra y cifra archivos importantes y sistemas, y que impiden que a una persona acceda a sus propios datos.	B0r0nt0k (ransomware de cifrado), Mado (programa malicioso).
Session Hijacking	Para obtener acceso no autorizado al servidor web. El ataque de secuestro de sesión interrumpe el token de sesión al robar o adivinar un token de sesión válido (token de sesión predecible).	Códigos JavaScript maliciosos, XSS, sesión de olfateo.

Tabla 3: Ataques Pasivos [13]

Nombre de Ataque	**Descripción**	**Ataque por: Paquete, Herramientas, etc.**
Ataques Pasivos		
Reconocimiento Activo	Un intruso se dedica a apuntar al sistema para adquirir información sobre vulnerabilidades (escaneo de puertos).	Nmap, Metasploit.
Reconocimiento Pasivo	Recopila información sobre computadoras y redes sin participar activamente con ellos (escuchar a escondidas, huellas digitales del sistema operativo).	Wireshark, Shodan.
Análisis de Tráfico	Método para recopilar y monitorear tramas, paquetes o mensajes para impulsar la información a los patrones de comunicación.	Sniffing tools (Herramientas de monitoreo y captura).

War Driving	Mapeo de los puntos de acceso inalámbricos con redes inalámbricas con vulnerabilidades en automóviles en movimiento	iStumbler, Global Positioning System (GPS), antenna, Wifiphisher.

Tabla 4: Ataques Bitcoin [13]

Nombre de Ataque	**Descripción**	**Ataque por: Paquete, Herramientas, etc.**
Ataques Pasivos **Ataque Bitcoin**		
Zero Access	Ataque que tiene un patrón desconocido o tiene como objetivo explotar una vulnerabilidad de seguridad de software potencialmente grave que el desarrollador o el personal de seguridad no tienen conocimiento.	Vulnerabilidades no descubiertas (más difícil de detectar).
Credential Stuffing	Especie de ciberataque en el que los atacantes irrumpen en un sistema, utilizando una lista de credenciales de usuario comprometidas (ataques de diccionario).	Bots para automatización, direcciones IP falsa.
Account Takeover	La apropiación de cuenta es como el robo de identidad en el que un delincuente accede de forma no autorizada a la cuenta de otra persona (phishing, fraude en el centro de llamadas)	Obtención de credenciales comprometidas.
Account Lockout	Atacante que no tiene acceso a credenciales de usuarios genuinos del sitio web, pero, sin embargo, les hace daño al tomar ventaja de los mecanismos de seguridad (ataque de fuerza bruta).	Bloquear una gran cantidad de cuentas de usuario.

Tabla 5: Brechas de Seguridad [13]

Nombre de Ataque	**Descripción**	**Ataque por: Paquete, Herramientas, etc.**
Ataques Pasivos **Brechas de Seguridad**		
Vulnerability Scanning	Proceso automatizado continuo para encontrar fallas de seguridad en sitios web en una red para explotar amenazas y atacar esos sitios web.	Bots que buscan problemas de seguridad y vincularlos a vulnerabilidades conocidas en una base de datos.
API Abuse	El abuso de API se define como el acceso no autorizado o ilegal a un API del servidor a través de aplicaciones móviles o de escritorio.	Robar códigos de aplicación para valiosa propiedad intelectual.

2.5 Técnicas de Prevención y Detección de Ataques

Los sistemas de seguridad y defensa están diseñados para identificar, defender y recuperarse de ataques a la red. La confidencialidad, la disponibilidad y la integridad son los tres objetivos principales de los sistemas de seguridad de la red de comunicación e información. Técnicas de prevención y detección de intrusiones en la red pueden clasificarse según el enfoque utilizado para detectar amenazas de red, prevenirlas, o una combinación de ambos.

Estas técnicas se desarrollan como software, hardware o una combinación de ambos. Se pueden clasificar en dos clases: sistemas de detección de intrusiones (IDS) y sistemas de prevención de intrusiones (IPS) [17], [18], [13]:

- Sistema de detección de intrusiones (IDS): también denominado IDS basado en red (NIDS). Este sistema monitorea intensamente las actividades maliciosas de la red y notifica a los usuarios si se produce un ataque. Se detecta sin capacidad de prevención. Detección basada en firmas y basada en anomalías son los dos enfoques más utilizados por IDS para identificar amenazas. Se aplican procedimientos basados en firmas para detectar solo amenazas conocidas, basándose en una base de datos que contiene una

lista de características preexistentes de ataques conocidos (firmas de ataques) para identificar eventos sospechosos. La base de datos debe actualizarse continuamente para incluir ataques emergentes. Por otro lado, los procedimientos basados en anomalías intentan diferenciar el tráfico malicioso del tráfico real en función de un cambio en la red tráfico; por tanto, pueden detectar amenazas desconocidas. Inconsistencias como tráfico de gran tamaño, latencia de red, tráfico de puertos poco comunes y rendimiento anormal del sistema, todos representan cambios en los comportamientos normales del sistema y pueden indicar la presencia de ataques a la red.

•Sistema de prevención de intrusiones (IPS): conocido también como detección y prevención de intrusiones de sistemas (IDPS). Escanea la red continuamente para detectar la presencia de atacantes, software malicioso entre otros, puntos de control que se detectan sobre la base de cambios en el comportamiento. El sistema toma automáticamente contramedidas para hacer frente a las amenazas y defender el sistema. El objetivo principal de un IDPS es evitar que paquetes y ataques maliciosos o no deseados causen daño. Un IDPS es más eficaz que IDS, ya que no solo detecta amenazas, también puede tomar medidas contra ellos.

Hay dos tipos de IDPS: sistemas de detección y prevención de intrusiones basados en la red (NIDPS) que analizan el protocolo de red para identificar cualquier actividad sospechosa, y sistemas de detección y prevención de intrusiones basadas en host (HIDPS) que se utilizan para monitorear las actividades del host en busca de eventos sospechosos dentro del anfitrión. Para identificar los ataques de manera eficaz y eficiente, se utilizan una variedad de enfoques de detección, en constante desarrollo sobre la base de técnicas inteligentes que incluyen Machine Learnig y Deeps Learning, que recientemente han ganado una inmensa popularidad en el campo de la seguridad de redes.

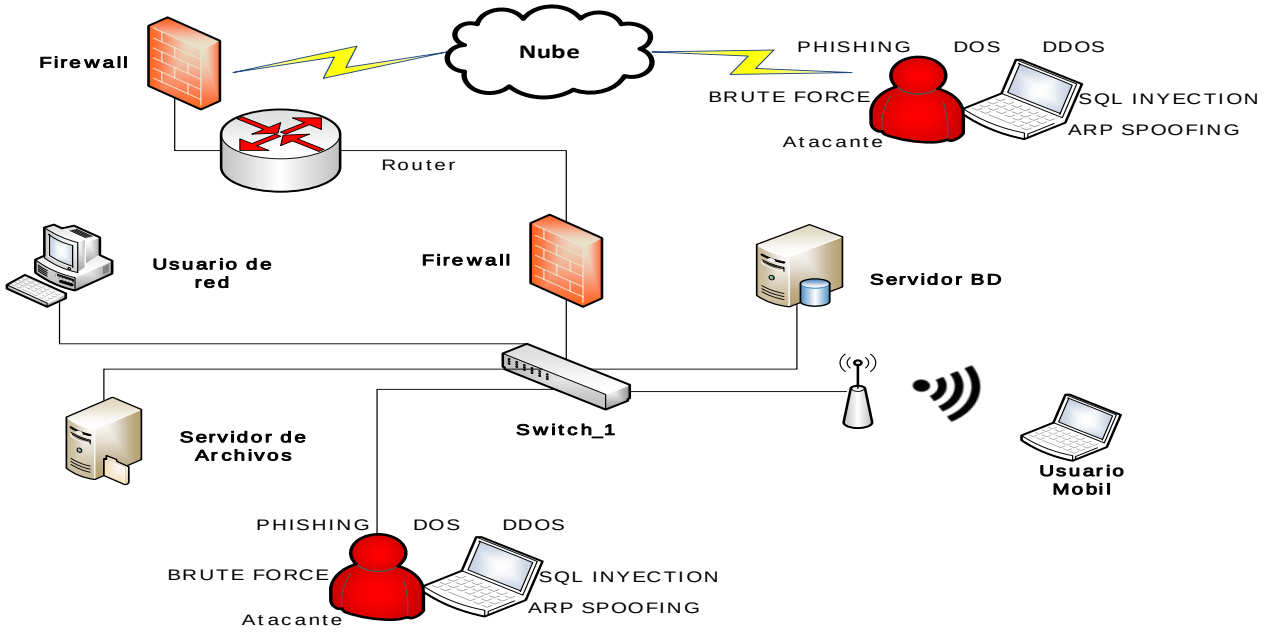
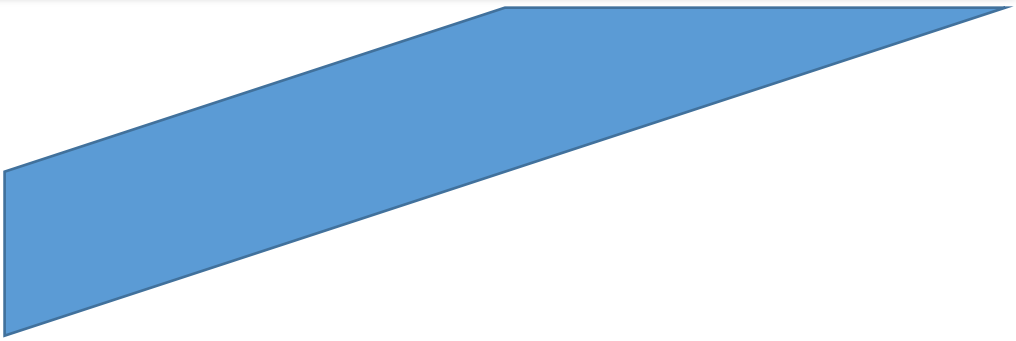


Ilustración 1: Ataque de Red

CAPITULO 3



3 SISTEMAS DE DETECCIÓN DE INTRUSIONES Y DE GESTIÓN DE EVENTOS E INFORMACIÓN DE SEGURIDAD

3.1 Sistema de Detección de Intrusiones

Los sistemas de detección de intrusiones son herramientas que puede ser utilizada para proteger una red de ataques informáticos. La detección de intrusiones es un mecanismo utilizado para identificar a un hacker cuando intenta una penetración. Idealmente, un sistema de este tipo solo emitirá una alarma cuando se realice un ataque con éxito. La detección de intrusiones también puede ayudar en la identificación proactiva de amenazas activas al proporcionar indicaciones y advertencias de que una amenaza está recopilando información para un ataque.

Las alarmas antirrobo y de coche también son formas de IDS. Si el sistema de alarma detecta un evento anómalo (como la rotura de cristal o la apertura de una puerta), se encienden las luces, suena una alarma o se llama a la policía; es decir, se proporciona la función disuasoria. Todos estos ejemplos comparten un único objetivo principal: detectar cualquier intento de penetrar el perímetro de seguridad del elemento (negocio, edificio, automóvil, etc.) que se está protegiendo. En el caso de un edificio o automóvil, el perímetro de seguridad es fácil de identificar. Las paredes del edificio, una cerca alrededor de la propiedad o las puertas y ventanas del automóvil definen claramente el perímetro de seguridad. Otra característica que todos estos ejemplos tienen en común es criterios bien definidos para lo que constituye un intento de penetración y lo que constituye el perímetro de seguridad.

Si traducimos el concepto de sistema de alarma al mundo informático, tenemos el concepto básico de un IDS. Ahora debemos definir cuál es el perímetro de seguridad de nuestra computadora, sistema o red. Claramente, el perímetro de seguridad no existe de la misma manera como una pared o cerca. El perímetro de seguridad de una red se refiere al perímetro virtual que rodea los sistemas informáticos. Este perímetro se puede definir mediante firewalls, puntos de demarcación de telecomunicaciones o computadoras de escritorio con módems. Eso

también puede ampliarse para incluir las computadoras personales de los empleados a los que se les permite teletrabajar o un socio comercial que tenga permiso para conectarse a la red. Un IDS está diseñado para diferenciar entre una entrada y una intrusión maliciosa.

Un IDS es como el guardia en la puerta de entrada de un centro comercial, observando a todos los clientes en la búsqueda de intenciones maliciosas (por ejemplo, portar un arma). Desafortunadamente, en el mundo virtual, el arma a menudo es invisible.

La segunda cuestión que debe abordarse es la definición de qué eventos constituyen una violación del perímetro de seguridad.

3.2 Clasificación de los Sistema de Detección de Intrusiones

Esta sección analiza los fundamentos de la tecnología de detección de intrusiones de acuerdo a los siguientes criterios: a) método de implementación, b) mecanismo de detección y técnica de análisis de datos, y c) arquitectura, así como las medidas más usadas y clases de ataques.

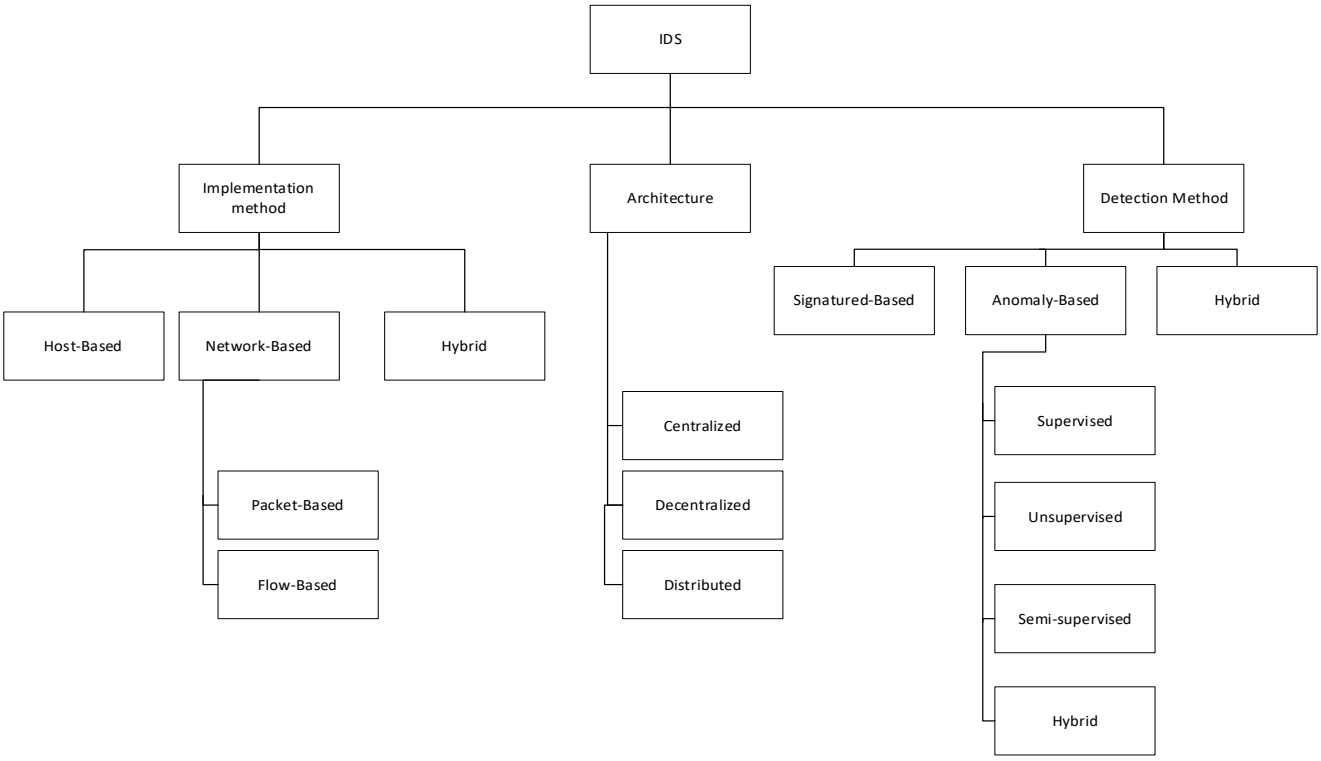


Ilustración 2: Clasificación de los IDS

3.2.1Métodos de Implementación

Con respecto al método de implementación, los IDS pueden ser dividido en dos categorías: a) basado en host y b) basado en red. Un sistema de detección de intrusiones basado en host (HIDS) despliega un agente local en cada host de la red. Un HIDS usa los agentes locales y los registros de la aplicación o las llamadas al sistema sin procesar como fuente de datos para detectar procesos fraudulentos, modificación de archivos de configuración críticos del sistema (claves de registro), privilegios escalados y cualquier otra acción no autorizada que esté en contra las políticas del sistema. Un HIDS tiene la ventaja de trabajar con datos de alta calidad que suelen ser muy informativos [17]. Sin embargo, procesar la pista de auditoría puede tener un impacto en el rendimiento del host, cuando los datos se procesan localmente [18], o en el ancho de banda de la red, cuando se utiliza una unidad de procesamiento remoto [19]. En [20], se redujo con éxito el costo computacional utilizando "información de sesión" en lugar de las pistas de auditoría tradicionales, también en [21], proponen un sistema basado en un modelo de Markov Oculto (Hidden Markov Model) con una etapa de preprocesamiento que elimina llamadas de sistema de subsecuencia similares. También en [22] y [23] proponen IDSs basados en host. Sin embargo, la protección ofrecida por HIDS no es suficiente, ya que esta restringida a un solo host. En comparación con un HIDS, un NIDS tiene las siguientes ventajas [17]:

1.Es más resistente a los ataques, ya que un HIDS depende de los registros producidos por el sistema y otras aplicaciones.
2.Es independiente del sistema operativo y la plataforma, lo que significa que el mismo NIDS funciona en cualquier plataforma sin necesidad de alguna modificación.
3.No afecta el rendimiento de la red, debido a que no agrega ninguna sobrecarga al tráfico de la red, simplemente realiza el monitoreo y procesamiento.

Un NIDS monitorea y analiza el tráfico de la red en un paquete o un nivel de flujo e intenta detectar anomalías, como accesos no autorizados o ataques DDoS. A nivel de paquete, un IDS realiza la denominada inspección profunda de paquetes (DPI), que analiza tanto el encabezado como la carga útil de cada paquete [17] , [24]. Aunque inspeccionar la carga útil de los paquetes puede ser solo informativo, con las redes de comunicación de alta velocidad actuales, este enfoque no solo requiere mucho tiempo, es ineficiente, y también computacionalmente costoso. Una de las características deseables de los IDS es el requisito

de tiempo real. Además, en el caso de paquetes cifrados, que se están volviendo frecuentes con el auge de darknet [25] y el uso de tecnologías como VPN, el análisis de la carga útil no es posible. Por el contrario, un NIDS basado en flujo, inspecciona solo los encabezados de los paquetes y usa datos de entrada en la forma de NetFlow o IPFIX [26] . Podría utilizarse una combinación de ambas técnicas para mejorar el rendimiento de un IDS. Por ejemplo, DPI podría aplicarse solo en los paquetes que fueron marcado como potencialmente malicioso por el IDS basado en flujo. Una de las principales desventajas del NIDS es la escalabilidad de la red, es decir, la capacidad del NIDS para ajustarse al tamaño y la complejidad de la red que está cambiando.

3.2.2 Mecanismo de Detección

Los IDS pueden clasificarse según su mecanismo de detección como [27]: 1) uso indebido o basado en firmas, 2) anomalías o basado en el comportamiento e 3) híbrido.

Los sistemas de uso indebido o basado en firmas mantienen una base de datos de firmas predefinidas (patrones) que corresponden a ataques conocidos y realizan la detección comparándolos con el flujo de datos de auditoría. Eso sigue siendo el método preferido en la industria actual, ya que tiene una tasa baja de falsos positivos y constituye una solución "fuera de caja”. Hay muchos IDS de código abierto basadas en firmas, que se utilizan ampliamente en el mundo empresarial, como SNORT [28], BRO [29] y Suricata [30]. Además, existen varias soluciones de seguridad que incluyen detección de intrusiones en la red proporcionadas por proveedores conocidos, como Unified Security Gestión (USM) [31] por AlienVault, Firepower NextGeneration IPS (NGIPS) [32] por CISCO y FireEye Network Seguridad [33] por FireEye.

A pesar de sus ventajas, un sistema de detección de intrusiones de uso indebido es tan bueno como su base de datos, que debe ser actualizada por un experto humano y por lo tanto no puede detectar ataques desconocidos [6]. En consecuencia, el rendimiento de tal sistema de detección de intrusiones está firmemente ligado a la calidad de su base de datos. Además, a medida que aumenta su tamaño, también lo hace el tiempo de procesamiento (tiempo de búsqueda) y costo. Según [34], el popular IDS de mal uso, como SNORT y BRO, consume una cantidad significativa de recursos (CPU y memoria) cuando operan en una red de alta velocidad. Además, en los actuales ecosistemas ciber físicos en constante cambio y evolución, aparecen diariamente nuevos ataques y variaciones más avanzadas de los antiguos. Como tal, mantener

un conjunto de reglas actualizado no es factible, así como es un proceso que requiere mucho tiempo y es insuficiente. En [35] se realizaron pruebas con cinco diferentes técnicas de evasión IDS, contra sistemas conocidos basados en mal uso como SNORT. Sus resultados mostraron que este tipo de IDS son vulnerables a ataques como mutación de carga útil y shellcode y variación simple de ataques más antiguos.

La detección de anomalías se refiere al problema de identificar instancias en un conjunto de datos que no se ajustan al comportamiento "normal". Un comportamiento se considera una anomalía o un valor atípico si la desviación "normal" excede un predefinido o dinámico umbral calculado. La detección encuentra una anomalía con un uso extensivo en una amplia variedad de aplicaciones, como la detección de fraudes para tarjetas de crédito, seguros o servicios de cuidado de la salud, detección de intrusiones para ciberseguridad, detección de fallas en sistemas críticos de seguridad y vigilancia militar para actividades de enemigos [36] . Es importante enfatizar que una anomalía no corresponde necesariamente a un ataque, sino a una observación sospechosa. Los IDS basados en anomalías no dependen de patrones, pero su objetivo es modelar el comportamiento / tráfico normal para detectar lo anormal, y así son capaces de detectar ataques conocidos y desconocidos. El precio por esto es la tasa alta de falsos positivos que producen y la etapa de sintonización que exigir. Hay muchos estudios en la literatura que apuntan a combinar las técnicas antes mencionadas para heredar sus dos ventajas, mejorando la tasa de detección y minimizando la tasa de falsos positivos. En [37], proponen uno de los sistemas híbridos más conocidos, denominados Análisis de Datos de Auditoría y Minería (ADAM). ADAM tiene dos etapas de detección: 1) utiliza la minería de reglas de asociación para la etapa basada en anomalías, 2) clasifica las conexiones sospechosas como normales, ataques conocidos y ataques desconocidos con un módulo de mal uso. En [38], combinan el árbol de decisiones C4.5 para el módulo de uso indebido con múltiples Máquinas de Vectores de Soporte (SVM) de una clase para modelar el comportamiento normal. De forma similar en [39], utilizan mapas autoorganizados (SOM) para el módulo de anomalías y árbol de decisión C4.5 para el módulo de mal uso. Otros sistemas híbridos son propuestos en [40] y [41].

3.2.3 Sistema de Detención de Intrusos Basados en Anomalías.

Los IDS basados en anomalías se pueden dividir en las siguientes categorías según el método que utilizan: estadístico, supervisado (clasificación), no supervisado (agrupamiento y valores

atípicos de detección), informática blanda, basada en el conocimiento y combinación de aprendices [42], [43]. En un IDS supervisado, un modelo está capacitado para aprender de ejemplos (datos etiquetados). Cuando se introduce una nueva instancia, un clasificador intenta asignarlo a una de las clases predefinidas. Varios algoritmos de clasificación, como el árbol de decisiones (C4.5), SVM, K-vecino más cercano, clasificador de Bayes, redes neuronales, etc., se han utilizado para tareas de detección de intrusiones basadas en red.

En [44], se muestra que los algoritmos supervisados exhiben una excelente precisión de detección y baja tasa de falsos positivos en la detección de ataques conocidos, logrando los mejores resultados con C4.5. Sin embargo, cuando se presentan ataques desconocidos en el conjunto de datos, la mayoría de los modelos de clasificación no los detectan, y SVM logra el mejor resultado. Además, los sistemas basados en clasificación tienen una desventaja similar a los basados en firmas, es decir, necesitan recibir formación periódica para preservar su alta tasa de detección. Esto no es factible ya que es extremadamente difícil obtener datos etiquetados, especialmente de forma regular. Además, incluso si existen datos etiquetados, es incierto si incluyen todos los nuevos ataques. En el pasado, se propusieron muchos modelos supervisados [45], [46].

La detección de anomalías no supervisada (detección basada en valores atípicos) utiliza técnicas de agrupación para identificar posibles instancias maliciosas en un conjunto de datos determinado, sin tener ningún conocimiento previo. El objetivo de la agrupación en clúster es separar un conjunto finito de datos sin etiquetar en un conjunto finito y discreto de "natural", estructuras de datos ocultos, en lugar de proporcionar una caracterización precisa de muestras no observadas generadas a partir de la misma distribución de probabilidad [47]. En otras palabras, los algoritmos de agrupación tienen como objetivo dividir los datos dados en grupos (clúster) que logran una alta disimilitud interior y exterior, sin ningún conocimiento previo. Para ello, todos los métodos de agrupación se basan en los siguientes supuestos. En primer lugar, el número de instancias normales en un conjunto de datos supera ampliamente en número el número de anomalías. En segundo lugar, las propias anomalías son cualitativamente diferentes de los casos normales [48]. Después de la formación de agrupaciones, las puntuaciones se asignan a la estructura construida de clúster. Si el puntaje de un clúster excede el predefinido o umbral calculado dinámicamente, se considera potencialmente malicioso. Respectivamente,

cuando se utiliza la agrupación en clústeres para detectar ataques a la red, se asume que a) el tráfico normal supera en número al malicioso y b) el tráfico normal se diferencia de alguna manera de lo malicioso.

Por estas razones, la selección del subconjunto de características adecuadas es de gran importancia. En otras palabras, uno tiene que seleccionar las características que describen lo suficientemente bien a los ataques para ser identificados. Con respecto al proceso de detección, el objetivo de la agrupación es agrupar los flujos o paquetes de la red sin ningún tipo de conocimiento, pero basado únicamente en las relaciones entre ellos. Como resultado, se crearán grandes grupos de tráfico normal mientras el tráfico malicioso formará clústeres más pequeños y valores atípicos, es decir, instancias que no pertenecen a ningún clúster. Basado en experimentos y calibración de los algoritmos en uso, un umbral dinámico o estático se puede utilizar para decidir qué grupos se consideran malicioso. La principal ventaja de los sistemas basados en clústeres es su capacidad para detectar ataques desconocidos sin ningún tipo de conocimiento, que elimina la necesidad de datos etiquetados. El principal inconveniente es la alta tasa de falsos positivos que producen.

Uno de los pasos más importantes de la detección de anomalías no supervisadas es la extracción o selección de características. Cada instancia, en un conjunto de datos está representada por una matriz de características, que se denominan características. La selección de funciones se refiere a el proceso de seleccionar un subconjunto de las características disponibles que son los más relevantes y menos redundantes. Por otra parte, la extracción de características tiene como objetivo crear (extraer) nuevas características de mejor calidad. Ambos procesos pueden afectar no solo la tasa de detección del sistema, sino también su rendimiento.

3.2.4 Arquitectura

La arquitectura de un IDS puede afectar su rendimiento, por lo que es una decisión importante durante el diseño del sistema. Esto es especialmente cierto debido a las redes de alta velocidad que la mayoría de las organizaciones utilizan hoy en día. Teniendo en cuenta la arquitectura del sistema, los IDS se puede dividir en las siguientes tres categorías:

1)Centralizado: los IDS centralizados constan de múltiples sensores en la red que monitorean y envían datos a la unidad central de procesamiento (CPU), donde el análisis de los datos recopilados y la detección se lleva a cabo. Esta arquitectura tiene dos desventajas principales. En primer lugar, no proporcionan escalabilidad de red, lo que significa que a medida que la red se expande, la CPU está sobrecargada y en algún punto puede volverse incapaz de mantenerse al día con la carga de trabajo. En segundo lugar, una CPU constituye un único punto de fallo (SPoF) del sistema [48].

2)Descentralizado: en esta arquitectura, varios sensores y múltiples unidades de procesamiento se encuentran dispersas por la red, siguiendo una estructura jerárquica. Los datos recopilados se envían a la unidad de procesamiento más cercana, donde se procesan previamente antes de que terminen en la unidad principal de procesamiento. De esta forma, se puede evitar el SPoF y los problemas de escalabilidad. El rendimiento del sistema también se ve reforzado debido a la etapa de preprocesamiento.

3)Distribuido: esta arquitectura consta de una superposición plana de múltiples agentes autónomos que actúan como sensores y unidades de procesamiento al mismo tiempo. Los datos se recopilan y procesan por medio de los agentes, que se comunican con otros a través de una arquitectura Peer-to-Peer (P2P) [45]. En esta arquitectura no hay unidad principal ni CPU y la carga de trabajo de procesamiento se distribuye entre todos los agentes, lo que aumenta el rendimiento y la escalabilidad del sistema. Tanto en arquitecturas descentralizadas como distribuidas, la comunicación entre los agentes es crucial para la detección de ciertos tipos de ataques. Por ejemplo, la pérdida de comunicación entre los agentes puede conducir a la incapacidad del sistema para detectar ataques distribuidos.

3.2.5 Medidas

En el pasado, se han utilizado diferentes métricas y conjuntos de datos para medir qué tan bueno es un sistema para identificar con éxito ataques y tráfico normal en un conjunto de datos, lo que dificulta la comparación de los resultados de los distintos sistemas propuestos. Las medidas de evaluación más habituales en cuanto a la detección y capacidad del sistema son las siguientes:

1)La matriz de confusión, también conocida como matriz de error, es una forma de visualizar la relación entre los resultados reales y los resultados previstos. Se utiliza principalmente en el aprendizaje supervisado para evaluar la precisión de la predicción de un clasificador. Cada fila de la tabla corresponde a un resultado predicho por el clasificador, mientras que cada columna corresponde a un resultado real.

2)Recall representa la parte de las instancias relevantes (verdaderos positivos) que se recuperan con éxito. Por el contrario, la precisión es la proporción de instancias recuperadas que están correctamente identificadas. Tanto “recall” como precisión se centran en las muestras positivas, pero ninguno de ellos captura que tan bien el modelo maneja los casos negativos. [48]. La media armónica de las dos mediadas anteriores se llama medida F (F1). Aunque F1 es defendido como una única medida para capturar la eficacia de un sistema, todavía ignora por completo los verdaderos valores negativos (TN) [49].

3)La precisión tiene en cuenta tanto los verdaderos positivos como los negativos y se define como la proporción de las muestras clasificadas al número total de instancias.

4)Sensibilidad, también conocida como tasa de positivos verdaderos (TPR), es la proporción de muestras positivas que se clasifican correctamente como tales. Por el contrario, la especificidad o tasa negativa verdadera (TNR) mide la proporción de instancias que se clasifican correctamente como negativas. De manera similar, la tasa de falsos positivos (FPR) representa la proporción de muestra que se identifica incorrectamente como anomalías.

5)La característica de funcionamiento del receptor (ROC) es una técnica utilizada originalmente en la teoría del procesamiento de señales para visualizar el TPR contra el FPR para diferenciar la configuración de parámetros. Representa las compensaciones relativas entre beneficios (verdaderos positivos) y costos (falsos positivos) [50]. Aunque la precisión de un IDS es uno de los requisitos más importantes, no es el único. El tiempo de respuesta del sistema es un factor significativo, ya que se utilizará en redes empresariales donde incluso una pequeña latencia puede resultar en pérdidas monetarias para una organización. Además, el costo computacional y de comunicación (entre los agentes y las unidades de procesamiento), no solo puede afectar negativamente la respuesta de tiempo, sino también el costo financiero de implementar

y mantener el sistema. Como las redes actuales son grandes y su tamaño dinámico y variable, el IDS debe poseer la capacidad de ajustarse a los cambios de tamaño y estructura de la red. Finalmente, un sistema destinado a proteger otros sistemas debe ser en sí mismo resistente a cualquier ataque que tenga como objetivo interrumpir su funcionamiento y tener un rendimiento estable y consistente bajo diferentes escenarios.

3.2.6 Clases de Ataques

Se han propuesto cuatro categorías principales de ataques en la literatura de detección de intrusiones, que un IDS necesita ser capaz de detectar:

1)Denegación de servicio (DoS): en un ataque DoS, el sistema objetivo está inundado con una gran cantidad de solicitudes que se originan en una sola conexión, hasta que todos los recursos están agotados y, por lo tanto, ya no es capaz de manejar solicitudes legítimas. En una denegación de servicios distribuidos (DDoS) el atacante está utilizando múltiples conexiones que se distribuyen a través de Internet y es probable que formen parte de una red de botnets. Ataques de este tipo, tienen como objetivo la disponibilidad de una infraestructura haciendo un servicio o recurso no disponible para sus usuarios. Los atacantes suelen utilizar ataques DoS y DDoS medianos como cortinas de humo para ocultar actividades maliciosas peligrosas o para eliminar la seguridad de dispositivos, como firewalls.

2)Sonda: este tipo de ataque (escaneo de puertos) se utiliza para explorar la red destino y recopilar información sobre los hosts, como puertos abiertos, servicios en ejecución, etc.

3)Usuario a Root (U2R): En este caso, el atacante ya ha conseguido acceso local al sistema objetivo y tiene como objetivo explotar una vulnerabilidad del sistema para escalar sus privilegios de los de un usuario simple a superusuario / administrador. Uno de los tipos de U2R más comunes es el desbordamiento de búfer, en el que el atacante intenta sobrellenar un búfer y ejecutar código malicioso bajo privilegios de root.

4)Remoto a local (R2L): en esta clase de ataque, el atacante no tiene una cuenta en la máquina destino e intenta obtener acceso local. Los ataques remotos a locales son generalmente combinados con ataques U2R. Un ejemplo de un ataque R2L es fuerza bruta SSH.

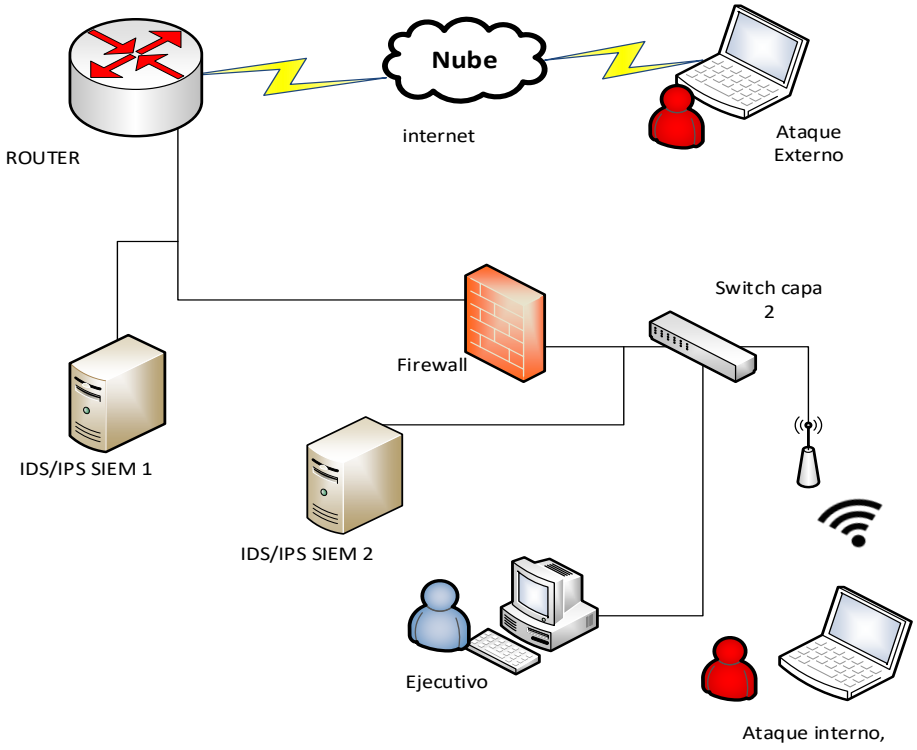


Ilustración 3: Red IDS

3.3 Sistema de Gestión de Eventos e Información de Seguridad.

Los riesgos de ciberseguridad que afectan a las tecnologías de información y comunicaciones (TIC) han aumentado enormemente durante los últimos años. Los atacantes se han vuelto más sofisticados y peligrosos y su detección adecuada y oportuna se ha convertido en un verdadero desafío. Algunos ejemplos de incidentes actuales de ciberseguridad que afectan a las TI y las TIC son [51]: ataques de ransomware; malware que tiene un impacto en la capacidad de la empresa de servicios públicos para realizar negocios y operaciones; campañas de phishing dirigidas a ejecutivos, asistentes ejecutivos, ingenieros SCADA, administradores de TI u otros usuarios privilegiados; incidentes de compromiso de correo electrónico comercial, incluida la toma de control de cuentas o la suplantación de ejecutivos; fuga y robo de datos; ingeniería social para recopilar información confidencial del personal.

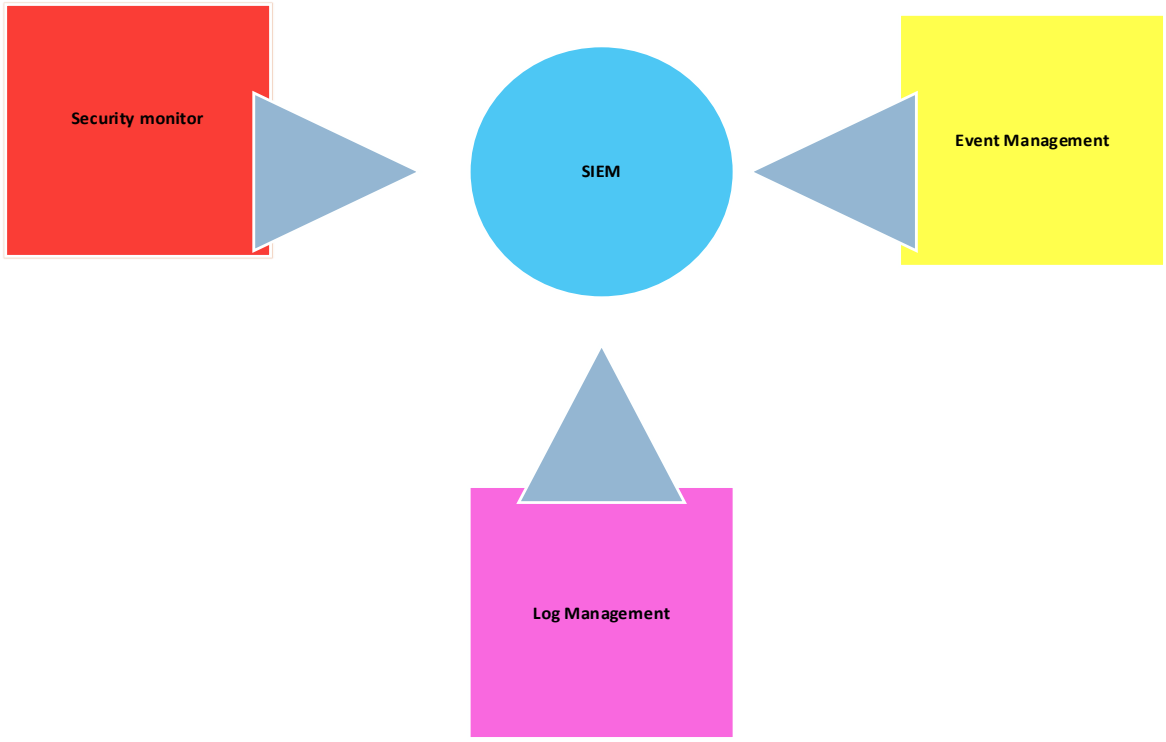


Ilustración 4 : SIEMS.

Los sistemas de gestión de eventos e información de seguridad (SIEM) se han desarrollado como respuesta para ayudar a los administradores a diseñar políticas de seguridad y gestionar eventos de diferentes fuentes. Generalmente, un SIEM simple se compone de bloques separados que pueden funcionar independientemente entre sí, pero sin que todos trabajen juntos, el SIEM no funcionará correctamente [6]. La Figura 5 muestra los componentes básicos de una solución SIEM regular.

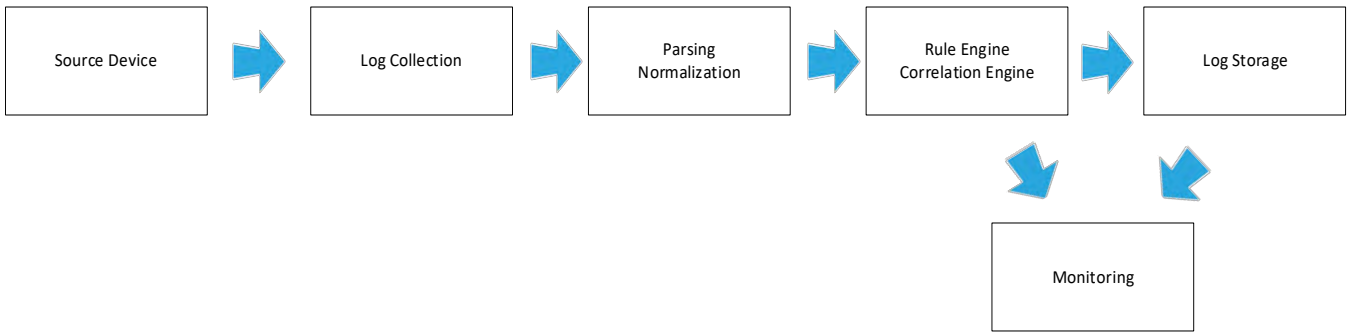


Ilustración 5: Componentes SIEM

Las plataformas SIEM proporcionan análisis en tiempo real de eventos de seguridad generados por aplicaciones y dispositivos de red. Además, aunque la nueva generación de SIEM proporciona capacidades de respuesta para automatizar el proceso de selección y despliegue de contramedidas, los sistemas de respuesta actuales seleccionan e implementan medidas de seguridad sin realizar un análisis de impacto integral de los ataques y escenarios de respuesta. Además de estas características comunes, los SIEM actuales presentan diferencias que los clasifican como líderes, retadores, jugadores de nicho o visionarios, según el informe anual SIEM Magic Quadrant de Gartner.

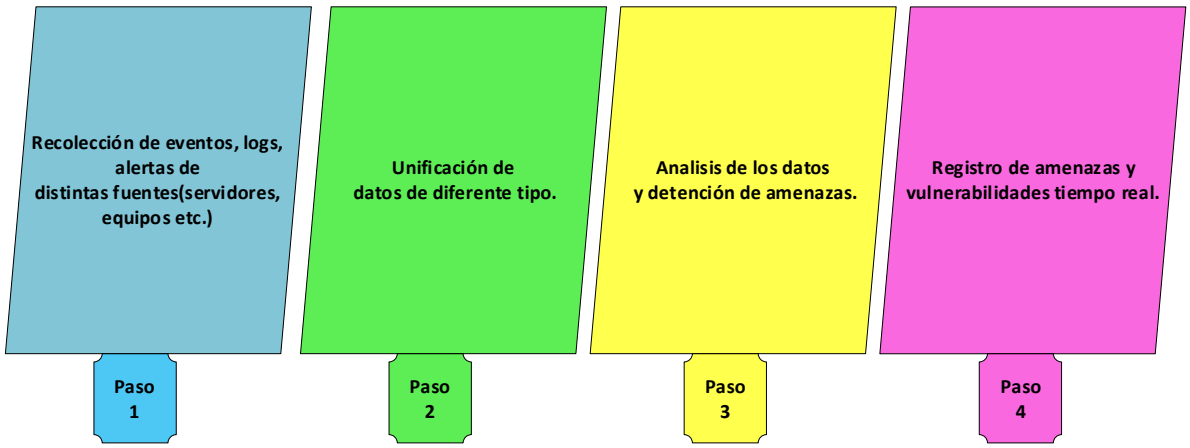


Ilustración 6: Funcionalidad SIEM

3.4 Clasificación de un Sistema SIEM

El análisis y la evaluación de los sistemas de seguridad se han propuesto ampliamente en la literatura. Mientras que algunas investigaciones se centran en los aspectos comerciales, otras se concentran en las características técnicas que podrían mejorarse en las soluciones SIEM actuales. Instituciones reconocidas como Gartner [52], por ejemplo, proponen un análisis comercial de SIEM Systems basado en el mercado y los principales proveedores, para lo cual se publica un informe anual para posicionar a los proveedores de SIEM como líderes del mercado, retadores, jugadores de nicho, o visionarios. Otras instituciones de seguridad (por ejemplo, e Info-Tech Research Group, han informó ampliamente sobre las capacidades de las soluciones SIEM y sobre la forma en que se pueden comparar y evaluar los proveedores de SIEM. Techtarget, por un lado, publica guías electrónicas periódicas sobre la seguridad de los

sistemas SIEM y cómo definir la estrategia, la gestión y el éxito de SIEM en la empresa. Info-Tech, por otro lado, proporciona informes técnicos sobre el panorama de proveedores de SIEM [53] centrados en los beneficios y desventajas de los principales SIEM comerciales. Ambas organizaciones toman el Cuadrante Mágico de Gartner como base para su análisis.

Durante la última década, Gartner ha clasificado las soluciones SIEM como líderes (organizaciones que se ejecutan bien frente a su visión actual y están bien posicionadas para el mañana), visionarias (organizaciones que entienden hacia dónde se dirige el mercado o tienen una visión para cambiar las reglas del mercado, pero no lo hacen). aún no se ejecutan bien), actores de nicho (organizaciones que se enfocan con éxito en un segmento pequeño o que no están enfocadas y no innovan ni superan a otros) y desafiantes (organizaciones que se desempeñan bien hoy o pueden dominar un segmento grande, pero no demuestran comprensión de la dirección del mercado).

La Tabla 6 muestra la evolución de las soluciones SIEM (y los proveedores de SIEM) de 2010 a 2020. De la Tabla 6, un símbolo de asterisco (*) indica aquellos que han estado liderando el mercado, los retadores se identifican con (**), los jugadores de nicho se identifican con (***) y los visionarios se identifican con(****). Es importante destacar que muy pocos de ellos aparecieron cada año en la evaluación de la clasificación más alta durante todo el período de la década. Este es el caso de RSA, la división de seguridad de EMC Corporation (Dell Technologies), que ofrece un SIEM evolucionado de NetWitness Platform; IBM, que ofrece una herramienta llamada Qradar; NetIQ / Microfocus / ArcSight, que ofrece ArcSight Enterprise Security Manager; McAfee / Intel, que ofrece McAfee Enterprise Security Manager y LogRhythm, que ofrece la plataforma Nextgen SIEM. Tenga en cuenta que algunas soluciones se han fusionado para adaptarse a los cambios y la evolución del mercado. Este es el caso de IBM y Q1 Labs (que ofrecieron una solución conjunta en 2012 y 2013); NetIQ y Novell (2012); HP y ArcSight (2013); AccelOps y Fortinet (2016), y más recientemente, Micro Focus y ArcSight, así como Micro Focus y NetIQ (2017). Algunas de estas soluciones unidas ya no estarán disponibles en 2021. Otro aspecto importante a tener en cuenta es que algunos proveedores de SIEM han estado en la lista de clasificación de Gartner superior desde que aparecieron por primera vez en el mercado (por ejemplo, Splunk, AlientVault / AT & T Cybersecurity, SolarWinds, EventTracker, Fortinet, MicroFocus). Algunos otros se han unido a la lista en los últimos años

con funciones de Análisis de comportamiento de entidades y usuarios (UEBA) (es decir, Manage Engine, Venustech, Rapid7, Exabeam, Secureonix, LogPoint y HanSight).

Un estudio reciente [53] considera 22 jugadores en el mapa de proveedores de SIEM 2020 basado en tres capacidades principales: (i) detección de inteligencia de amenazas, (ii) cumplimiento y (iii) gestión de registros. Además de la inteligencia de amenazas, el cumplimiento y la gestión de registros, los desarrolladores de SIEM están considerando las capacidades de UEBA y los paneles inteligentes como innovaciones que se agregarán a sus soluciones. Como resultado, los nuevos sistemas SIEM ayudarán a los administradores de seguridad con cuadros de mando preconstruidos, informes, flujos de trabajo de respuesta a incidentes, análisis avanzados, búsquedas de correlación e indicadores de seguridad. Además, un análisis en profundidad de la extensibilidad de SIEMs [54] reveló que las soluciones SIEM actuales necesitan mejorar características como análisis de comportamiento, análisis de riesgo e implementación, visualización, almacenamiento de datos y capacidades de reacción, para mantenerse al día con el mercado.

Tabla 6: Clasificación SIEM

SIEm Vendor	2010	2011	2012	2013	2014	2015	2016	2017	2018	2020
HP/ArcSigth/HPE	*	*	*	*	*	*	*	*		
RSA/EMC	*	*	**	**	**	**	**	**	*	*
SenSAGE	*	****	***	***						
LogLogic	*	*	**							
Symantec	*	*	**	**						
Q1Labs	*	*	*	*						
Novell	*	*	*							
IBM	**	**	*	*	*	*	*	*	*	*
QUEST Software	**	**								
CA	**									
Tenable	**	****	**	**	**					
Prism MicroSystems	**	****	**							
LogMatrix	**									

NetIQ /Microfocus	****	**	*	***	***	**	**	**	**	**
McAfee/Intel	****	*	*	*	*	*	*	*	*	*
Trustwave	****	****	****	***	***	***	***	***		
LogRhythm	****	****	*	*	*	*	*	*	*	*
TriGeo	****	****								
netForensics	****	****								
elQnetworks	****	****	****	****	***					
Splunk	***	**	*	*	*	*	*	*	*	*
Tripwire	***									
AlientVault/At&t	***	****	****	****	****	****	***	***	***	
Cibersecurity										
Correlog	***	***								
521sec	***	***								
Tango/04	***	***								
Tier-3	****	****								
SolarWinds		****	**	***	***	***	***	***	***	**
Tibco-LogLogic			****	***						
EventTracker			***	***	***	***	***	***	***	***
AccelOPS/fortinet					***	***	***	***	***	***
Blackstratus					***	***	***	***	***	***
Manage Engine							***	***	***	***
FireEye								***		***
Venustech								***	***	
Rapid7								****	****	*
Exabeam								****	*	*
Securonix								****	*	*
LogPoint								***		****
HanSight										***

* Leader ** Challenger *** Niche Player **** Visionary. SIEM [55]- [54]

3.5 Herramientas de un Sistema SIEM

Teniendo en cuenta la información anterior sobre SIEM, la Tabla 7 resume algunos de los SIEM más prometedores hasta la fecha.

Tabla 7: Mercado SIEMS

ArcSight Enterprise Security Manager (MicroFocus/ HPE/ NetIQ)	Proporciona una interfaz gráfica para el equipo del Centro de operaciones de seguridad (SOC) y un conjunto de aplicaciones o comandos externos que ayudan a los procesos de correlación y / o investigación.	Opciones de visualización limitadas e intrincadas reglas de correlación. La información asociada a los eventos es inmutable, con evidentes déficits a la hora de adaptar el producto a los procesos y necesidades de la empresa.
Qradar (IBM)	Se puede implementar como hardware, software o dispositivo virtual, así como como software como servicio (SaaS) en la nube de IBM. Proporciona una interfaz de usuario para visualización y eventos en tiempo real, informes, infracciones, información de activos y gestión de productos. Ofrece soporte para fuentes de inteligencia de amenazas.	Proporciona capacidades de reacción básicas que incluyen funciones de notificación y alerta. El monitoreo de terminales para la detección y respuesta de amenazas, o la integridad básica de archivos, requiere el uso de tecnologías de terceros.
McAfee Enterprise Security Manager (McAfee/ Intel)	Permite una arquitectura SIEM escalable y versátil, que ofrece análisis forense en tiempo real, monitoreo integral del tráfico / contenido de aplicaciones y bases de datos, reglas avanzadas y correlación basada en riesgos para detección de incidentes históricos y en tiempo real y reacción automática.	Requiere el uso de soluciones adicionales (por ejemplo, McAfee Active Response). El análisis predictivo y otras funciones integradas, como el análisis de comportamiento, están poco desarrolladas.
LogRhythm Next GEN SIEM Platform (LogRhythm)	Proporciona monitoreo de punto final, análisis forense de redes, análisis de comportamiento de usuarios y entidades y capacidades de respuesta. Se puede implementar en un dispositivo, software o instancia virtual que admita arquitecturas descentralizadas escalables.	No es adecuado para organizaciones con infraestructuras críticas, aunque se pueden implementar extensiones para mejorar las capacidades de SIEM. Requiere un alto grado de automatización y contenido listo para usar.
USM and OSSIM (AT&T Cybersecurity/ AlienVault)	Ofrece tanto soluciones comerciales (es decir, Alienvault Unified Security	Análisis limitado del comportamiento de usuarios o entidades, así como

	Management-USM) como soluciones SIEM de código abierto. (es decir, OSSIM). Incluye un Interfaz gráfica basada en web para administración, informes y gestión de eventos de seguridad.	capacidades de aprendizaje automático. Capacidades de reacción básicas (por ejemplo, enviar correo electrónico, ejecutar script, abrir ticket) y limitadas al conjunto predefinido de condiciones asociadas a una política de seguridad.
RSA Netwitness Platform (Dell)	Analiza los datos y el comportamiento de las personas y los procesos dentro de una red en los registros, paquetes y puntos finales de una empresa. Secentra en la detección avanzada de amenazas. Proporciona sólidas capacidades de monitoreo de OT.	Requiere una amplia comprensión de la amplitud de las opciones y las implicaciones de costo, funcionalidad y escalabilidad.
Splunk Enterprise Security (Splunk)	Plataforma líder en el mercado de Inteligencia Operativa. Ofrece capacidades de recopilación, indexación y visualización de datos para el monitoreo de eventos de seguridad. Utiliza análisis de seguridad avanzados, que incluyen capacidades de comportamiento del usuario y aprendizaje automático no supervisado.	Utiliza reglas de correlación predefinidas básicas para los requisitos de supervisión y generación de informes. Las capacidades de reacción se limitan a las notificaciones por correo electrónico. Requiere integración con aplicaciones de terceros para la automatización de tareas y flujos de trabajo.
SolarWinds Log and event Manager (SolarWinds)	Proporciona una recopilación de registros centralizada y normalización, amenaza automatizada detección y respuesta, intuitivo visualización e interfaz de usuario, así como correlación en tiempo real y búsqueda de registros para apoyar la investigación.	Carece de apoyo para monitorear al público IaaS o SaaS de los servicios en la nube. No apoyar la redacción de informes personalizados y personalización de fuera de la caja plantillas de informes de cumplimiento.

3.6 Funciones y Capacidades de un Sistema SIEM

Básicamente, todos los SIEM tienen la capacidad de recopilar, almacenar y correlacionar eventos generados por una infraestructura administrada. Además de estas capacidades clave, existen muchas diferencias entre los sistemas existentes que normalmente reflejan las diferentes posiciones de los SIEM en el mercado.

Reglas de correlación: el éxito de la detección de un evento mediante un SIEM depende del poder de las reglas de correlación. Si bien la mayoría de los SIEM poseen reglas de correlación básicas, pocos de ellos tienen capacidades de búsqueda sólidas y admiten lenguajes de procesamiento de búsqueda para escribir búsquedas complejas que se pueden utilizar en los datos del SIEM.

Fuentes de datos: una de las características clave de un sistema SIEM es la capacidad de recopilar eventos de múltiples y diversas fuentes de datos en la infraestructura administrada. La mayoría de los SIEM admiten varios tipos de fuentes de datos de forma nativa, incluidos los sensores admitidos y los tipos de datos admitidos (por ejemplo, inteligencia de amenazas). Para otras soluciones (por ejemplo, QRadar, USM), esta característica podría ser compatible con componentes adicionales integrados al SIEM. Esta función evalúa las fuentes de datos compatibles de forma nativa y la posibilidad de que un SIEM las personalice automáticamente.

Procesamiento en tiempo real: esta característica considera la capacidad de un SIEM para manejar datos en tiempo real bajo cambios constantes. Evalúa las capacidades de control, monitoreo y canalización en tiempo real implementadas por la herramienta para prevenir o reaccionar ante incidentes de ciberseguridad, así como las capacidades de cálculo de rendimiento que tienen los SIEM para analizar millones de eventos en tiempo real. Todos los SIEM estudiados tienen capacidades avanzadas de procesamiento en tiempo real. Volumen de datos: analizar grandes volúmenes de datos provenientes de diferentes fuentes es importante para obtener más información de los eventos recopilados y tener un mejor monitoreo. Sin embargo, mantener grandes volúmenes de datos recopilados en un sistema SIEM activo a menudo es costoso y poco práctico. Esta función evalúa la posibilidad de que los sistemas actuales admitan grandes volúmenes de datos para operaciones de correlación, indexación y almacenamiento.

Visualización: Uno de los factores clave que dificultan el análisis de los eventos de seguridad es la falta de soporte para los métodos adecuados de visualización de datos y el poco soporte proporcionado para la exploración interactiva de los datos recopilados. Por lo tanto, es importante comprender las capacidades de los sistemas analizados en términos de creación de nuevos métodos de visualización de datos y cuadros de mando personalizados.

Análisis de datos: las versiones más recientes de los SIEM líderes admiten una amplia integración con detectores de anomalías basados en aplicaciones y usuarios. Estas capacidades incluyen el análisis del comportamiento de los empleados, terceros contratistas y otros colaboradores de la organización. Para ello, el SIEM debe comprender la gestión de perfiles de usuario / aplicación y el uso de técnicas de aprendizaje automático para la detección de malas conductas.

Rendimiento: esta función evalúa el rendimiento de una solución SIEM en términos de capacidad computacional, capacidades de almacenamiento de datos (por ejemplo, lectura / escritura), procesamiento de correlación de reglas (por ejemplo, motor de correlación de alto rendimiento), así como búsqueda de datos, índice y monitoreo.

Forense: además de las capacidades de registro, algunos SIEM (p. Ej., ArcSight, LogRhythm) ofrecen capacidades forenses de red integradas que incluyen capturas de paquetes de sesión completa de conexiones de red consideradas maliciosas con el objetivo de convertir paquetes de datos en documentos, páginas web, voz sobre IP y otros archivos reconocibles. Algunos otros productos (p. Ej., QRadar, Splunk) pueden guardar paquetes individuales de interés cuando lo solicita un analista de seguridad, pero no guardan automáticamente las sesiones de red de interés, el resto de las soluciones estudiadas no tienen una red incorporada. capacidades forenses.

Complejidad: los SIEM son conocidos por ser difíciles de implementar y administrar. Sin embargo, es importante comprender si el sistema analizado se puede instalar para realizar pruebas con un esfuerzo bajo o moderado. Por ejemplo, ArcSight es la herramienta con mayor complejidad para la implementación y la administración, mientras que LogRhythm y Splunk se consideran herramientas fáciles y amigables de instalar, implementar y usar.

Escalabilidad: esta característica considera la capacidad de una implementación SIEM para crecer no solo en términos de hardware, sino también en términos de la cantidad de eventos de seguridad recopilados en el borde de la infraestructura SIEM. La nueva transformación digital conduce a más sensores y más dispositivos (por ejemplo, servidores, agentes, nodos) conectados a la misma red.

Análisis de riesgos: las versiones recientes de los sistemas SIEM líderes (por ejemplo, QRadar, LogRhythm, Splunk) incluyen funciones para realizar análisis de riesgos en los activos de la infraestructura administrada. Esta función evalúa si el SIEM admite de forma nativa el análisis de riesgos o si se puede integrar con dispositivos externos para ese propósito.

Almacenamiento: Teniendo en cuenta que los SIEM generalmente almacenan información durante no más de 90 días, esta función evalúa el tiempo durante el cual las tecnologías SIEM actuales mantienen los datos almacenados en sus sistemas para su procesamiento posterior y operaciones forenses.

Precio: esta función evalúa el método de licencia asociado a la solución SIEM (por ejemplo, empresarial, gratuita, beta, premium) y los límites en el número de usuarios, consultas, volúmenes de índices, alertas, correlaciones, informes, paneles y acciones correctivas automatizadas. La mayoría de las soluciones estudiadas son muy caras, a excepción de LogRhythm, USM y SolarWinds, con costos más razonables y la posibilidad de utilizar soluciones de código abierto con capacidades más limitadas.

Resiliencia: la resiliencia o tolerancia a fallas es una característica importante de cualquier sistema de monitoreo crítico. Es importante comprender cuáles son las capacidades de tolerancia a fallas de los SIEM existentes, por ejemplo, si el motor de correlación admite la tolerancia a fallas; la forma en que la recuperación de desastres y la replicación son compatibles con el almacenamiento de eventos; si los conectores admiten funciones de alta disponibilidad.

Capacidades de reacción y generación de informes: esta función estudia las acciones que son compatibles de forma nativa con el SIEM para reaccionar frente a incidentes de seguridad (incluidas las capacidades para compartir e informar) y la forma en que dichas acciones se expresan en el motor de correlación.

UEBA: esta función evalúa si la solución SIEM presenta la capacidad nativa de análisis de comportamiento de entidades y usuarios (UEBA), o si proporciona integración con soluciones UEBA de terceros.

Seguridad: esta función evalúa la capacidad de implementar la automatización de la seguridad, así como las capacidades de cifrado nativas presentes en el SIEM durante el monitoreo, detección, correlación, análisis y presentación de los resultados.

3.7 ELK Stack

ELK stack es un paquete de tres aplicaciones. Elasticsearch, Logstash también, Kibana [56]. La Ilustración 7 muestra el paquete de Elastic Stack. A pesar de que Logstash y Elasticsearch pueden funcionar de forma independiente, los tres elementos están pensados para ser utilizados como un arreglo incorporado, al que en la actualidad se alude como Elastic Stack. El Elastic Stack se puede enviar en las instalaciones o, por otro lado, se puede utilizar como software como servicio (SaaS) proporcionado por una organización externa en la nube [2]. La pila se mantiene actualmente y se mantiene de manera efectiva, bajo un permiso de código abierto, por la organización llamada Elastic.

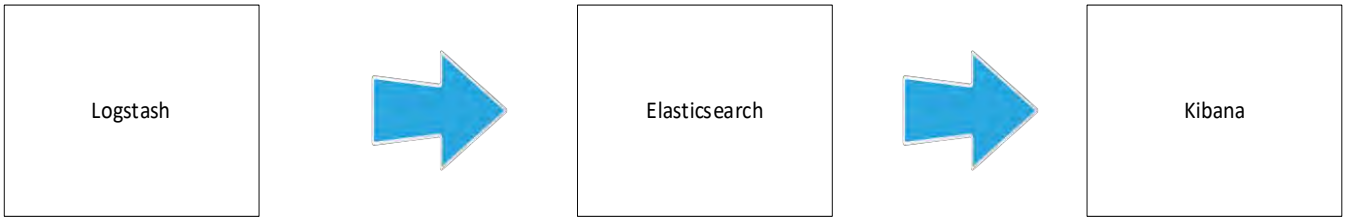


Ilustración 7: Stack EIK

3.7.1 Elasticsearch

Elasticsearch es una herramienta de búsqueda de código abierto excepcionalmente adaptable. Permite al usuario mantener e investigar un volumen extraordinario de todo tipo de datos para todos los efectos de forma progresiva. Elasticsearch trabaja con documentos de registros JSON. Utilizando una estructura interna, puede analizar Logstash ElasticSearch Kibana la información prácticamente en curso para buscar los datos necesarios [57]. Los conceptos principales de Elastic Search son Indexación y Mapeo.

i. Indexación

Es un conjunto de varios tipos de informes y propiedades de archivo de registro.También utiliza la idea de fragmentos para mejorar la ejecución Un grupo de ElasticSearch puede contener numerosos índices (bases de datos), que por lo tanto contienen diferentes tipos (tablas). Estos tipos contienen numerosos Documentos (líneas) y cada registro tiene Propiedades (secciones).

ii. Mapeo

Define los campos para documentos de un tipo específico, el tipo de datos (como cadena y número entero) y cómo los campos deben indexarse y almacenarse en Elasticsearch. Un mapeo puede caracterizarse inequívocamente o producirse en consecuencia cuando se ordena un informe utilizando diseños.

3.7.2 Logstash

Logstash es una herramienta basada en los patrones de filtro / canalizaciones para recopilar, procesar y generar los registros o eventos. Ayuda a centralizar y realizar análisis en tiempo real de registros y eventos de diferentes fuentes. Logstash está compuesto en el lenguaje de programación JRuby que sigue ejecutándose en la JVM, de ahora en adelante podemos ejecutar Logstash en varias plataformas. Maneje una amplia gama de información de registro Consiga de manera efectiva una gran cantidad de registros web como Apache y registros de aplicaciones como log4j para Java. Captura muchas otras posiciones de registro como syslog, administración de sistemas y registros de firewall. Es un túnel de tres etapas, entrada-> filtro-> salida.

I. Entrada (Input):

Puede manejar datos y tipos de entrada heterogéneos, manejará el siguiente tipo de entrada:

- Archivo
- Protocolo de red (tcp)
- Multilínea (JSON) Entrada estándar

II. Filtrar:

Los filtros de las medidas de Logstash manipulan y crear eventos como Apache Access. Muchos complementos de filtro utilizado para gestionar los eventos en Logstash.

- asimilar: analizar y estructurar contenido autoafirmante .asimilar es a partir de ahora la ruta más ideal en Logstash para analizar información de registro no estructurada en algo organizado y apto para consultas.

- Mutar: realiza cambios generales en instancias de los campos. Podemos renombrar, evacuar, suplantar y alterar campos en todos los eventos.

III. Producción (Output):

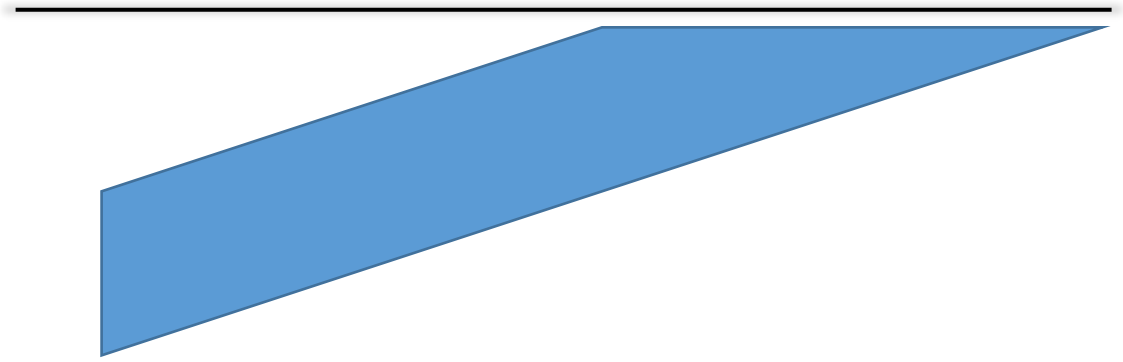
Logstash puede almacenar los registros filtrados en un archivo, Elasticsearch Engine, stdout, AWS CloudWatch y pronto. Protocolos de red como TCP, UDP, Websocket también se puede utilizar en Logstash para intercambiar las ocasiones de registro a los marcos de almacenamiento remoto.

3.7.3 Kibana

Kibana es una plataforma de examen y percepción de código abierto destinada a funcionar con Elasticsearch. Utilice Kibana para explorar, ver y conectarse con la información almacenada en los registros de Elasticsearch. Sin duda, podemos realizar un examen de información impulsado y representar su información en una variedad de esquemas, tablas y mapas. Kibana deja en claro grandes volúmenes de información. Su sencilla interfaz basada en programas le permite crear y compartir rápidamente cuadros de mando dinámicos que muestran cambios en las preguntas de Elasticsearch de forma continua.

La sección de gestión de Kibana está dividida en 4 segmentos principales: Descubrir, Visualizar, Paneles y Gestión [58]. Es donde se deben establecer los diseños de listas. En el caso de que los ejemplos de la lista contengan ocasiones de base de tiempo, también es necesario determinar la base de campo de marca de tiempo en la que Kibana ordenará y canalizará la información. Kibana, basado en un conjunto de listas que satisfacen el diseño de lista elegido, demuestra el resumen de los campos de archivo junto con su clasificación y propiedades. Además, es posible ajustar los formateadores de campo para modificar cómo se muestra la estima de campo en la GUI de Kibana.

CAPITULO 4



4 ELECCIÓN DE LA LÍNEA DE DEFENSA

El presente proyecto se desarrolló mediante software libre, en la parte de análisis de red y detección de intrusiones se utilizó Suricata, ya que es uno de los sistemas más completos que existen.

Suricata trabaja mediante reglas y alertas preconfiguradas o personalizadas, crea archivos log, eve.log entre otros, para el análisis de información y detección de anomalías. Sin embargo, la información generada por Suricata no es fácil de interpretar para la detección de ataques a la red, por tal motivo se propone el uso del Sistema de Gestión de Eventos e Información de Seguridad ELASTIC Stack (conjunto de software informático formado por Elasticsearch, Logstash, Kibana). El SIEM propuesto permite cargar los archivos tipo log, pcap, eve, etc; leer e interpretar la información de dichos archivos y visualizarla de forma gráfica en tiempo real [59], [60].

Para la instalación, del IDS y el SIEM es posible utilizar alguno de los siguientes sistemas operativos: Linux, Frebsd, Openbsd, Macosx o Windows.

Ilustración 8: Sistemas Operativos

4.1 Sistema Operativo

Como se menciona anteriormente, los sistemas elegidos en el desarrollo del presente proyecto son: para la detección de intrusiones se utilizó Suricata y como sistema de gestión de eventos e información de seguridad se utilizó el stack de Elasticsearch. Por tal motivo hemos elegido el sistema operativo Ubuntu (GNU/Linux) [61]. Ubuntu utiliza una versión con entorno gráfico para todo tipo de usuario, desde trabajos escolares, oficina o empresarial, así como su versión dedicada a servicios de red para el entorno empresarial; se consideró los requerimientos mínimos de hardware para su ejecución:

- Procesador de doble núcleo 2 GHz o superior.
- 2 GB Memoria Ram (para el caso de operación de suricata y ELK, requerimos un mínimo de 8 gb de ram).
- 25 GB de disco duro en espacio libre
- Acceso internet

4.2 Instalación del Software Libre

Suricata es el principal motor independiente de detección de amenazas de código abierto. Al combinar la detección de intrusiones (IDS), la prevención de intrusiones (IPS), la supervisión de la seguridad de la red (NSM) y el procesamiento PCAP, Suricata puede identificar, detener y evaluar rápidamente los ataques más sofisticados. Para su implementación en Ubuntu 20.04, se usó la terminal con la instrucción “apt”, como indica [61] y las líneas de comando para su instalación como se indica en [60], [62].

4.2.1 Suricata IDS

Para la configuración e instalación de la última versión estable de Suricata se realizó lo siguiente:

- #sudo apt-get.
- # sudo add-apt-repository ppa: oisf/suricata-stable.
- # sudo apt-get update.
- #sudo apt-get install suricata.

Para probar que efectivamente se instaló en el sistema operativo, se consultó la siguiente ruta de instalación:

#cd/etc/suricata.

Ahí se encuentran los archivos de configuración principal: # suricata.yaml.

4.3 Elasticsearch

Para la instalación de Elasticsearch, se utilizó el terminal con los paquetes binarios preconfigurados. Antes de iniciar se verifico que se contará con el Java Developer Kit (JDK), ya que el SIEMS elegido trabaja bajo la plataforma de Java.

Se aplicaron las siguientes líneas:

- #apt-get update.
- # java -version (si no se cuenta con java, aplicar otro comando).
- sudo apt install default-jre.

Una vez instalado JDK, se descargó e importó la clave de GPG publica de Elasticsearch en la terminal y luego agregaron las fuentes en el archivo source.list [63]:

- wget -qO - https://artifacts.elastic.co/GPG-KEY-elasticsearch | sudo apt-key add –.
- echo "deb https://artifacts.elastic.co/packages/6.x/apt stable main" | sudo tee -a /etc/apt/sources.list.d/elastic-6.x.list.

Se actualizó con el comando #apt-get update e instalo con:

- #apt-get install elasticsearch.

Se verificó su instalación en la siguiente rura:

- #cd /etc/elasticsearch.

En esta ruta se puede consultar el archivo de configuración: elasticsearch.yml

4.4 Kibana

Como kibana pertenece al mismo stack de elastic, ya no fue necesario importar las claves GPG y las fuentes de Elastic, solo se instaló con el siguiente comando:

- #apt-get install kibana.

Instalado Kibana, se verificó la ruta donde se encuentra el archivo de configuración: #cd /etc/kibana, y el archivo es kibana.yml.

4.5 Filebeat

Logstash es un pipeline de procesamiento de datos gratuito y abierto del lado del servidor que ingesta datos de una multitud de fuentes, los transforma y luego los envía a tu "escondite" favorita. Logstash procesa cualquier tipo de información e interpreta sin importar su complejidad. Para nuestro caso práctico, el stack de elastic tiene un software de la misma proporción de Logstash, con la diferencia que procesa información de sistemas exclusivos ya preconfigurados, beat reenvía log de forma liviana y procesa de una forma más rápida.

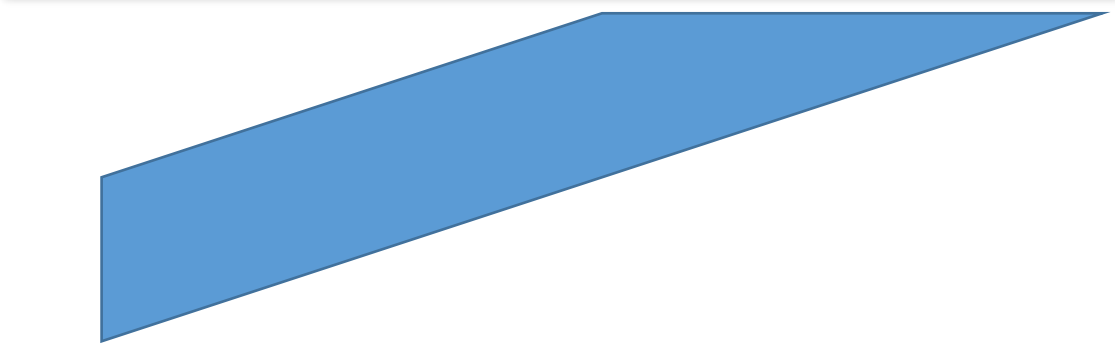
Para uso practico de este trabajo de investigación, se usó un beat preconfigurado del stack de Elastic, Filebeat tiene una métrica configurada para recopilar los logs de suricata IDS.

Para la instalación se aplicó el comando desde la terminal de la misma forma que los anteriores:

#apt-get install Filebeat, y se verificó la ruta de instalación en: #cd /etc/Filebeat, donde se encuentra el archivo Filebeat.yml.

Ilustración 9: ELK IDS

CAPITULO 5



5 IMPLEMENTACIÓN Y ANALISIS DE LOS SISTEMAS DE DEFENSA

Como hemos mencionado con anterioridad, debido a que en la actualidad las amenazas de ataques cibernéticos crecen día con día, se han desarrollado herramientas (software y hardware) capaces de mitigar estas actividades ilícitas que puede producir un atacante informático. En este trabajo se documenta la implementación de un IDS, [64], y la implementación de un SIEM, que tienen la capacidad de almacenar y procesar archivos de tipo log,pcap, eve.son, etc. [59]. Estos sistemas, juegan un papel muy importante en la seguridad informática de cualquier red, permiten a los equipos de seguridad controlar todo lo que está pasando en la red en tiempo real para así poder reaccionar rápidamente ante posibles ataques y/o vulnerabilidades. También, no solo se controla el tráfico externo de la red sino también el interno, por ejemplo, de los propios trabajadores, que webs se visitan, que inicios de sesión se producen y si han sido víctimas de algún ataque [65].

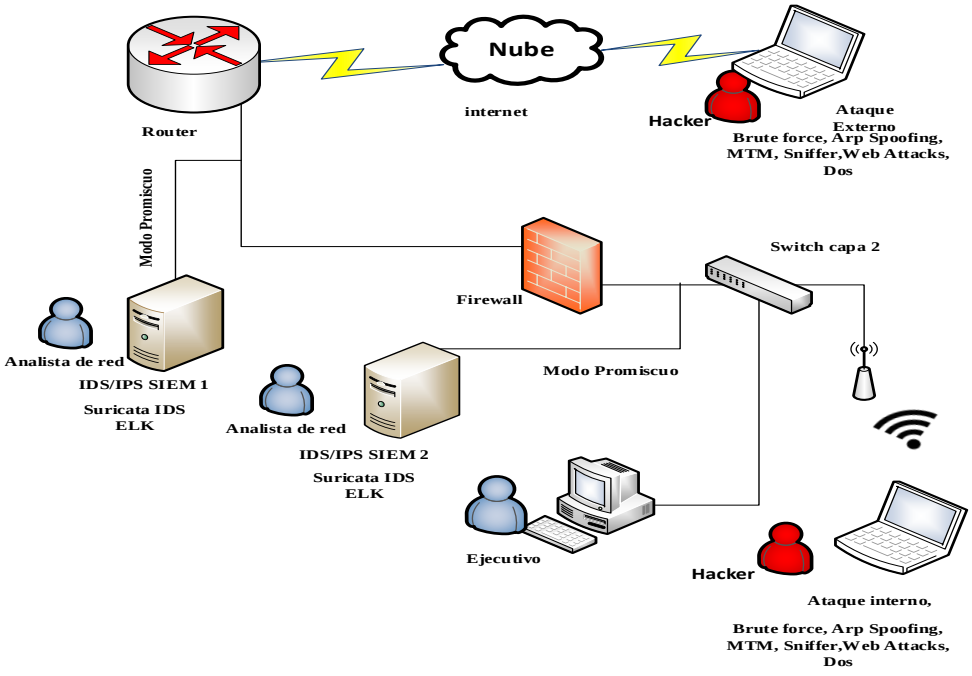


Ilustración 10: Topología de Red

Como se mostró en el capítulo anterior, en el desarrollo de esta monografía, se trabajó con el sistema operativo Ubuntu para realizar la instalación y configuración del IDS suricata y el SIEM ELK.

Para configurar suricata, se realizó la edición correspondiente en el archivo de configuración suricata.yaml desde la terminal mediante el comando **nano /etc/suricata/suricata.yaml.**

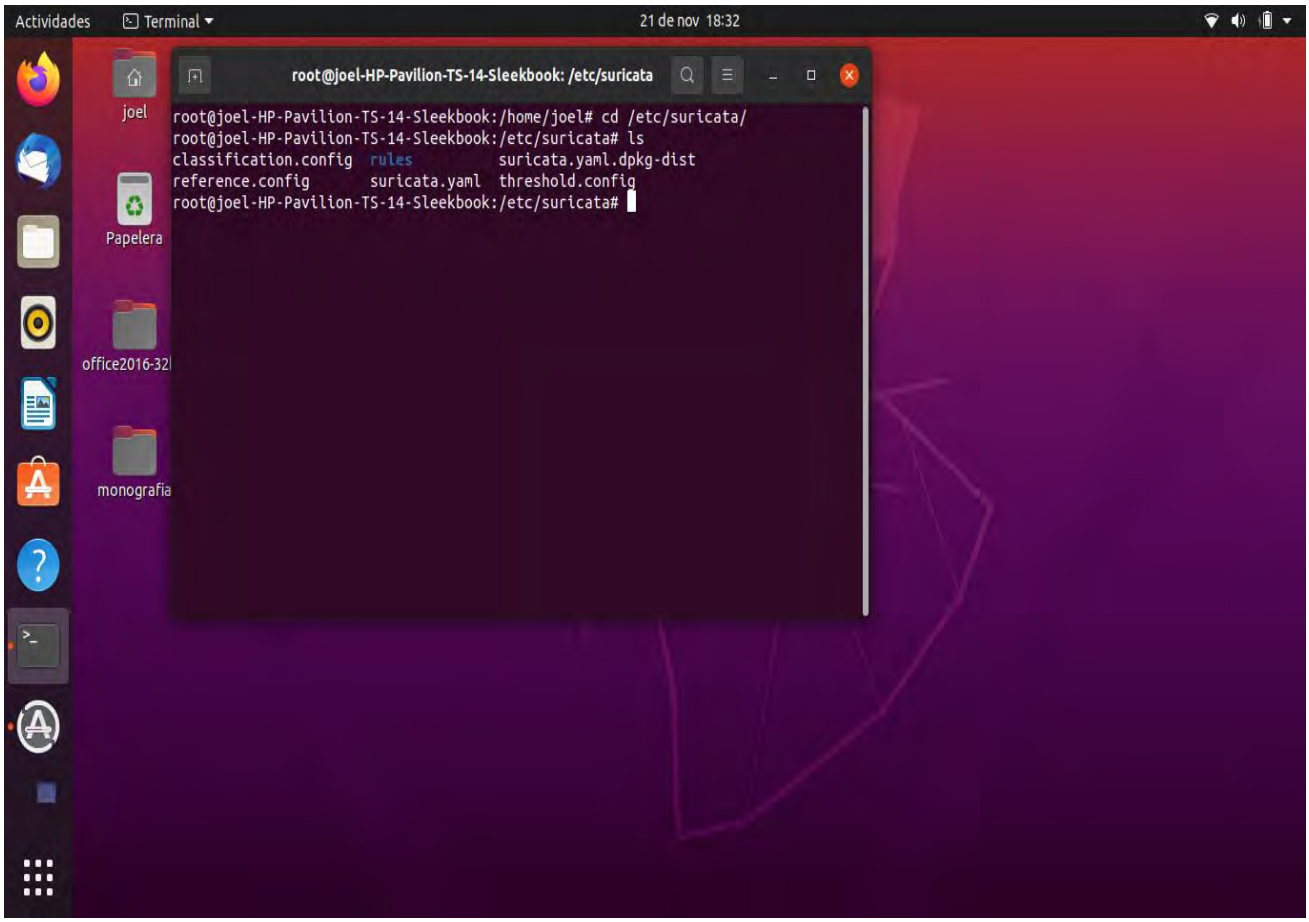


Ilustración 11: Ruta Suricata

Dentro de este archivo de configuración, se verificó que no estén comentadas las líneas de configuración y de no existir se agregan las líneas de configuración necesarias para su operación optima.

HOME_NET: "[192.168.0.0/16,10.0.0.0/8,172.16.0.0/12]"

EXTERNAL_NET: "!$HOME_NET".

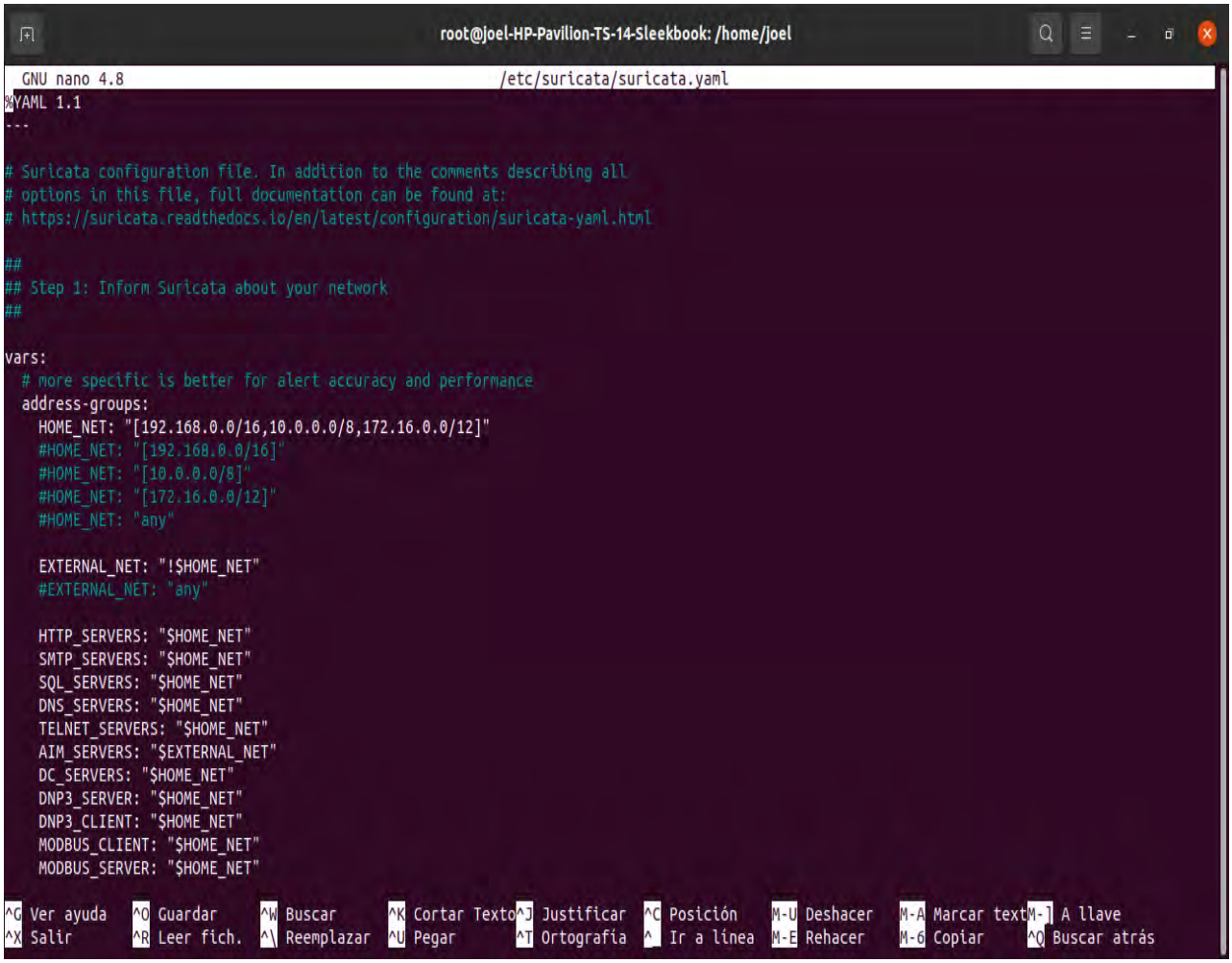


Ilustración 12: Configuracion de Red Suricata

Se indicó en el apartado dafault-rule-path, la ruta donde se encuentran las reglas preconfiguradas para la detección de amenazas o anomalías ya existentes, creadas por la comunidad OISF y reglas personalizadas que apuntan a los archivos indicados dentro la sección rules-files.

Default-rule-path: /var/lib/suricata/rules

Rules-files:

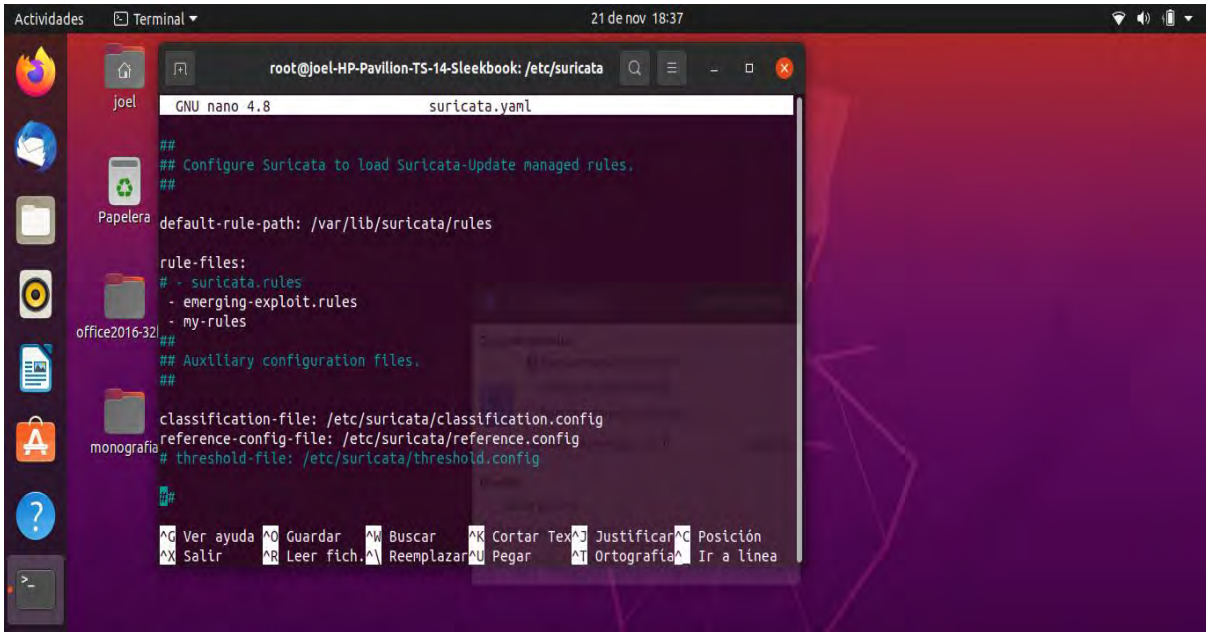


Ilustración 13: Ruta de Reglas Suricata

Se editó la sección af-packet en la cual se indicó la interfaz de red encargada de procesar e interceptar todo el tráfico de red (poner en modo promiscuo la interfaz de red para pasar todo el tráfico por el servidor suricata). Se creó una regla básica en la ruta #nano /etc/suricata/rules/rules.

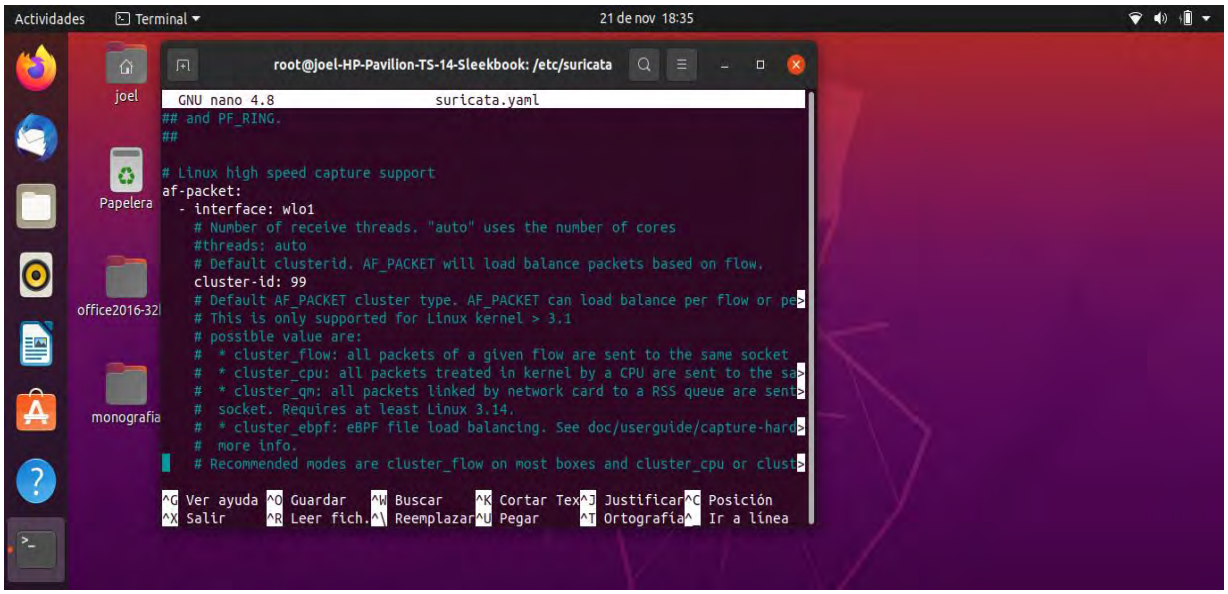


Ilustración 14: Interfaz de Red

Para poner en operación a Suricata se utilizó el siguiente comando:

suricata -c /etc/suricata/suricata.yaml -i wlo1(donde wlo1 es la interfaz de red del servidor Ubuntu).

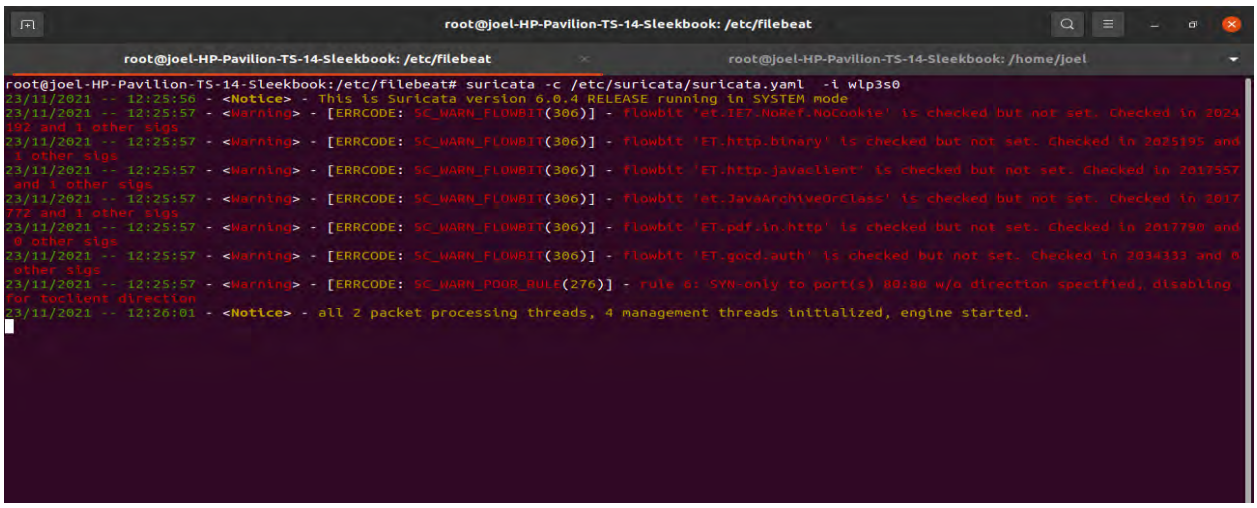


Ilustración 15: Modo Escucha Suricata

En la ruta /var/log/, se encuentran los archivos eve.json, fast.log, suricata.log que contienen información del análisis de la red que está monitoreando suricata IDS.

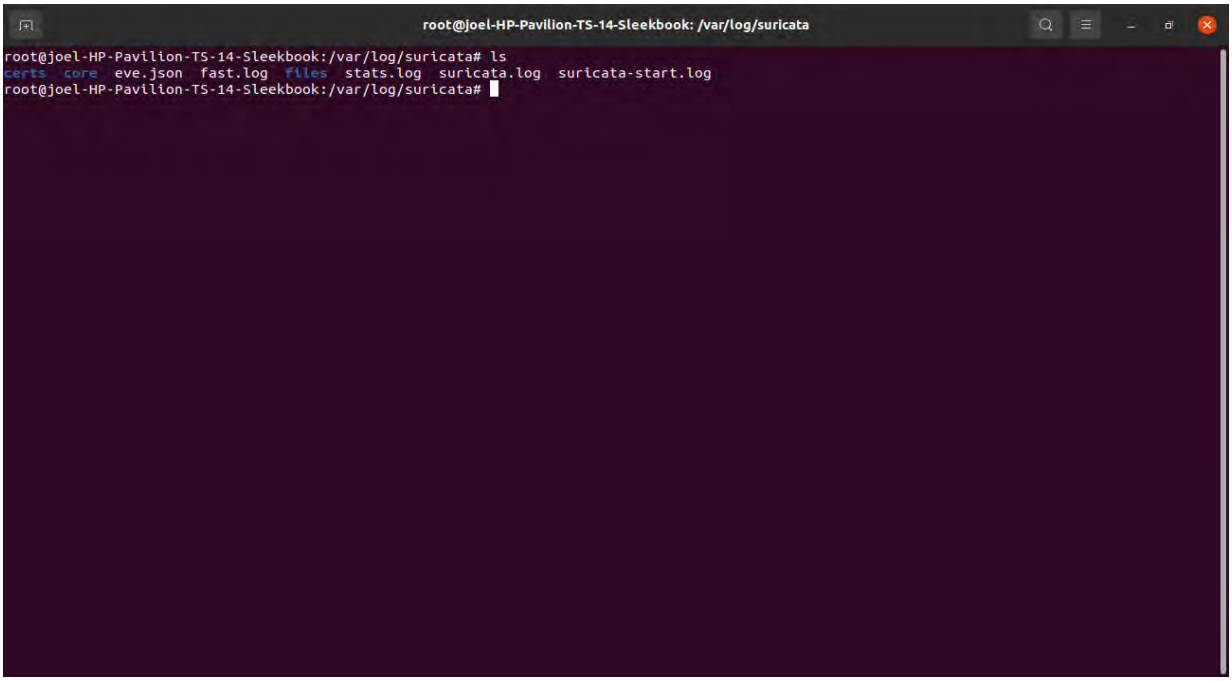


Ilustración 16:Ruta Archivos Log

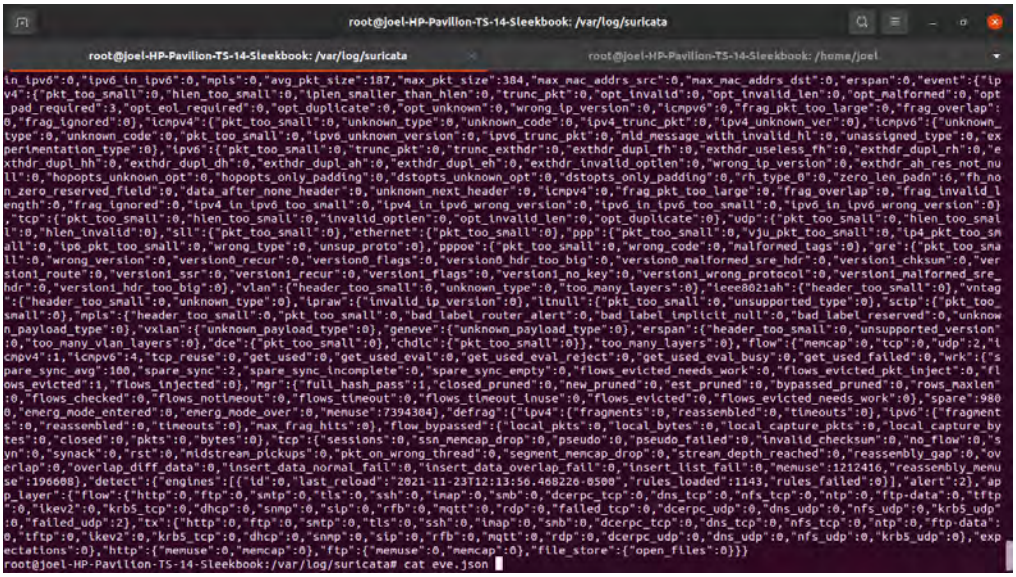


Ilustración 17: Leyendo Archivo eve.json

5.1 Configuración y Conexión a ELK

Una vez configurado suricata IDS para el escaneo de red, se configuró Elasticsearch y kibana para que puedan buscarse entre ellos.

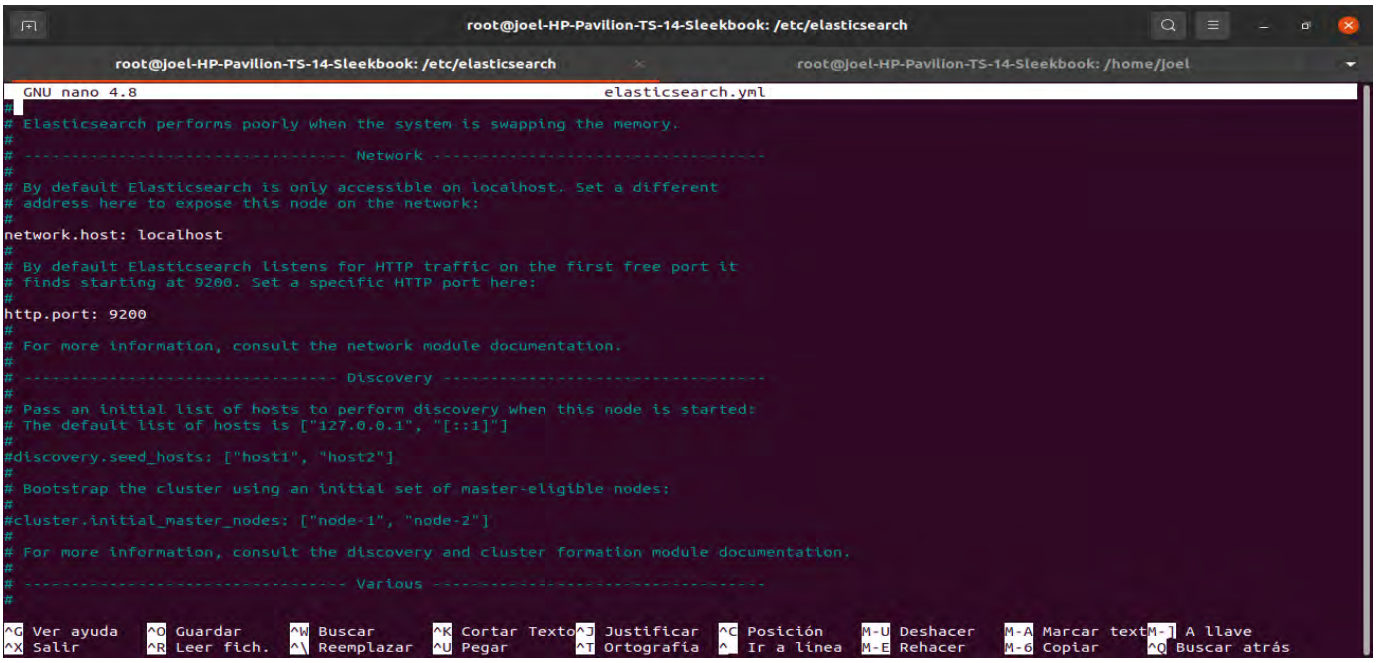


Ilustración 18: Habilitando Puerto e IP a Elastisearch

Se corroboró que el servicio se encuentre funcionando, usando el comando: **curl -x GET "localhost:9200"**

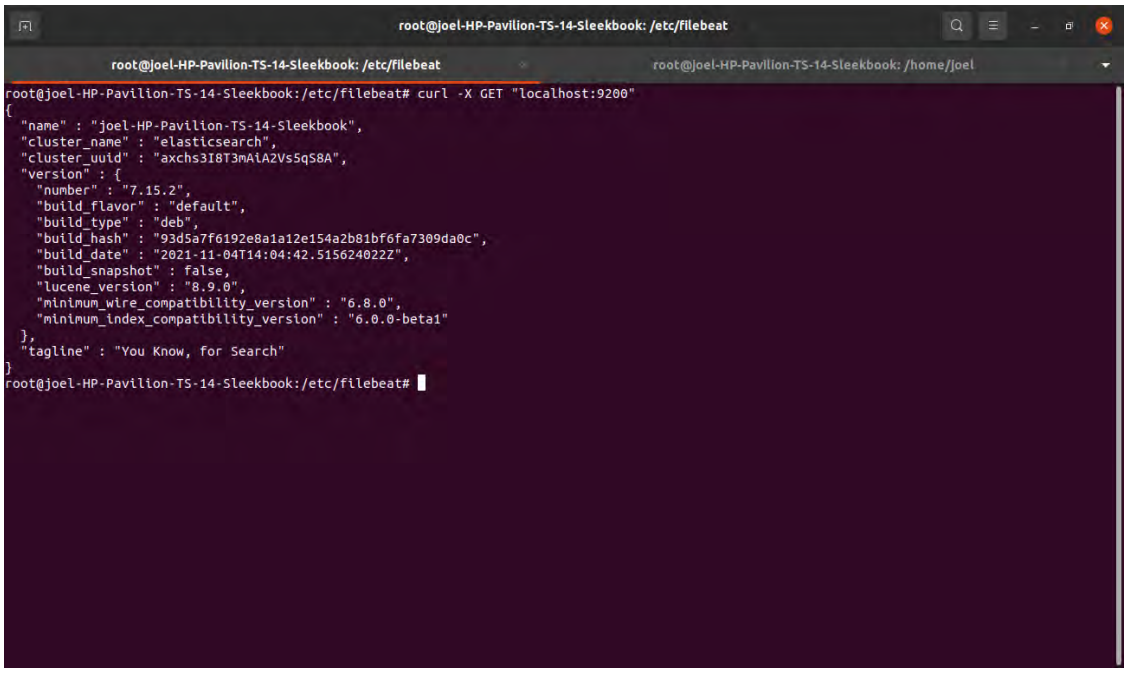


Ilustración 19:Elasticsearch

Para el caso de kibana el archivo de configuración se encuentra en la ruta **/etc/kibana**, ahí se editó el archivo **kibana.yml** con el comando nano **kibana.yml.**

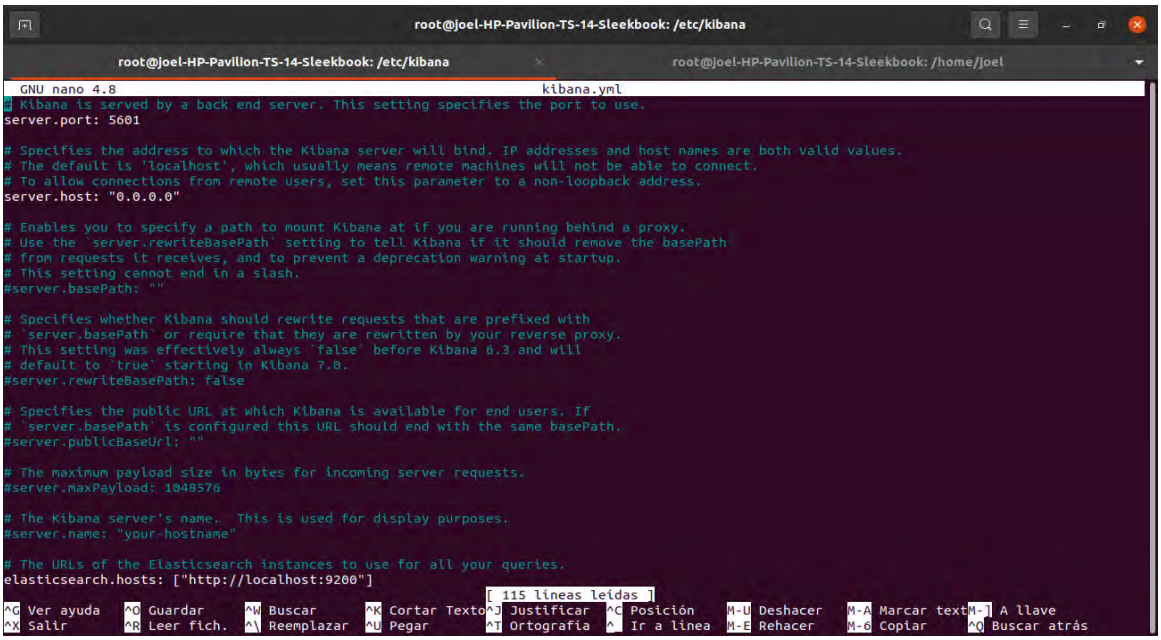


Ilustración 20: Edición kibana.yml

Se verificó que este habilitado la línea elasticsearch.host: [“http://localhost:9200”], para que kibana pueda establecer conexión con elasticsearch y pueda cargar los log almacenados e indexados. Para verificar que el sistema kibana se encuentre funcionando se usa el comando systemctl status kibana, si se encuenra activo podremos usar cualquier navegador de internet y entrar mediante IP. Se ingreso al visualizador Kibana mediante **localhost:5601** o con la IP del servidor **192.168.1.65:5601,** donde 5601 es el puerto de comunicación que utiliza kibana, esta IP se registró en el archivo de configuración **kibana.yml** en el la línea **server.host**, si en el archivo de configuración se deja como IP 0.0.0.0, indica al sistema que se puede usar cualquier IP que tenga el servidor kibana, si está configurado por DHCP.

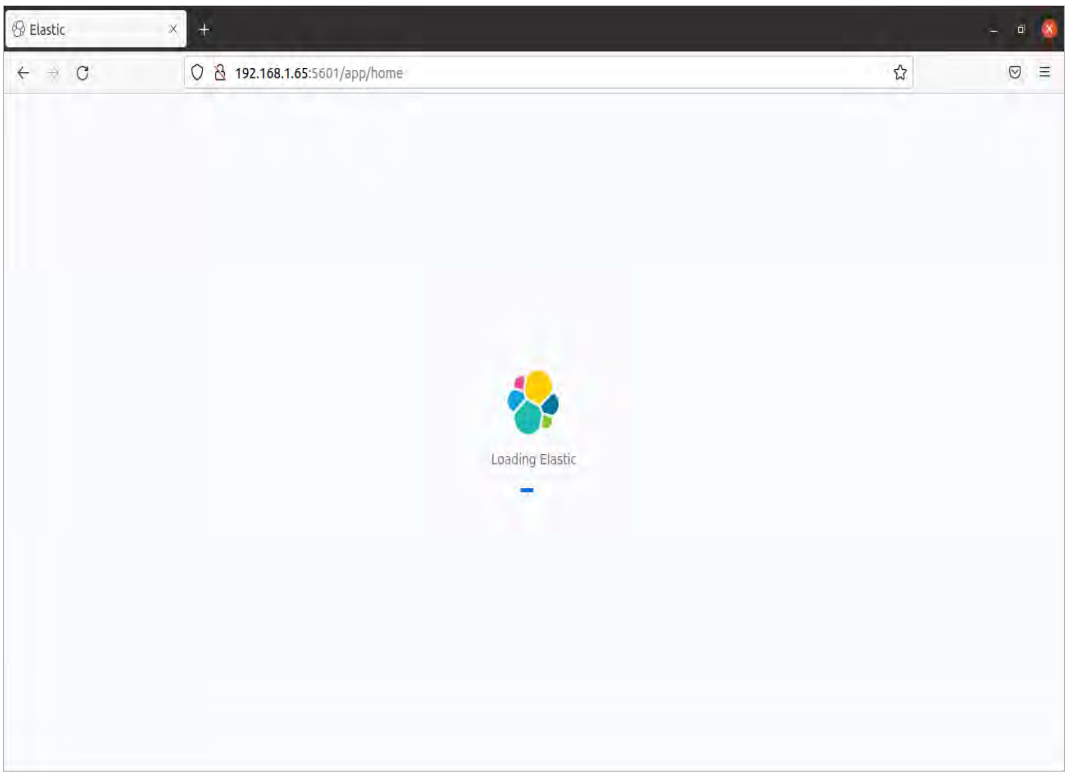


Ilustración 21:Entorno Web kibana

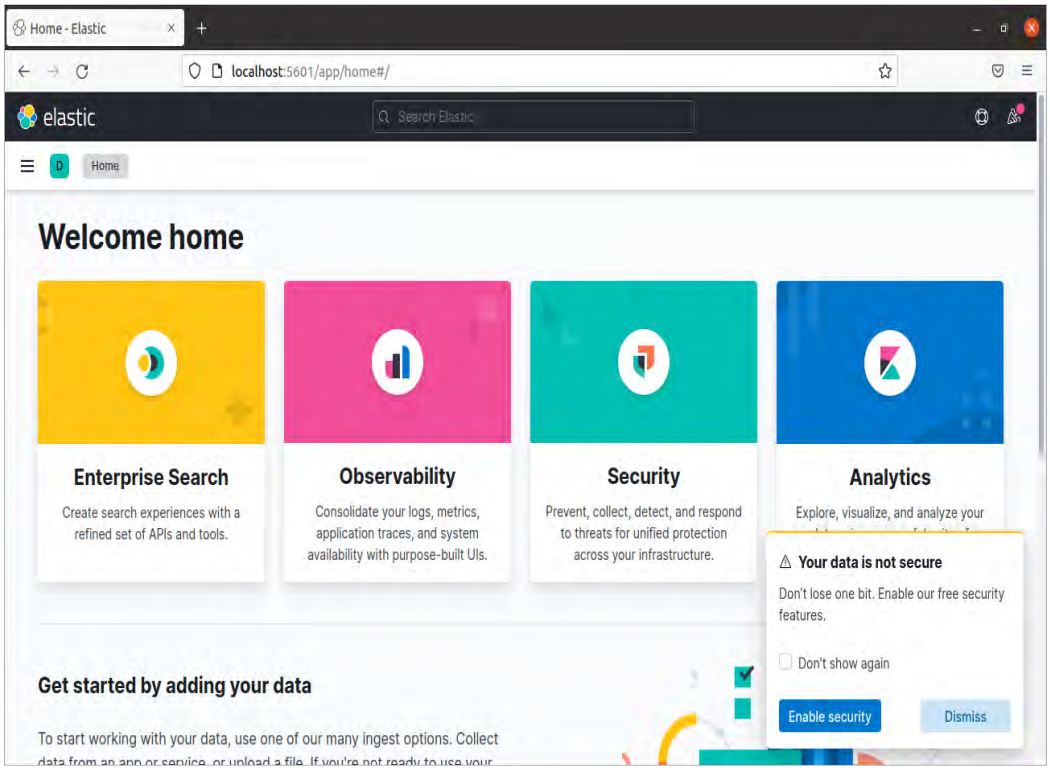


Ilustración 22:Bienvenidos Kibana.

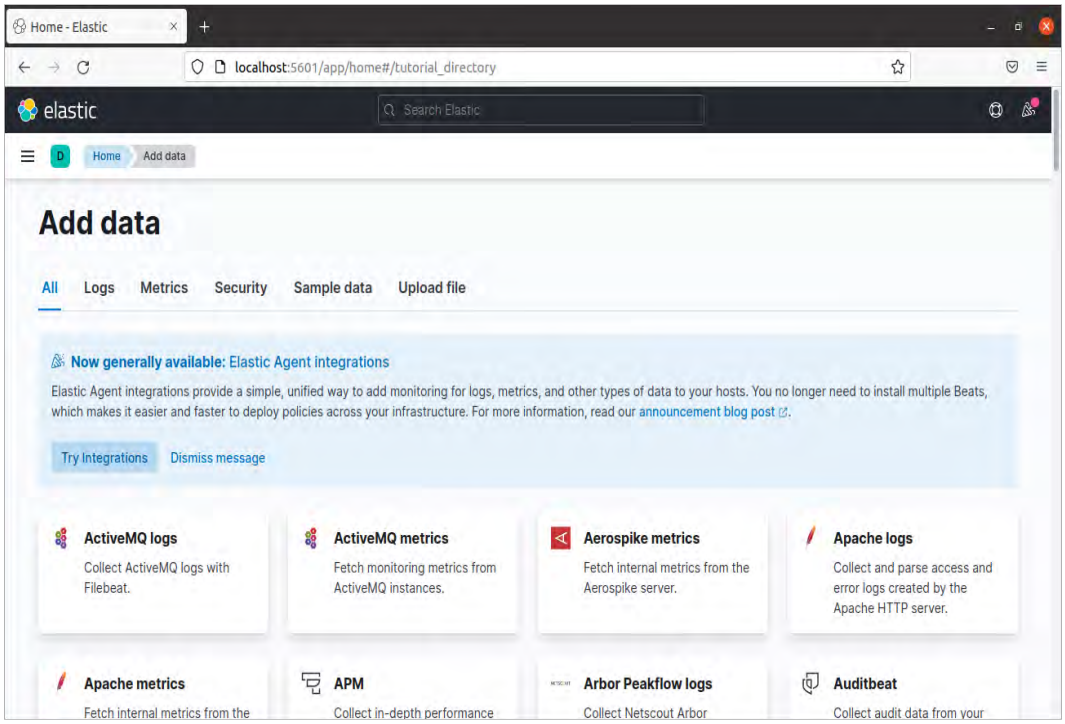


Ilustración 23: Agregamos Datos

Una vez en el entorno web de kibana, se agregó el módulo de suricata logs para visualizar la información que proporcionan los log, eve,json de suricata.

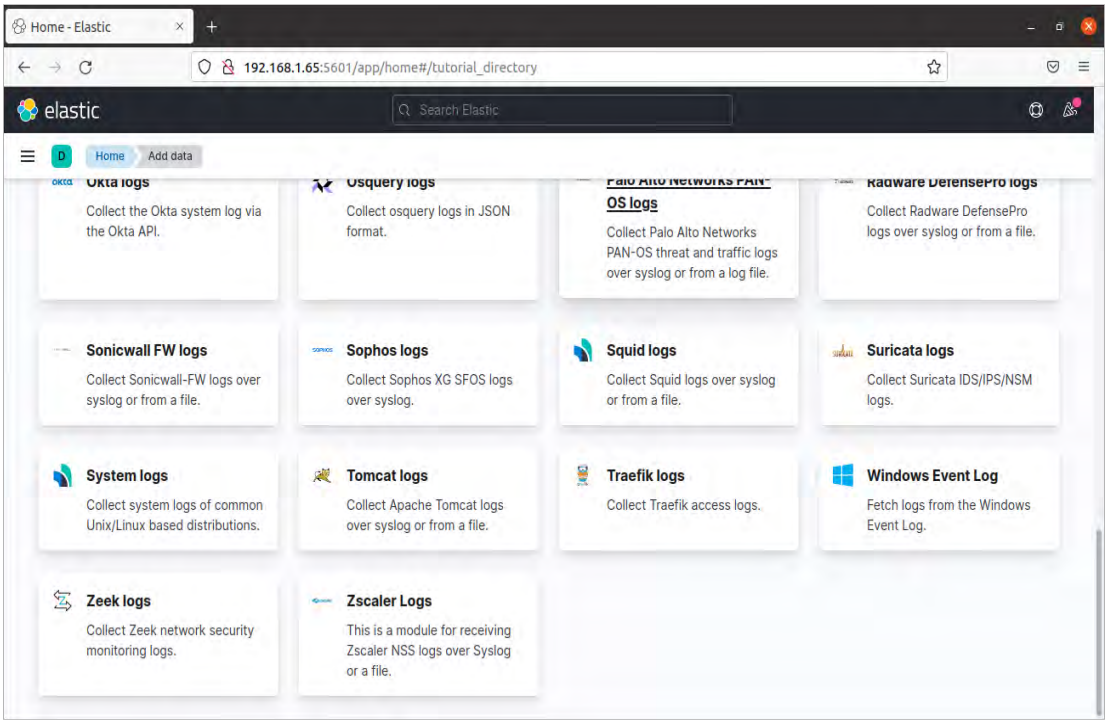


Ilustración 24:Métrica Suricata log.

Para hacer el llamado del log entre kibana elasticsearch y suricata, se usó Filebeat, Filebeat es un cargador ligero para reenviar y centralizar datos de registro. Instalado como un agente en los servidores, Filebeat monitorea los archivos de registro o las ubicaciones que se especifiquen, recopila eventos de registro y los reenvía a Elasticsearch o Logstash para su indexación. De esta manera monitoreamos en tiempo real todos los eventos producidos por suricata IDS.

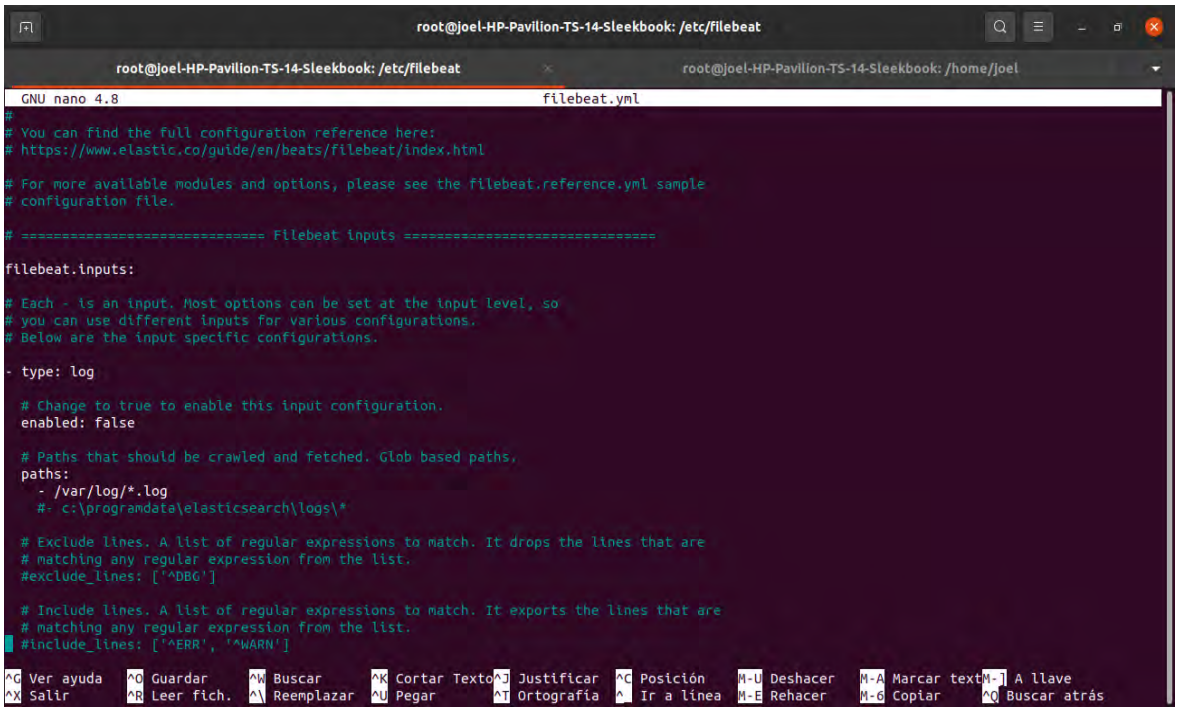


Ilustración 25: configuracion Filebeat

Se verificó que se encontrará la línea con la sintaxis para comunicarse con kibana. Dentro de Filebeat.yml la línea **host: "localhost:5601"**

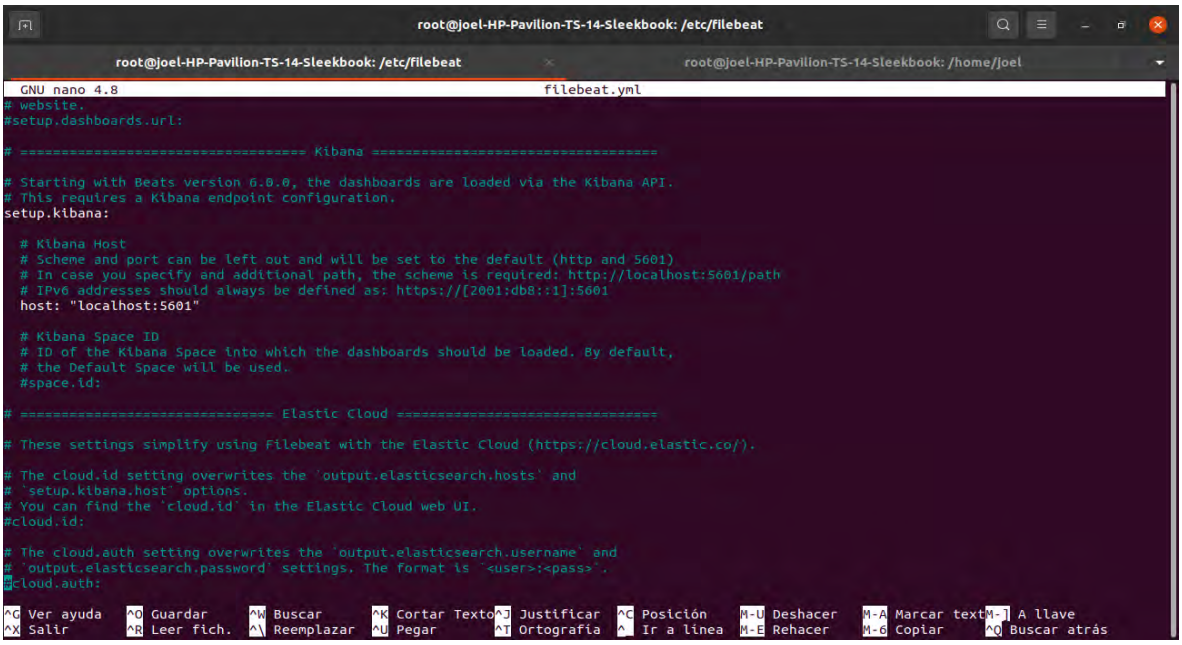


Ilustración 26:Filebeat-Kibana

De igual forma con elasticsearch **output.elasticsearch: hosts: [“localhost:9200”]**

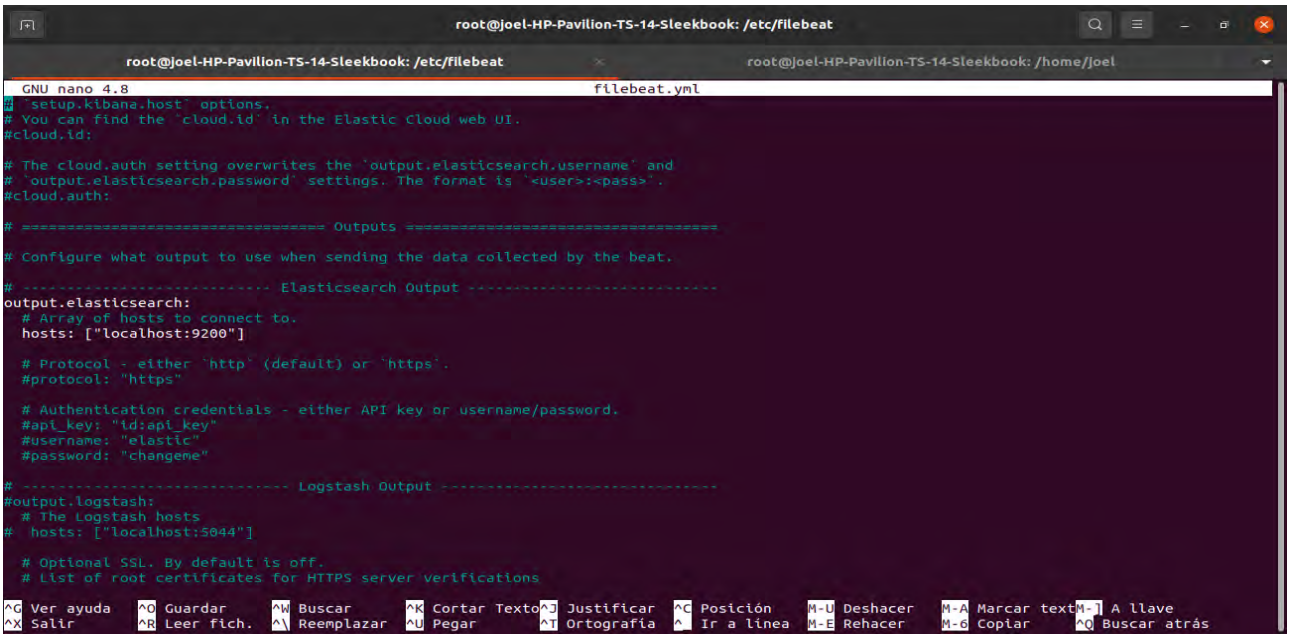


Ilustración 27: Filebeat-Elasticsearch

Una vez configurado, se cargó el módulo de suricata Filebeat para que kibana pueda proporcionar los dashboard de los logs capturados. En la terminal Shell se usó el siguiente comando:

filebeat modules enable suricata

filebeat setup

En el módulo de métricas de seguridad desde el entorno web de kibana, cargamos el dashboard, para visualizar las gráficas que procesa kibana en tiempo real de los eventos y alertas que produce Suricata.

5.2 Resultados de Configuración Final

Para comprobar las configuraciones de suricata y que, mediante el SIEM, se generen resultados positivos, se creó la regla sencilla de enviar una alerta al momento de hacer un ping a la red (ICMP) desde cualquier host que esté conectado. Suricata capturara el tráfico generado por cualquier equipo y generara un archivo log, que este a su vez lo recuperara Filebeat, lo

almacenara Elasticsearch para indexar el log y poder visualizarlo en nuestro sistema de monitoreo Kibana:

- alert icmp any any -> $HOME_NET any (msg:"ICMP connection attempt"; sid:1000002; rev:1;) <------ detecta los pings enviamos a él host configurado con suricata
- alert tcp any any -> $HOME_NET 22 (msg:"SSH connection attempt"; sid:1000003; rev:1;) <---------- detecta conexiones al puerto 22 donde generalmente se encuentra el servicio Secure Shell
- alert tcp any any -> $HOME_NET 80 (msg:"DDoS Unusually fast port 80 SYN packets outbound, Potential DDoS"; flags: S,12; threshold: type both, track by_dst, count 500, seconds 5; classtype:misc-activity; sid:6;) <------- detecta cuando hay un envió excesivo de paquetes al puerto 80

Los dashboards en Kibana permiten crear vistas que reúnen gráficos, mapas y filtros para mostrar el panorama completo de los datos generados por suricata y almacenados en Elasticsearch, así como monitoreo de amenazas en tiempo real.

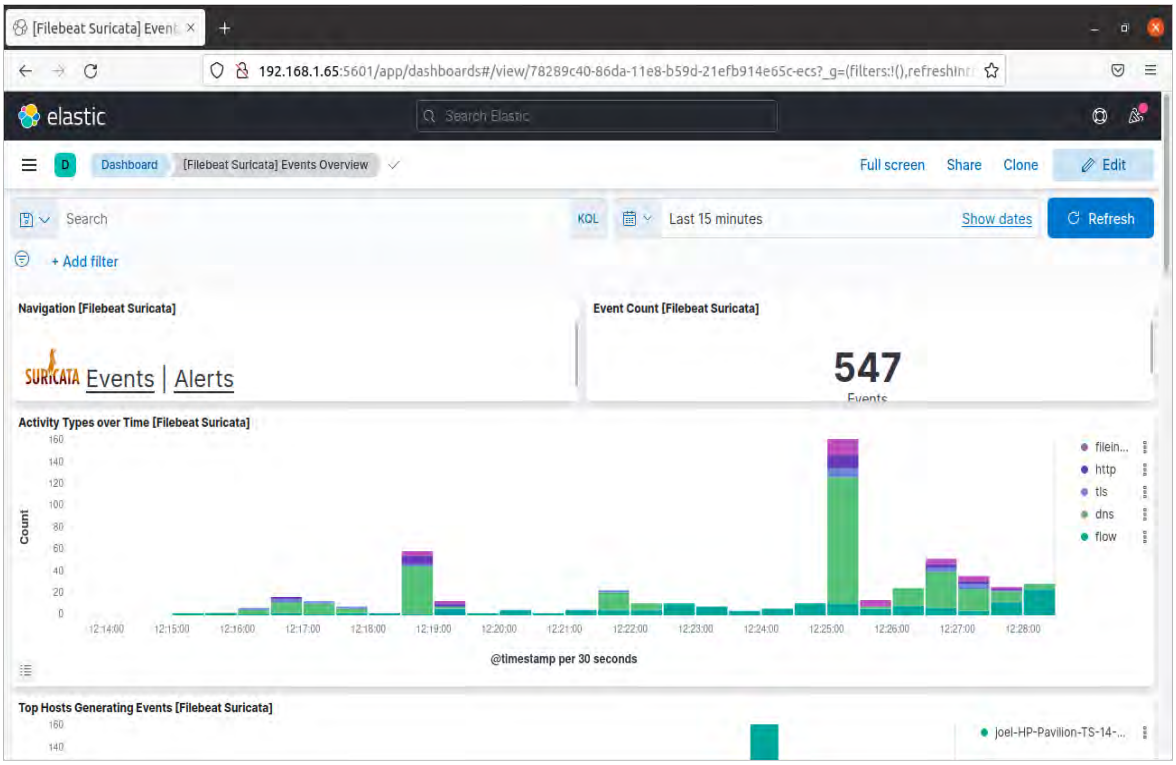


Ilustración 28:Dashboard 1 Suricata

En la Ilustración 56 se observa el recuento de eventos generado en los logs de suricata y la actividad por protocolos de red en función del tiempo.

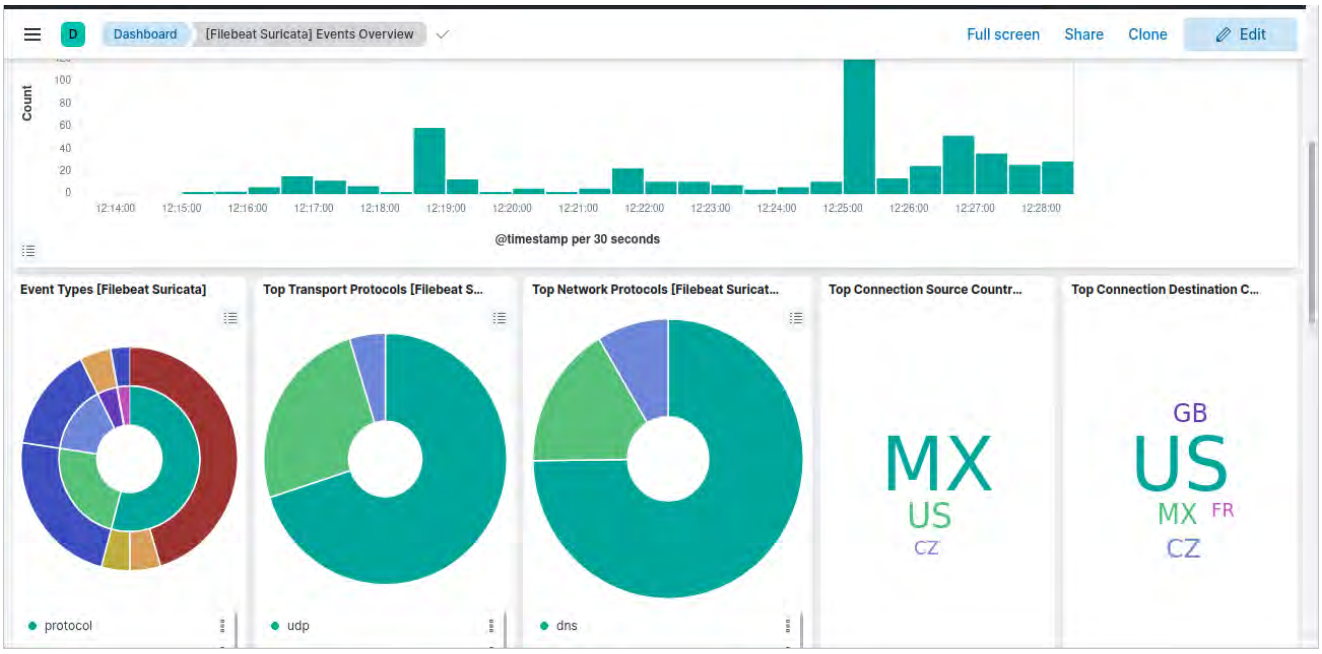
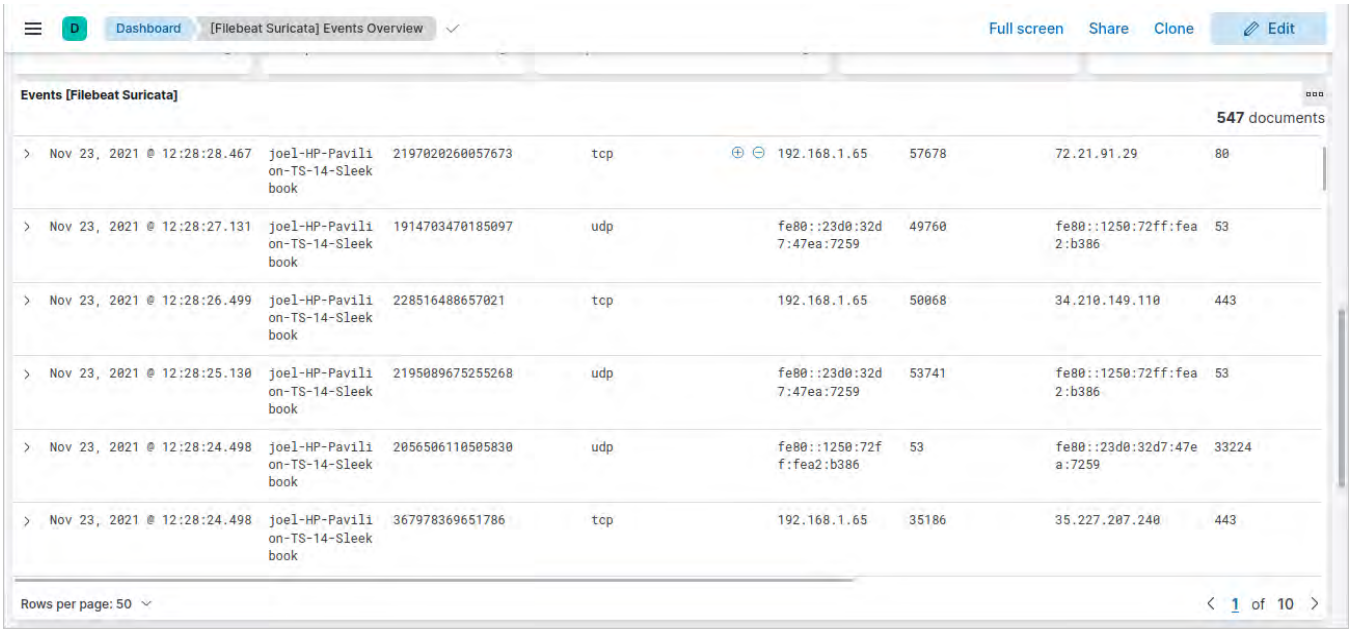


Ilustración 29:Dashboard Suricata 2

En el dashboard de la Ilustración 56 se visualizan los protocolos más utilizados en la red, así como los orígenes y destino de conexiones a las salidas de internet.



Nov 23, 2021 @ 12:28:28.467	joel-HP-Pavilion-TS-14-Sleekbook	2197020260057673	tcp	192.168.1.65	57678	72.21.91.29	80
Nov 23, 2021 @ 12:28:27.131	joel-HP-Pavilion-TS-14-Sleekbook	1914703470185097	udp	fe80::23d0:32d7:47ea:7259	49760	fe80::1250:72ff:fea2:b386	53
Nov 23, 2021 @ 12:28:26.499	joel-HP-Pavilion-TS-14-Sleekbook	228516488657021	tcp	192.168.1.65	50068	34.210.149.110	443
Nov 23, 2021 @ 12:28:25.130	joel-HP-Pavilion-TS-14-Sleekbook	2195089675255268	udp	fe80::23d0:32d7:47ea:7259	53741	fe80::1250:72ff:fea2:b386	53
Nov 23, 2021 @ 12:28:24.498	joel-HP-Pavilion-TS-14-Sleekbook	2056506110505830	udp	fe80::1250:72ff:fea2:b386	53	fe80::23d0:32d7:47ea:7259	33224
Nov 23, 2021 @ 12:28:24.498	joel-HP-Pavilion-TS-14-Sleekbook	367978369651786	tcp	192.168.1.65	35186	35.227.207.240	443

Ilustración 30: Dashboard Suricata 3

En la Ilustración 58 se observa la fecha y hora de cada evento creado, se proporciona información referente a el protocolo de red utilizado, el nombre del equipo en la red que está generando alertas, la IP utilizada por el equipo, el puerto por el que esta interactuando, así como otros datos de importancia a la hora de analizar amenazas o comportamientos de los usuarios en red.

Se observa también cuántos paquetes, eventos o alertas del tipo suricata.eve se han producido hasta el momento, así esto servirá para el monitoreo en tiempo real.

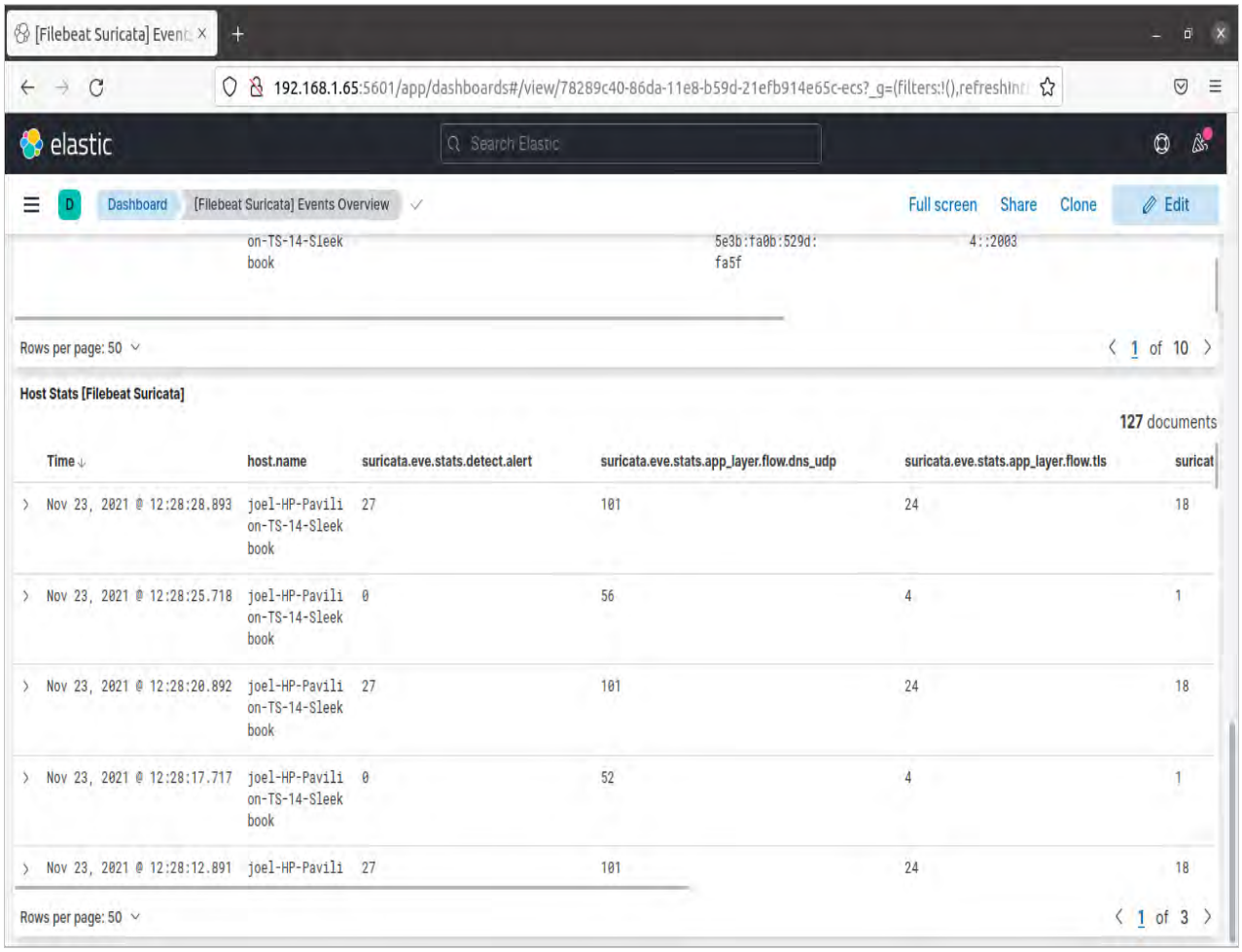


Ilustración 31: Dashboard Suricata 4

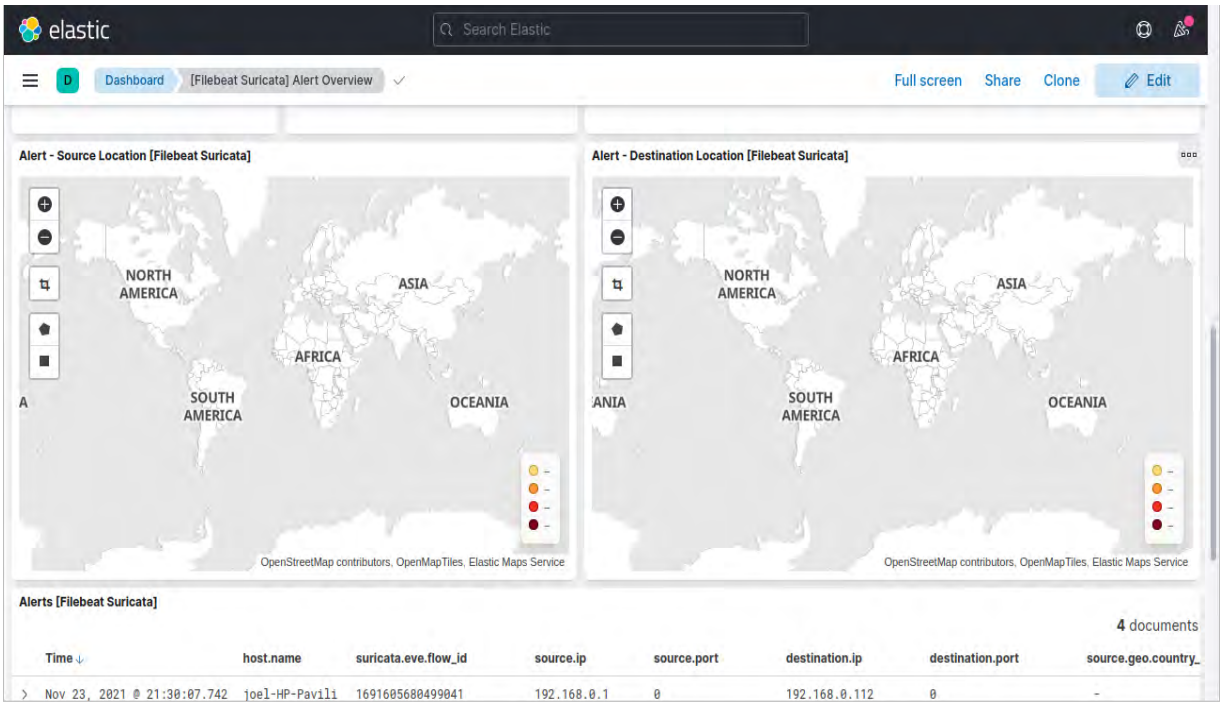


Ilustración 32: Dashboard Suricata 5

La Ilustración 59 muestra los dashboards tipo mapa que permiten visualizar por geolocalización los intentos de ataques.

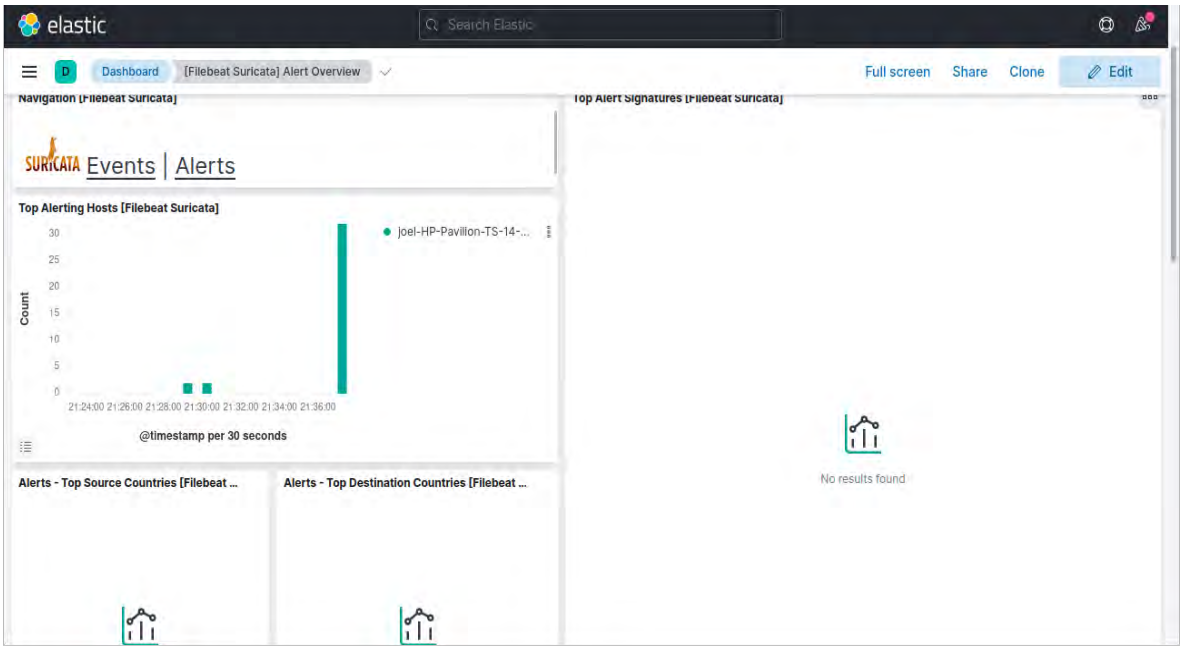


Ilustración 33: Dashboard Suricata 6.

La Ilustració 60 muestra los equipos que más alertas activan en el paso del tiempo.

CONCLUSIONES

En este trabajo monográfico se presenta la documentación sobre la implementación de un IDS (Suricata OISF), en combinación con nuevas herramientas de código abierto, como los sistemas de Gestión de Eventos e Información de Seguridad.

Se implementó Suricata como herramienta de análisis, para alertar de posibles ataques cibernéticos como: **DOS**,**DDOS**, **Backdoors**, **Inyection SQL**, entre muchos más que existen. Suricata es una herramienta muy completa que trabaja con archivos del tipo json, los cuales son archivos de textos que sistema puede interpretar debido a que usa una sintaxis estructurada. Una ventaja de utilizar este sistema de defensa es que tiene una comunidad de desarrollo muy activa, en constantes actualización sobre los nuevos tipos de ataques que se generan día a día.

Para complementar a los IDS, se han creado los sistemas SIEM, que son programas de computadora capaces de gestionar eventos, procesar información de logs y generar gráficos de fácil acceso en un entorno web. Existen diversas herramientas comerciales que integran sistemas SIEM, tales como QRadar de IBM, RSA enVision, Security MARS, empow o Alien Vault USM; las cuales, tienen un alto costo. Sin embargo, en este trabajo se propone el uso de la solución de código abierto como ELK Stack (Elasticsearch-Logstash-Kibana).

Abreviaturas

Tabla 8: Abreviaturas

IDS	Intrusion Detected System	Sistema de detención de intrusos
SIEM	Security Information Event Management	Información sobre seguridad y gestión de eventos
IP	Internet Protocol	Protocolo de Internet
OISF	Open Information Security Foundation	Fundación de la seguridad libre
DOS	Denial of service	Denegación de servicio
DDOS	Distributed Denial Of Service.	Ataque de red distribuido
SEM	Security event management	Gestión de eventos de seguridad
SIM	Security information mnagement	Gestión de información de seguridad
FTP	File Transfer Protocol	Protocolo de transferencia de archivos
VPN	Virtual Private Network	Red privada virtual
UDP	User datagram protocol	Protocolo de datagramas de usuario
TCP	Transmision Control Protocol	Protocolo de transmisión de control
ICMP	Internet Control Message Protocol	Protocolo de control de mensajes de internet
HTTPS	Hyper Text Transfer Protocol Secure	protocolo seguro de transferencia de hipertexto
SSL	Security socket layer	Capa de socket seguro
TLS	Transport Layer Security	Capa de trasporte segura
SMTP	Simple Mail Transfer Protocol	Protocolo de transferencia simple de correo
SSH	Security shell	Protocolo de comunicación segura

DNS	Domain Name Service	Servicio de nombres de dominio
IBM	International Business Machines	Máquina de negocios internacionales
JDK	Java Developer Kit	Kit de desarrollo de java

BIBLIOGRAFÍA

[1] J. Erickson, Hacking: the art of exploitation., Pollock, 2008.

[2] J. C. Santos, Seguridad informática, Ram-ma editorial, 2015.

[3] G. González-Granadillo, S. González Zarzosa y R. Diaz, «Security Information and event management (SIEM): Analysis,,» *Sensors,* 2021.

[4] Center y s. a. a. W. Information, «15 Cybersecurity fundamentals for water and,» *WaterIscac,* p. 56, 2019.

[5] F. Alessandro, G. G. Battista, P. Fabio, M. P. Girdinio y Mario, «Toward the Integration of Cyber and Physical Security,» *Sensors,* 2021.

[6] D. Miller, S. Harris, A. Harper, S. VanDyke y C. Blask, Security Information and Event Management (SIEM) Implementation, New York: McGraw Hill Professional, 2010.

[7] K. Kent y M. Souppaya, Guide to computer security log management, NIST Spec, 2006.

[8] W. Stallings, Network security: Application and Standards, Prentice Hall, 2011.

[9] Rediris, «Sistema de detencion de intrusos,» 12 11 2011. [En línea]. Available: https://www.rediris.es/cert/doc/unixsec/node26.html.

[10] M. Nieles, K. Nieles y V. Yan, An Introduction to Computer Security : The NIST Handbook., 1995.

[11] W. Stallings y L. Brown, Computer security principles and practice, Pearson, 2008.

[12] F. Hussain, H. S, R. Hussain y E. Hossain, Machine learning for resource management in cellular and IoT networks:Potentials, current solutions, and open challenges, 2020.

[13] Aljabri, S. Malak, A. Sumayh, M. R. Mustafa, S. M. H Almotiri S, M. Samiha, Fatima Anis, A. Menna, M. A. Dorieh y H. A. Dina, «Intelligent techniques for detecting network attacks: Review and research directions,» *Sensors,* p. 43, 2021.

[14] M. Shaikh y R. Vadivel, Cloud computing: Major challenges and counter acts, 2018.

[15] S. Goudos, P. Dallas, S. Chatziefthymiou y S. Kyriazakos, A Survey of IoT Key enabling and future technologies: 5G, Mobile IoT, sematic web and applications, 2017.

[16] J. F. R. Buendia, Seguridad Informática, España: Macgraw Hill , 2013.

[17] A. Lazarevic, V. Kumar y J. Srivastava, “Intrusion detection: A survey”, Springer, 2005.

[18] D. Yeung, “Host-based intrusion detection using dynamic and static behavioral models,”, 2003.

[19] S. E. Smaha, “Haystack: An intrusion detection system,”, Orlando Florida, 1988.

[20] P. Lichodzijewski, A. N. Zincir-Heywood y M. I. Heywood, “Host-based intrusion detection using self-organizing maps,”, Honolulu, 2002.

[21] H. Jiankun, Y. Xinghuo, D. Qiu y C. Hsiao-Hwa, “A simple and efficient hidden markov model scheme for host-based anomaly intrusion detection,”, 2009.

[22] G. Creech and J. Hu, “A semantic approach to host-based intrusion detection systems using contiguous and discontiguous system call patterns,”, 2014.

[23] A. J. Hoglund, K. Hatonen y A. S. Sorvari, “A computer host-based user anomaly detection system using the self-organizing map,", italy, 2010.

[24] H. Debar, M. Dacier y A. Wespi, “Towards a taxonomy of intrusiondetection systems”, 1999.

[25] P. Biddle, P. England, M. Peinado y B. Willman, “The darknet and the future of content distribution, 2002.

[26] A. Sperotto, “An overview of IP flow-based intrusion detection,”, 2010.

[27] W. Lee y Stolfo, “Data mining approaches for intrusion detection”, San Antonio TX, 1998.

[28] M. Roesch, “Snort: Lightweight Intrusion detection for networks,”, 1999.

[29] V. Paxson, “Bro: A system for detecting network intruders in real-time,”, 1999.

[30] «Suricata-IDS.,» [En línea]. Available: https://suricata-ids.org. [Último acceso: 12 Septiembre 2017].

[31] Alienvault, «Alienvault IDS unified security management,» [En línea]. Available: www.alienvault.com/ids. [Último acceso: 12 Septiembre 2017].

[32] Cisco, «Firepower next-generation IPS (NGIPS).,» [En línea]. Available: https://www.cisco.com/c/en_uk/products/security/ngips/index.html. [Último acceso: 12 Septiembre 2017].

[33] FireEye, «Security Solutions.,» [En línea]. Available: https://www.fireeye.com/products/nx-network-security-products.html. [Último acceso: 12 Septiembre 2017].

[34] H. Dreger, F. A, V. Sommer y R. Paxson, “Operational experiences with high-volume network intrusion detection”, 2004.

[35] C. Tsung-Huan, L. Ying-Dar, L. Yuan-Cheng y P.-C. Lin, “Evasion techniques:Sneaking through your intrusion detection/prevention systems,”, 2012.

[36] C. V, A. Banerjee y V. Kumar, “Anomaly detection: A survey,” ACM Comput., 2009.

[37] B. D., Couto, J. S., L. Popyack y N. Wu, “ADAM:Detecting intrusions by data mining,”, 2001.

[38] G. Kim, S. Lee y S. Kim, “A novel hybrid intrusion detection method integrating anomaly detection with misuse detection., 2014.

[39] O. Depren, M. Topallar, E. Anarim y M. K. Ciliz, “An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks,”, 2005.

[40] J. Zulkernine y M. Zhang, “A hybrid network intrusion detection technique using random forests,”, Vienna, 2006.

[41] D. Anderson, T. Frivold y A. Valdes, “Next-generation intrusion detection expert system, 1995.

[42] M. H. Bhuyan, D. K. Bhattacharyya y J. K. Kalita, “Network anomaly detection: Methods, systems and tools,”, 2014.

[43] P. Garcia-Teodoro, J. Diaz-Verdejo, G. M. Fernández y E. Vázquez, “Anomaly-based network intrusion detection: Techniques, systems and challenges,”, 2009.

[44] P. Laskov, P. Düssel, K. Schäfer y C. Rieck, Learning intrusion detection: Supervised or unsupervised, 2005.

[45] T. M. Pattewar and H. A. Sonawane, “Neural network based intrusion detection using Bayesian with PCA and KPCA feature extraction,”, 2015.

[46] I. Kang, M. K. Jeong y D. Kong, “A differentiated one-class classification method with applications to intrusion detection,”, 2012.

[47] R. Xu and D. Wunsch, “Survey of clustering algorithms,”, 2005.

[48] E. Vasilomanolakis, S. Karuppayah, M. Mühlhäuser y M. Fischer, “Taxonomy and survey of collaborative intrusion detection,”, 2015.

[49] Powers y D. M. W., “Evaluation: From precision, recall and F-measure to ROC informedness, markedness and correlation, 2011.

[50] T. Fawcett, “An introduction to ROC analysis,”, 2006.

[51] W. I. s. a. a. Center, «15 Cybersecurity fundamentals for water and wastewater utilities. best practices to reduce exploitable weaknesses,» *Water Iscac,* p. 56, 2019.

[52] Gartner, «7 Macro Factors That Will Shape the 2020s,» [En línea]. Available: https://www.gartner.com/en.

[53] Solutions, «https://solutionsreview.com/security-information-event-management/security-information-event-management-vendor-map/,» [En línea]. Available: https://solutionsreview.com/security-information-event-management/security-information-event-management-vendor-map/. [Último acceso: 14 Diciembre 2020].

[54] DiSIEM, I. Galan-Corroto, E. Robla, S. Prieto-Perez, A. Gonzalez-Zarzosa, A. Bessani, J. Respicio, L. Alves y A. Ferreira, Turkay: In-depth analysis of SIEMs extensibility; DiSIEM technical report, Portugal, 2017.

[55] M. Nicolett y K. Kavanagh, «Magic Quadrant for Security Information and Event Management, Gartner Technical Report,» [En línea]. Available: http://docplayer.net/2407833-Magic-quadrant-for-security-information-and-event-management.html. [Último acceso: 10 Noviembre 2010].

[56] Y. K. Sung Jun Son, «Performance of ELK Stack and Commercial System in Security Log Analysis.,» *IEEE,* 2017.

[57] A. Talaş, F. Pop y G. Neagu, Elastic stack in action for smart cities: making sense of Big data, 2017.

[58] D. P. M. Ibrahim Yahya Mohammed AL-Mahbashi, «Network Security Enhancement through Effective Log Analysis Using ELK,» de *International Conference on Computing Methodologies and Communication.*, 2017.

[59] Elasticsearch B.V, «Elastic,» 2021. [En línea]. Available: https://www.elastic.co/es/what-is/elk-stack. [Último acceso: Septiembre 2021].

[60] T. o. I. s. Foundation, «Suricata,» 2021. [En línea]. Available: https://suricata.io/. [Último acceso: 10 septiembre 2021].

[61] Canonical, «Ubuntu,» 2021. [En línea]. Available: https://ubuntu.com/. [Último acceso: Septiembre 2021].

[62] oisf y suricata, «github,» [En línea]. Available: https://github.com/OISF/suricata/blob/master/doc/userguide/install.rst.

[63] Kalsin, Vadym y J. Ellingwood, «Digital ocean community,» [En línea]. Available: https://www.digitalocean.com/community/tutorials/how-to-install-elasticsearch-logstash-and-kibana-elastic-stack-on-ubuntu-18-04-es.

[64] Rediris, «Sistema de detencion de intrusos,» 12 11 2011. [En línea]. Available: https://www.rediris.es/cert/doc/unixsec/node26.html.

[65] Sanchez Martinez, R. Fernandez, Y. García Moran y J. Paul, Hacking y Seguridad en Internet, RA-MA, 2011.

[66] L. A. Long, Profiling Hackers, Retains Full Rights, 2012.

[67] Cisco, «¿Qué es un Firewall?,» [En línea]. Available: https://www.cisco.com/c/es_mx/products/security/firewalls/what-is-a-firewall.html.

[68] Elasticsearch B.V, «Elastic,» [En línea]. Available: https://www.elastic.co/guide/en/elasticsearch/reference/current/targz.html.

[69] J. F. R. Buendia, Seguridad Informática, España: Macgraw Hill , 2013.

[70] R. Rahman y D. Tomar, Security attacks on wireless networks and their detection techniques., Sigapore, 2018.

[71] E. Maiwald, Network Security: A Beginner’s Guide, McGraw-Hill, 2001.

[72] D. Miller, S. Harris, A. Harper, S. Van Dyke y C. Blask, Security Information and Event Management (SIEM) Implementation, Graw Hill, 2010.

[73] S. Review, «https://solutionsreview.com/security-information-event-management/security-information-event-management-vendor-map/,» [En línea]. Available: https://solutionsreview.com/security-information-event-management/security-information-event-management-vendor-map/. [Último acceso: 14 Diciembre 2020].

[74] N. Azeez, T. Bada, S. Misra, A. Adewumi, C. der Vyver y R. Ahuja, Intrusion detection and prevention systems: An updated review,. In data management, analytics and Innovation, Sigapore, 2020.

[75] M. Bajer, Building an IoT data hub with Elasticsearch, Logstash and Kibana, 2017.

[76] G. G. Escriva, Seguridad Informática, Macmillan Iberial S.A, 2013.

[77] W. Stallings, Network security: Applications and standards, Prentice Hall, 2011.

[78] C. Santos Jesus, Seguridad Informática, RAM-MA Editorial, 2015.

ANEXO A: INSTALACIÓN DEL SISTEMA OPERATIVO UBUNTU

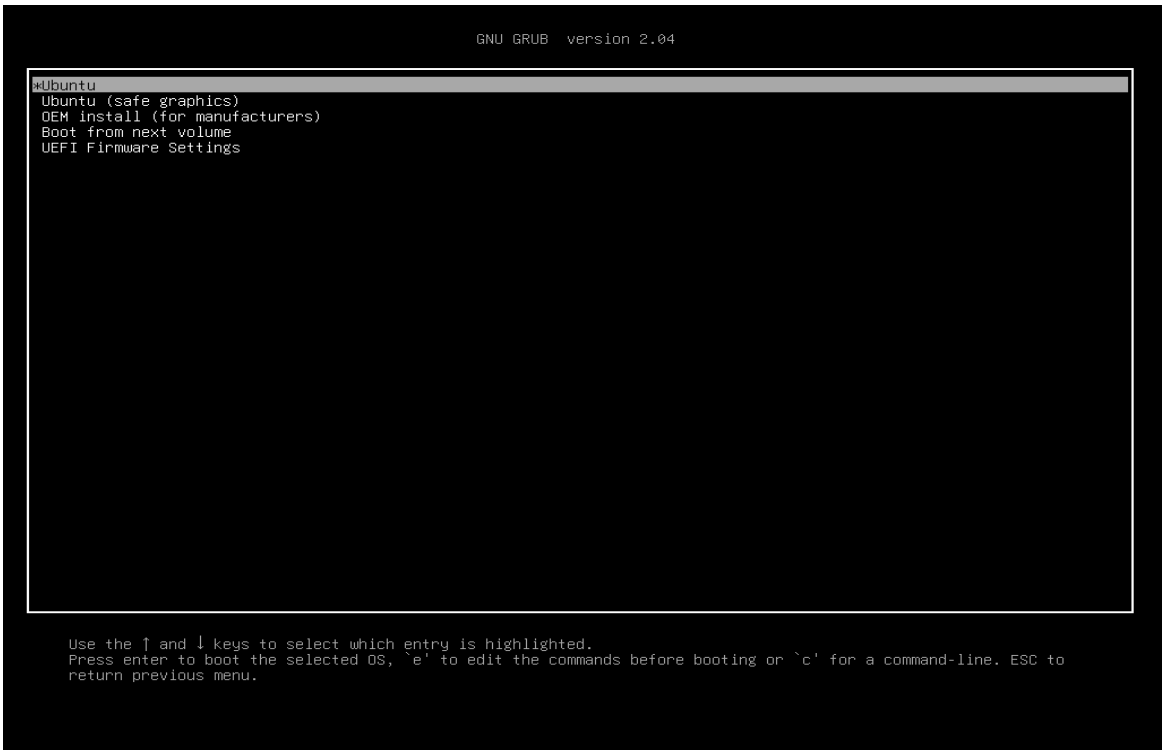


Ilustración 34: Pantalla Instalación Ubuntu

Al iniciar la lectura de nuestro Dvd o Usb donde tenemos el boot del sistema operativo Ubuntu, tendremos la siguiente ventana, dando enter sobre la primera línea.

Ilustración 35: Opción Idioma

Escogeremos del siguiente menú, el idioma de preferencia y daremos instalar Ubuntu.

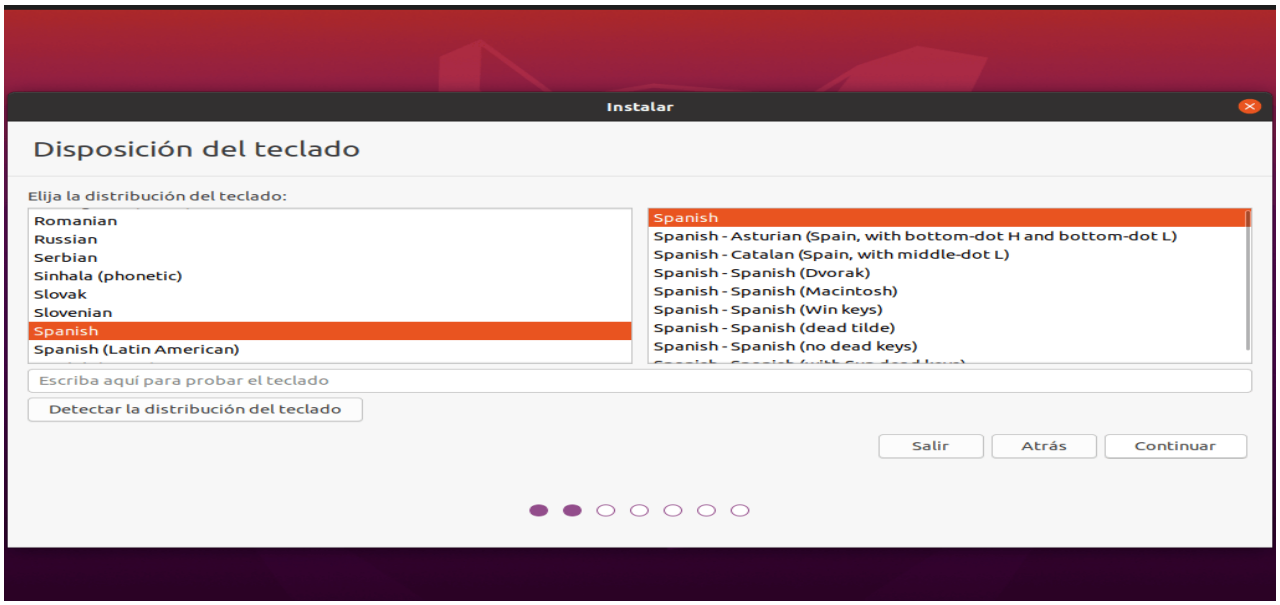


Ilustración 36: Distribución Idioma del Teclado

Escogemos la distribución del teclado, esto dependerá del tipo que tenga nuestro equipo de cómputo.

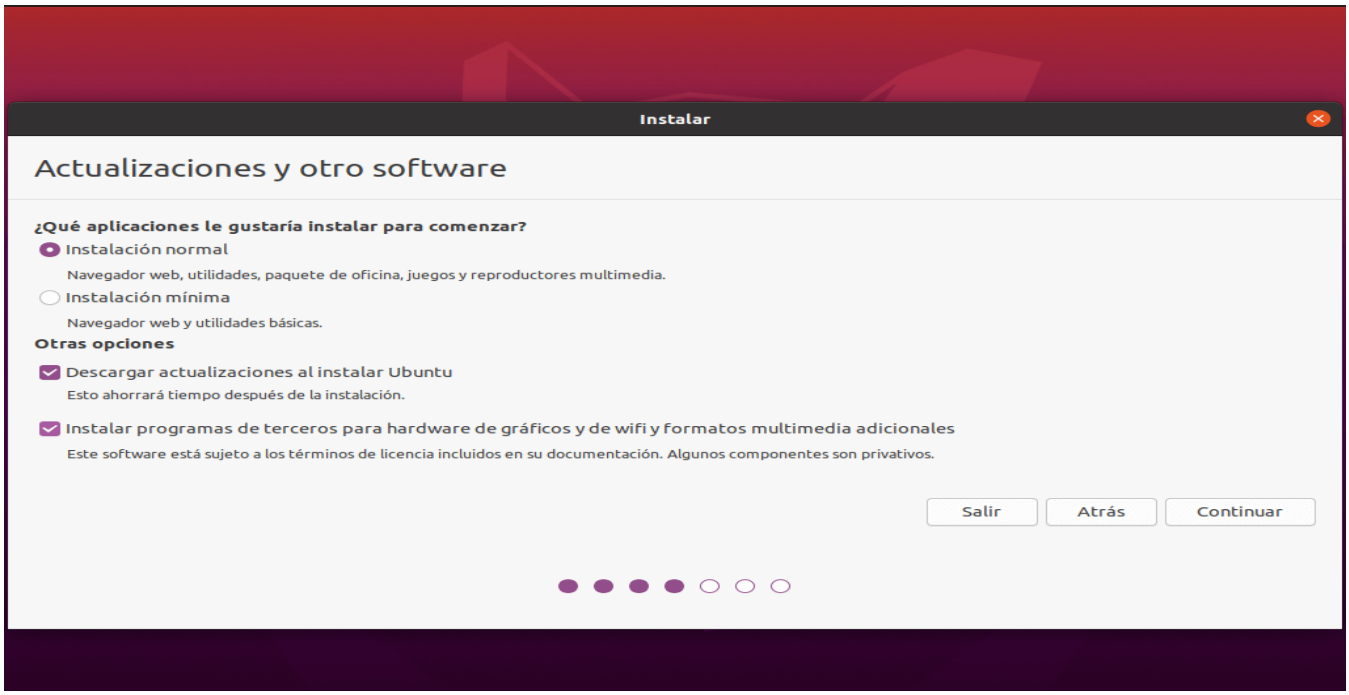


Ilustración 37: Instalación y Actualizaciones

En este apartado, nos pedirá el tipo de instalación, normal, mínima o completa, de igual forma actualizar el sistema mientras instala en nuestro equipo.

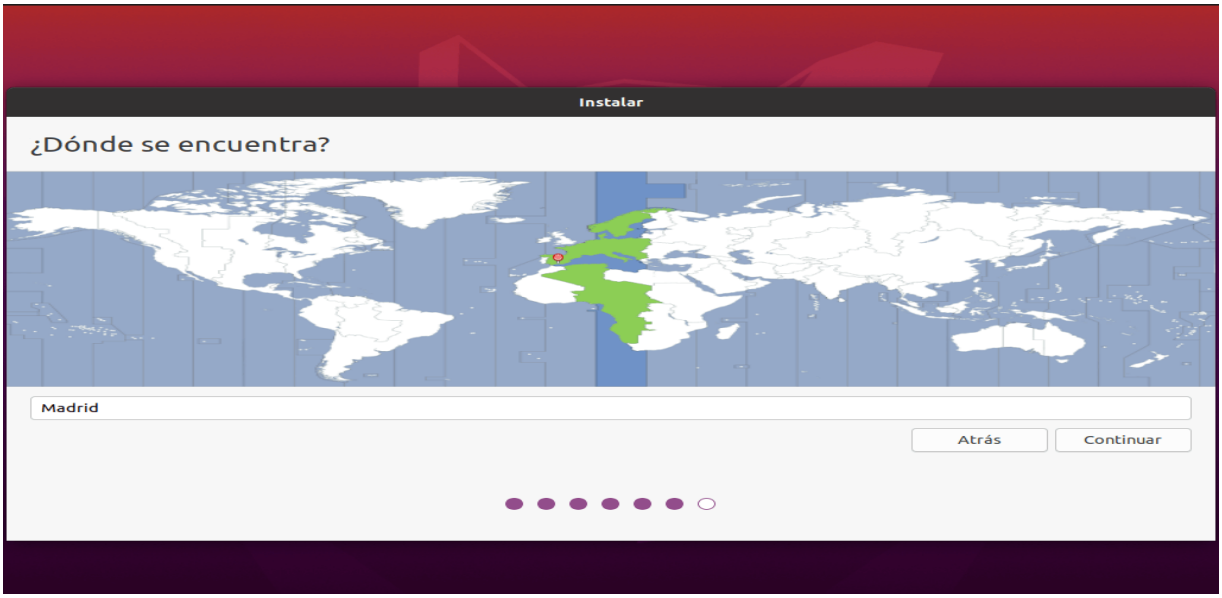


Ilustración 38: Región Geográfica

Escogemos la ubicación geográfica para la configuración de uso horario.

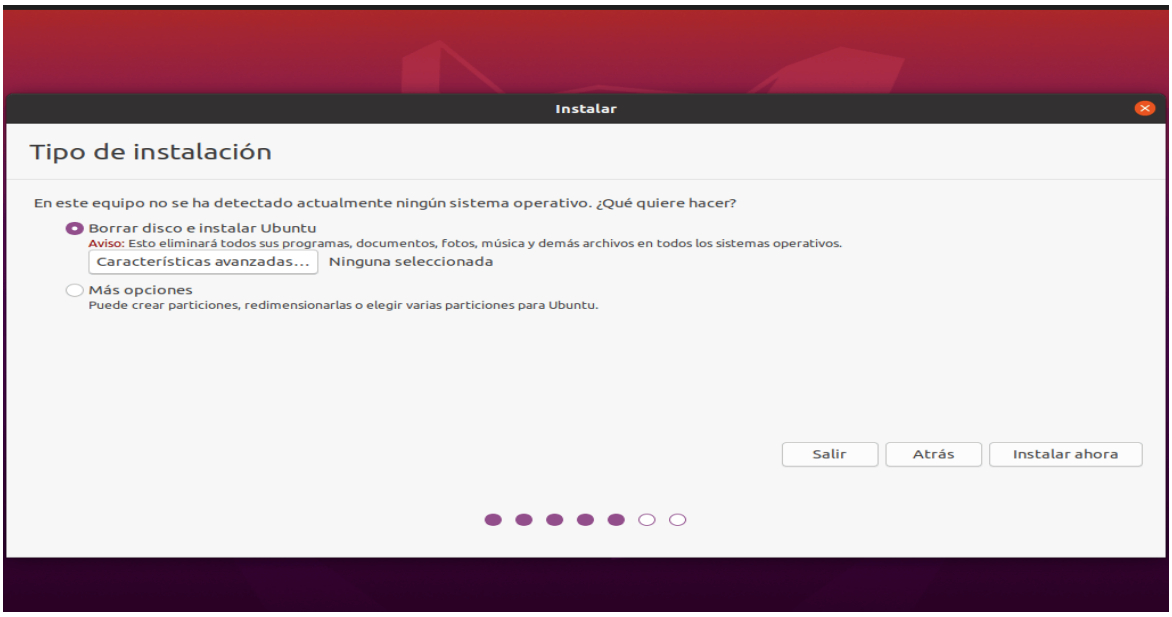


Ilustración 39: Particiones y Disco Duro

Ahora pedirá escoger la unidad de disco duro para la instalación, dando la opción de crear particiones sin en el caso compartiremos el HDD con algún sistema Windows, si no es así daremos en instalar sistema Ubuntu completamente.

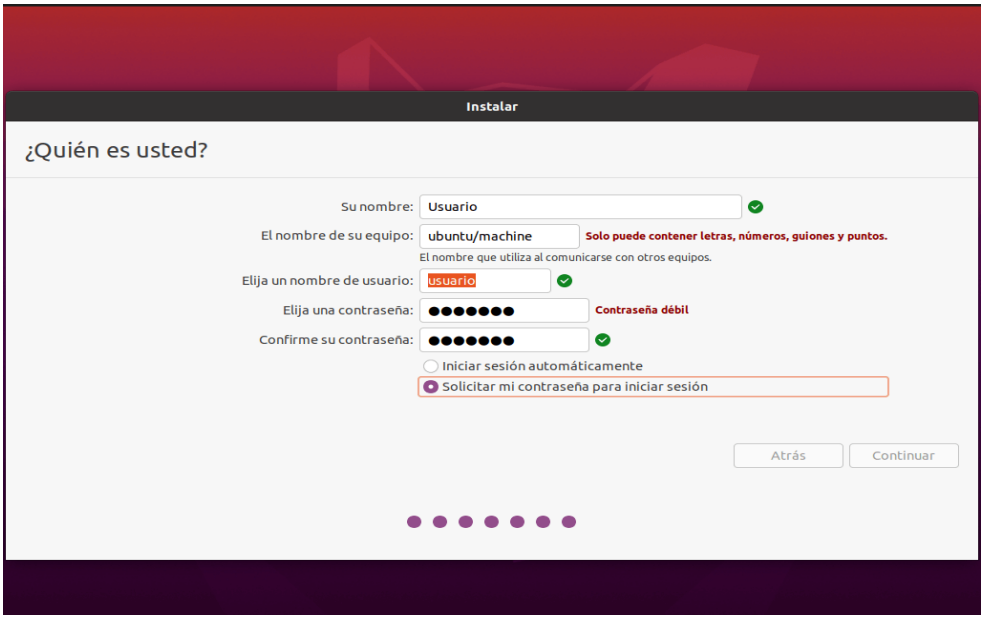


Ilustración 40: Creación de Usuario

Crearemos un usuario para operar nuestro sistema.

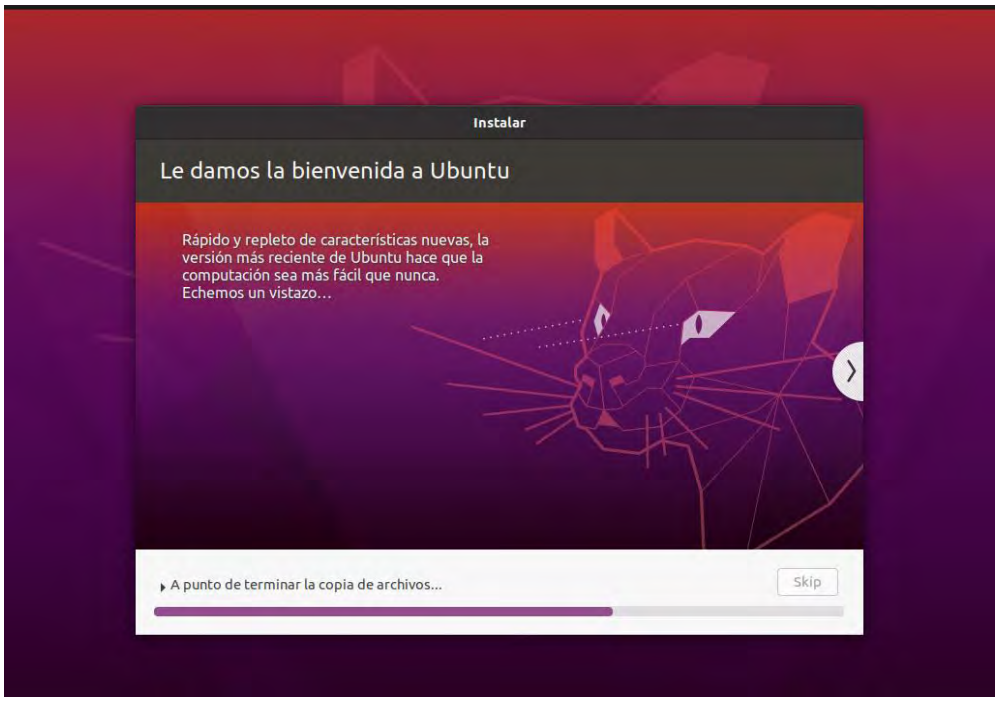


Ilustración 41: Carga de Archivos de Sistema

Esperamos que se cargue todo lo necesario para la configuración, actualización e instalación del sistema Ubuntu.

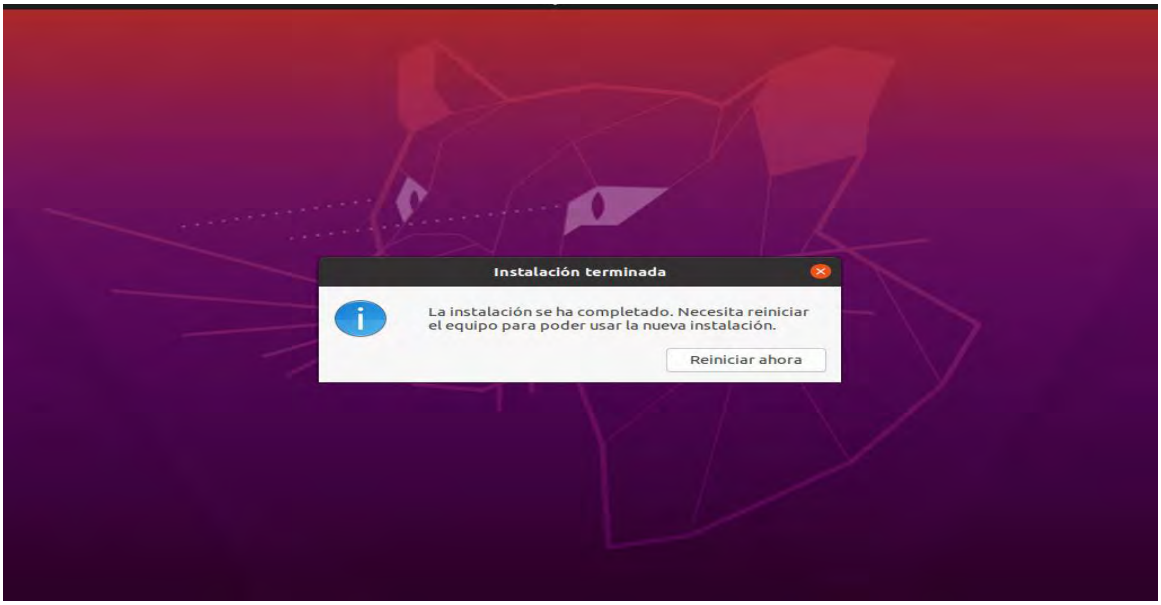


Ilustración 42: Instalación Terminada

Al terminar la instalación nos pedirá reiniciar el equipo para poder iniciar la interfaz del sistema.

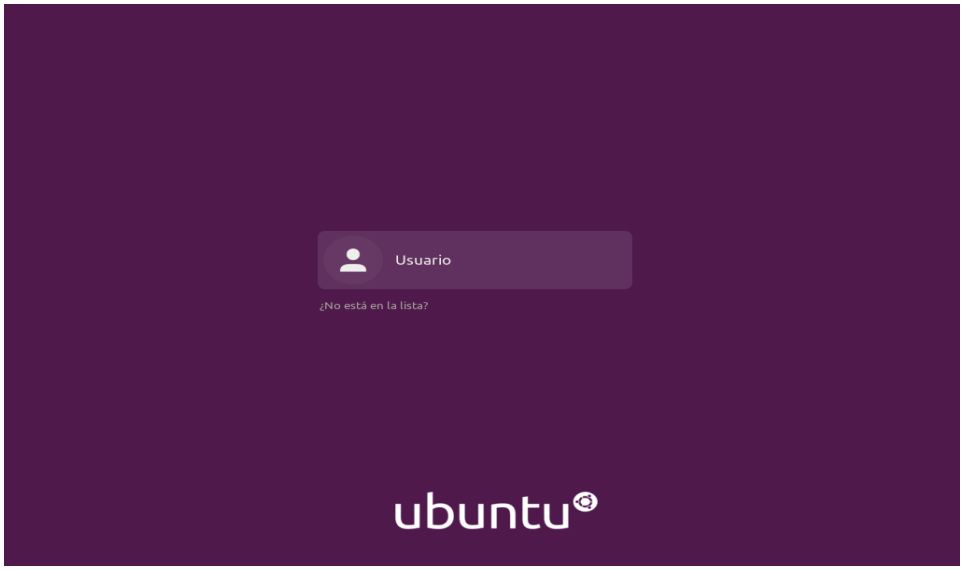


Ilustración 43: Pantalla Inicio de Sesión

Nos autentificamos para iniciar sesión.

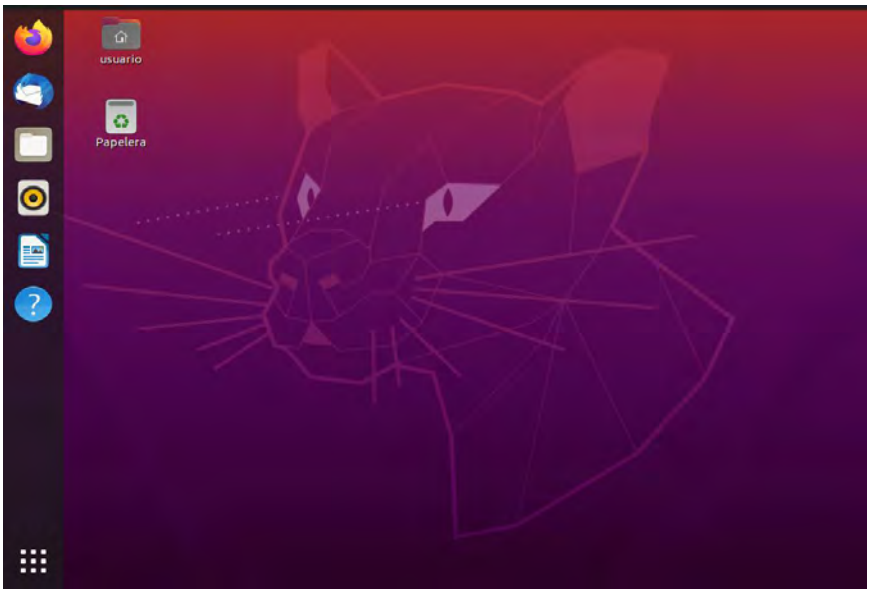


Ilustración 44: Escritorio Ubuntu 20.04

Bienvenido al sistema operativo Ubuntu, ahora queda hacer las configuraciones básicas como actualizar repositorios etc.

ANEXO B: INSTALACIÓN DE UN IDS

Antes de instalar el software IDS (SURICATA), tendremos que considerar los prerrequisitos necesarios para su optima operación. Dependiendo de la configuración de su red y de cómo pretenda utilizar Suricata, es posible que necesite un procesador y memoria ram para su servidor. En general, cuanto más tráfico planee inspeccionar, más recursos debe asignar a Suricata. En un entorno de producción, planee usar al menos 2 CPU y 4 u 8 GB de RAM para empezar [60]. A partir de ahí, puede escalar los recursos de acuerdo con el rendimiento de Suricata y la cantidad de tráfico que necesita procesar.

Para la instalación de nuestro servidor IDS Suricata, usaremos la terminal de comandos que nos otorga el sistema operativo Ubuntu 20.04 y ejecutaremos los comandos [60], que nos proporciona la comunidad de Suricata software.

Actualizaremos nuestros repositorios actuales con el comando Apt-get update y por consiguiente aplicaremos el comando que nos permitirá cargar los repositorios necesarios para poder instalar suricata IDS y una vez más usaremos el comando anterior para cargar el nuevo repositorio. #apt-get update.

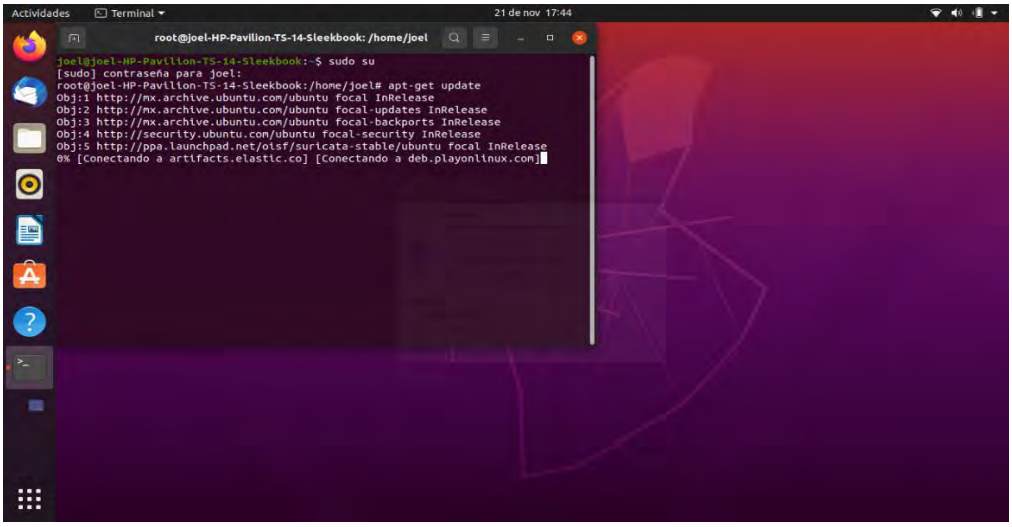


Ilustración 45: terminal actualización de repositorios

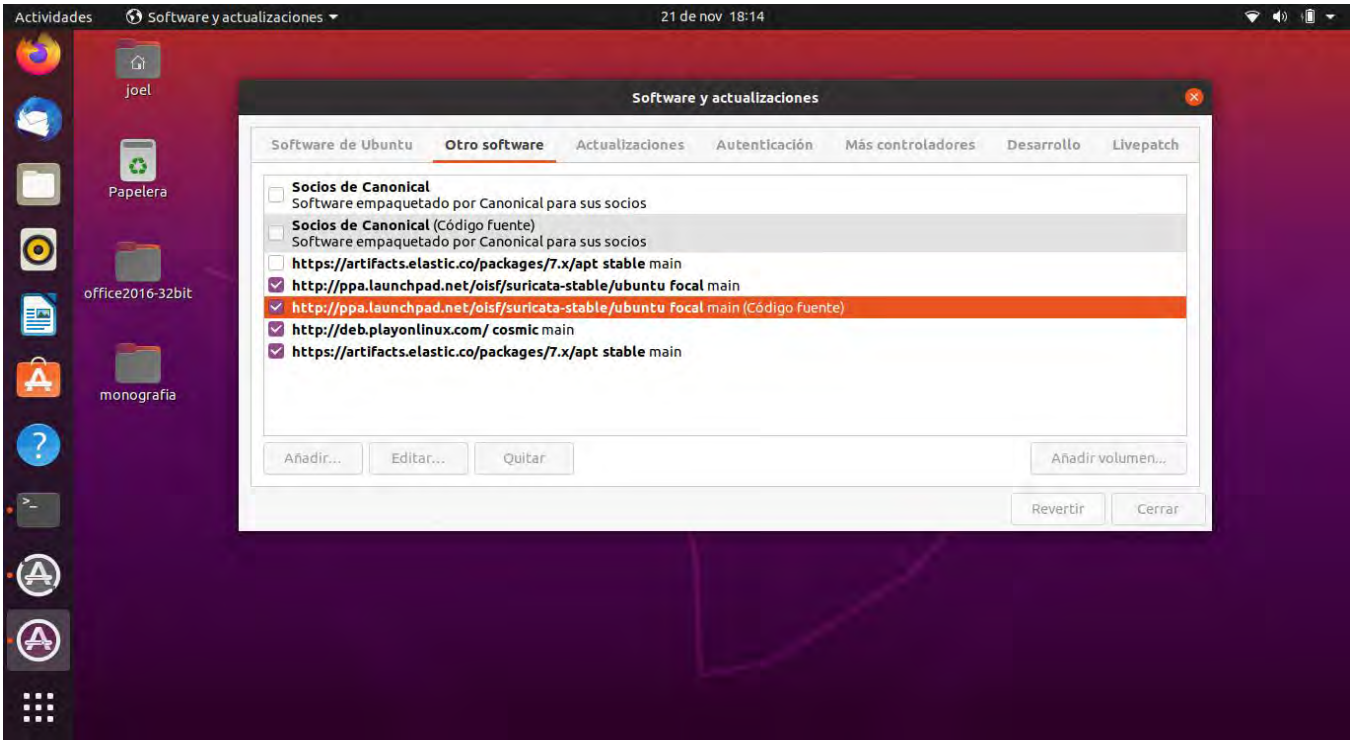


Ilustración 46: Repositorio OISF

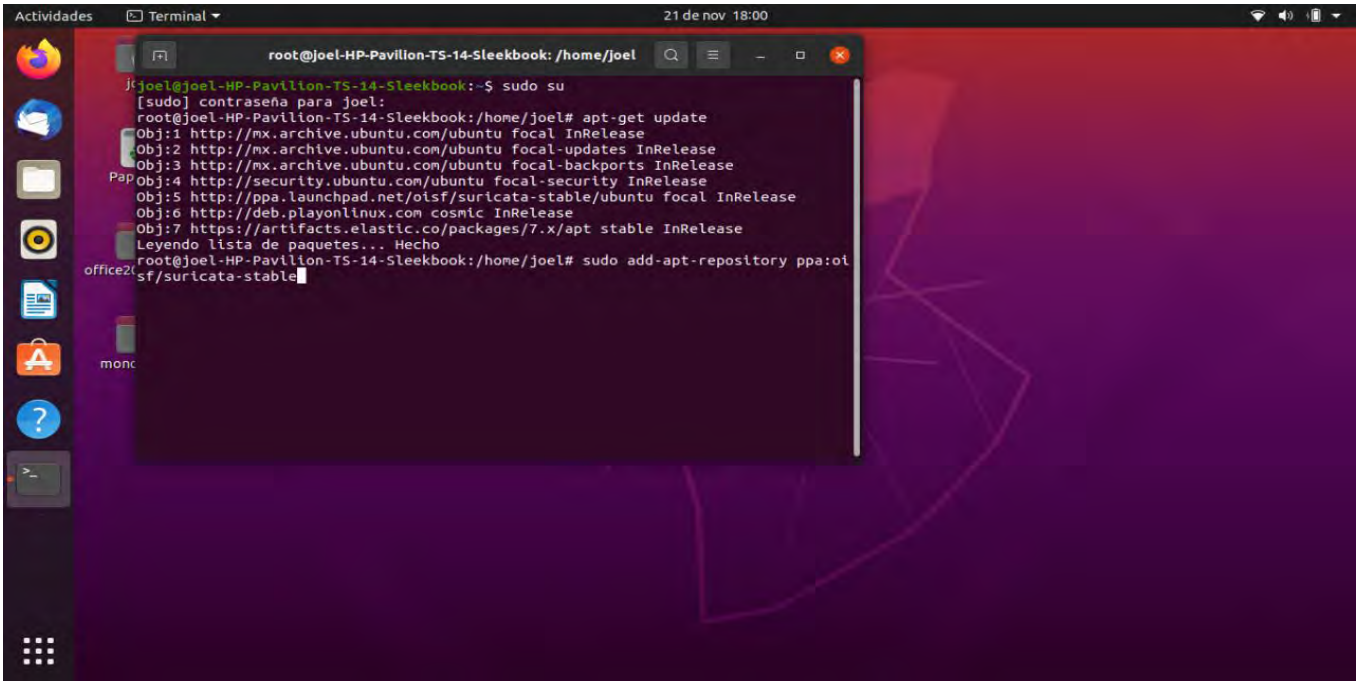


Ilustración 47: Comando Agregar Repositorio

Con el comando desde la terminal agregaremos los repositorios necesarios para la instalación de Suricata IDS: **sudo add-apt-repository ppa:oisf/suricata-stable**, y una vez más usaremos el comando **apt-get update** para instalarlo. Luego comprobaremos que se haya anexado al sistema.

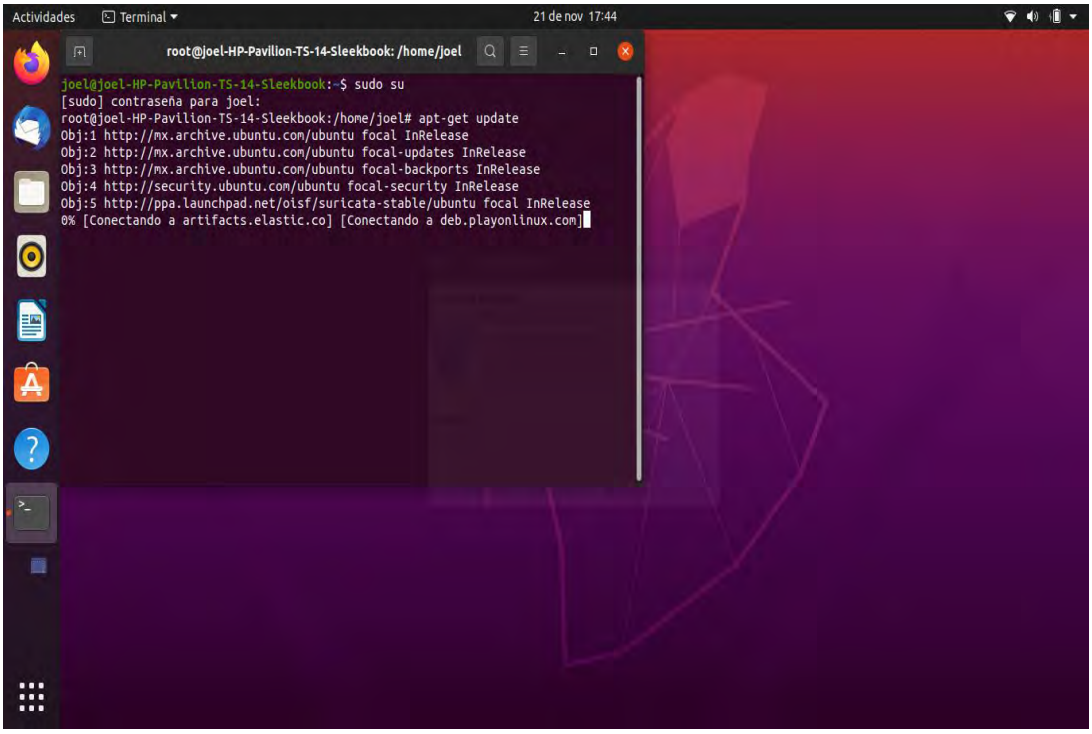


Ilustración 48: Actualización de Repositorios

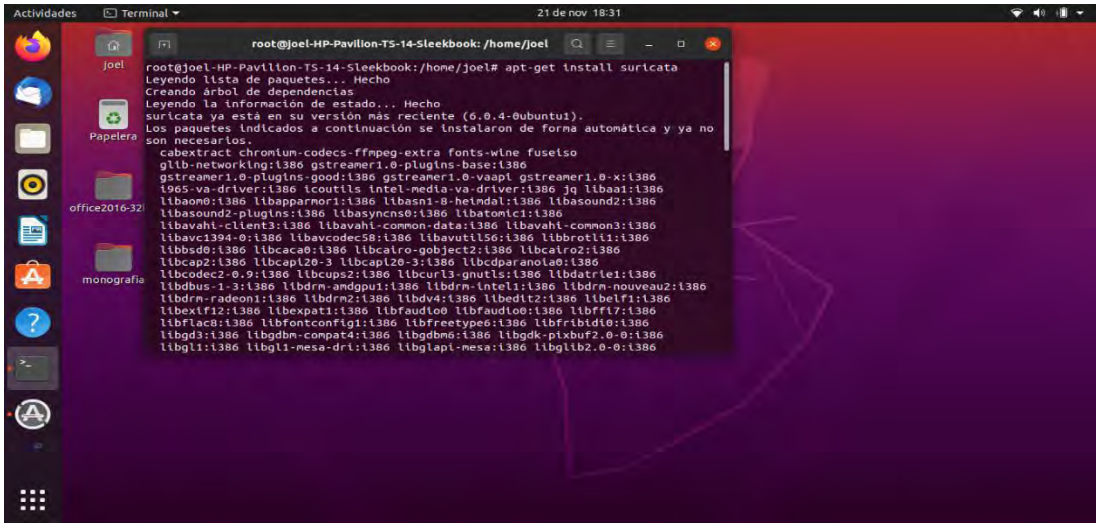


Ilustración 49: Instalación Suricata IDS

Una vez cargado los repositorios necesarios y actualizado. Insertaremos el comando siguiente para iniciar la descarga e instalación de suricata IDS con el siguiente comando:

- Apt-get install suricata.

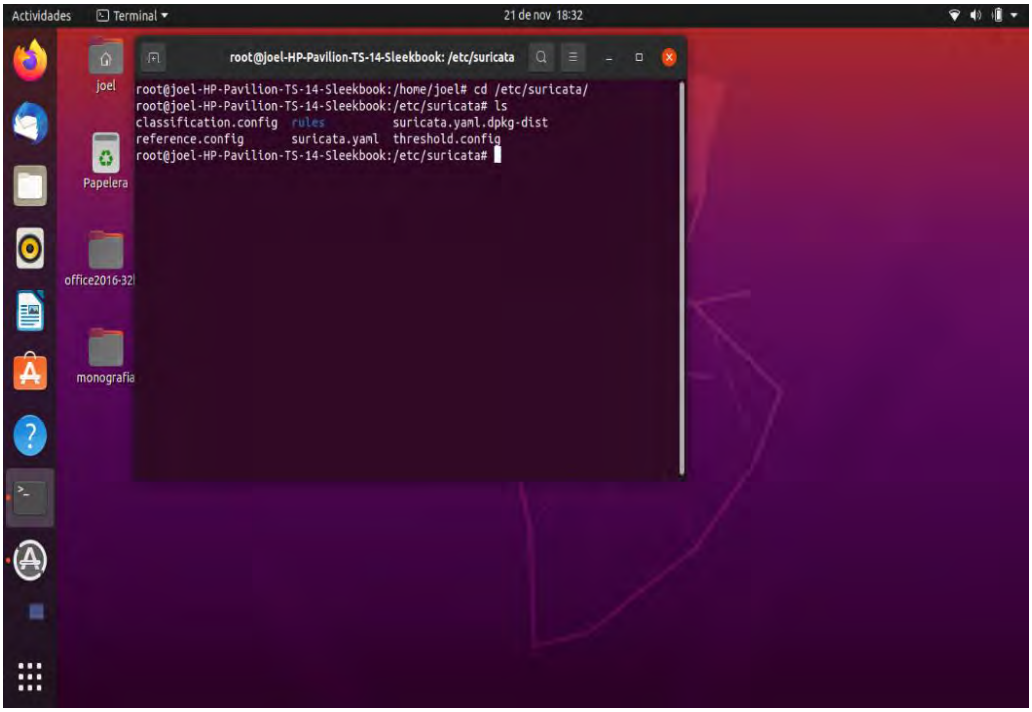


Ilustración 50: Ruta de Configuración Suricata

Ya instalado nuestro IDS iremos a la ruta de archivos de configuración de Suricata con el comando **cd /etc/suricata/,** en donde se encuentran los archivos de configuración.

ANEXO C: INSTALACIÓN DE ELASTICSEARCH

Para instalar nuestro sistema SIEMS, abriremos nuestra terminal Shell del sistema operativo Linux y agregaremos las llaves de licencia y dependencias del stack de elk [59].

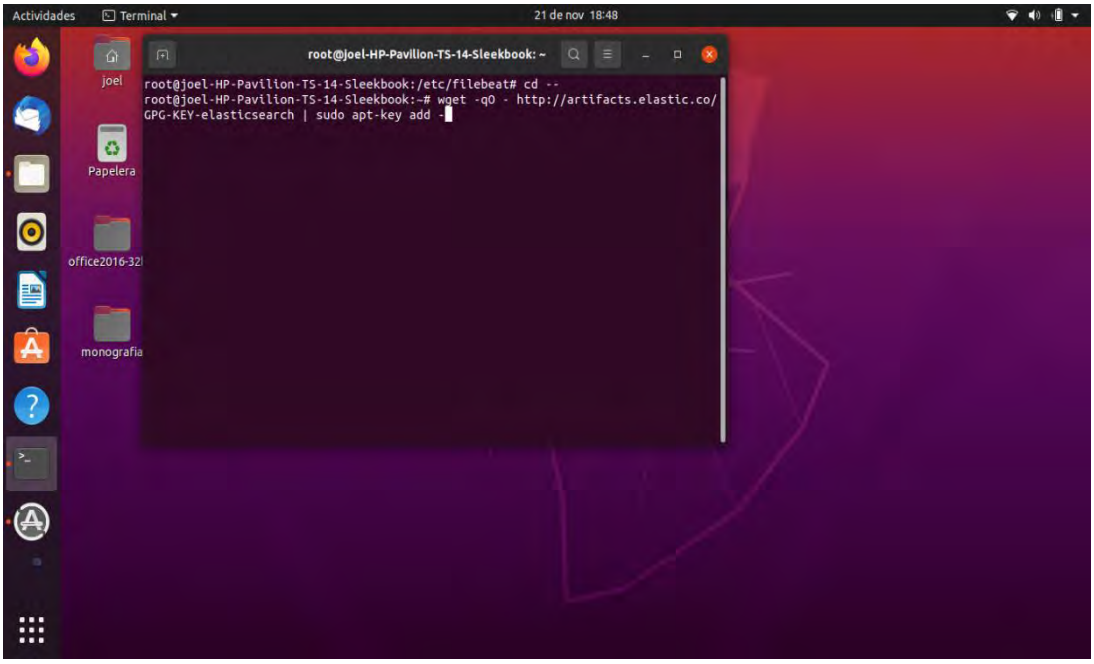


Ilustración 51: llaves y Dependencias ELK

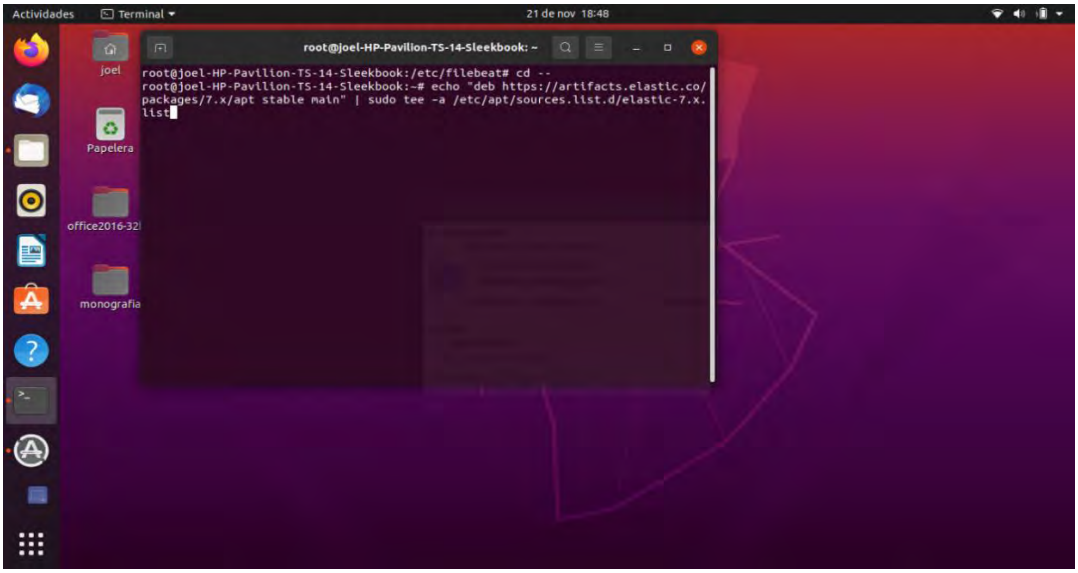


Ilustración 52: Repositorio ELK

Agregaremos los repositorios necesarios para poder instalas la paquetería de herramientas que ofrece el stack de ELK. Antes de instalar ELK, debemos tener instalado en el sistema el JDK (**java developer kit**) ya que este SIEM, está desarrollado en lenguaje de programación java. Una vez instalado verificamos que ya contemos con java.

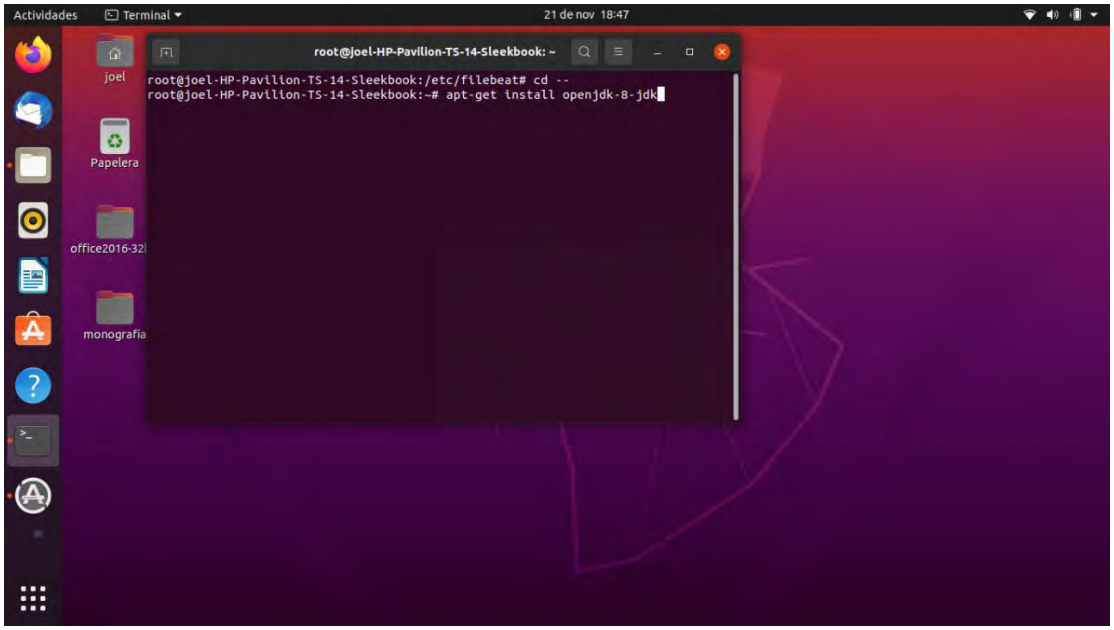


Ilustración 53: Instalación JDK

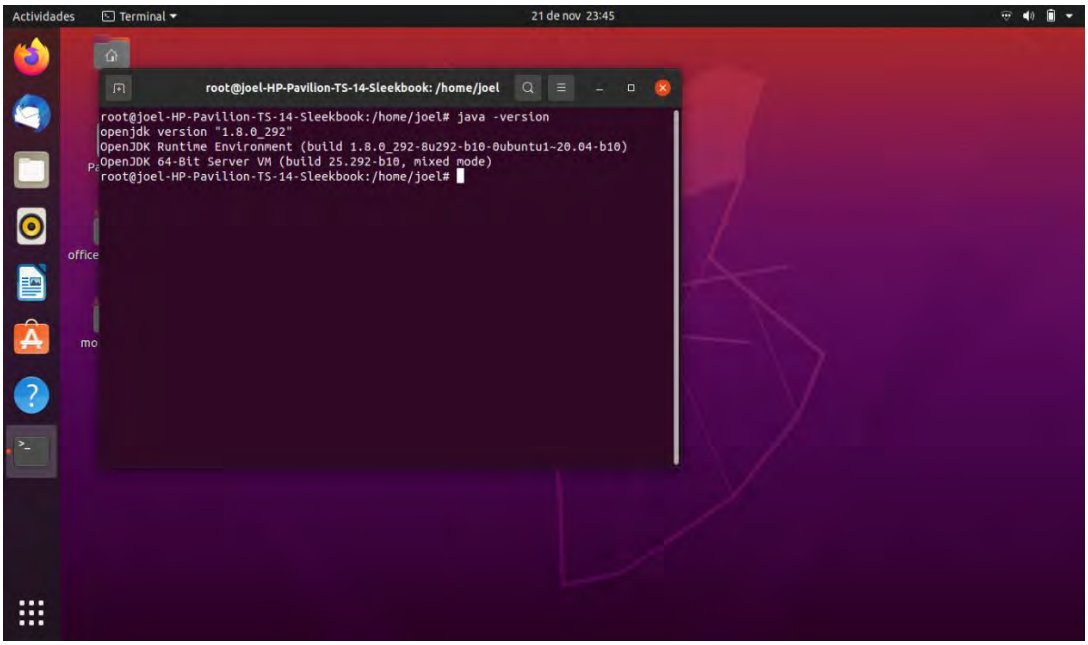


Ilustración 54: Verificación JDK

Una vez instalado el kit de desarrollo JDK usaremos el siguiente comando para instalar Elasticsearch:

- #apt-get install elasticsearch.

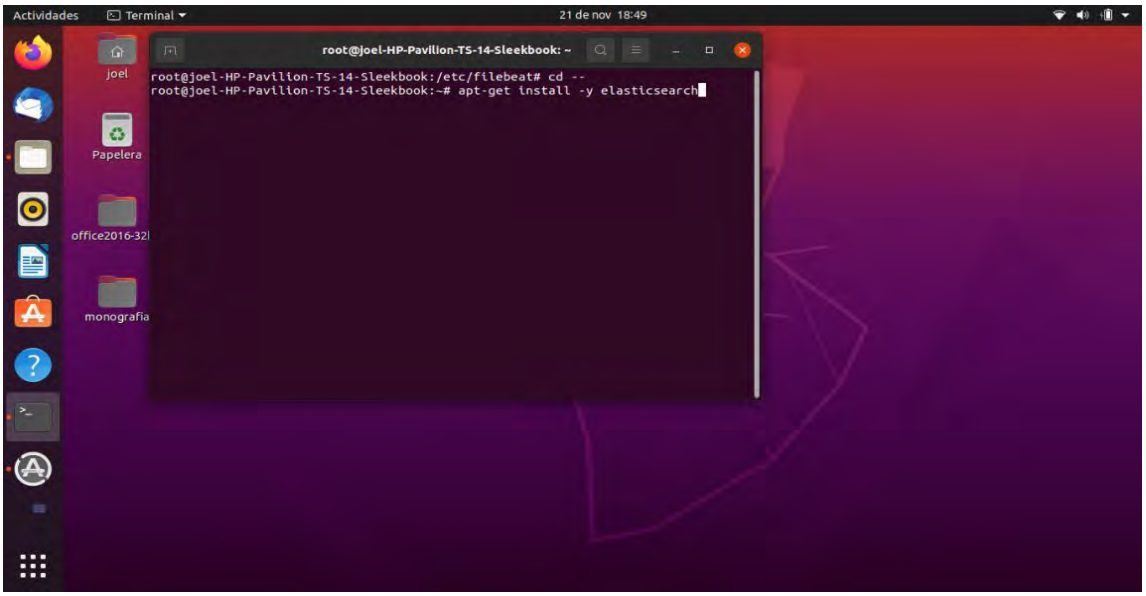


Ilustración 55: Apt-Get Install Elasticsearch

Una vez instalamos verificamos la ruta de instalación, ya verificado tendremos instalado el software de indexación. #cd /etc/elasticsearch

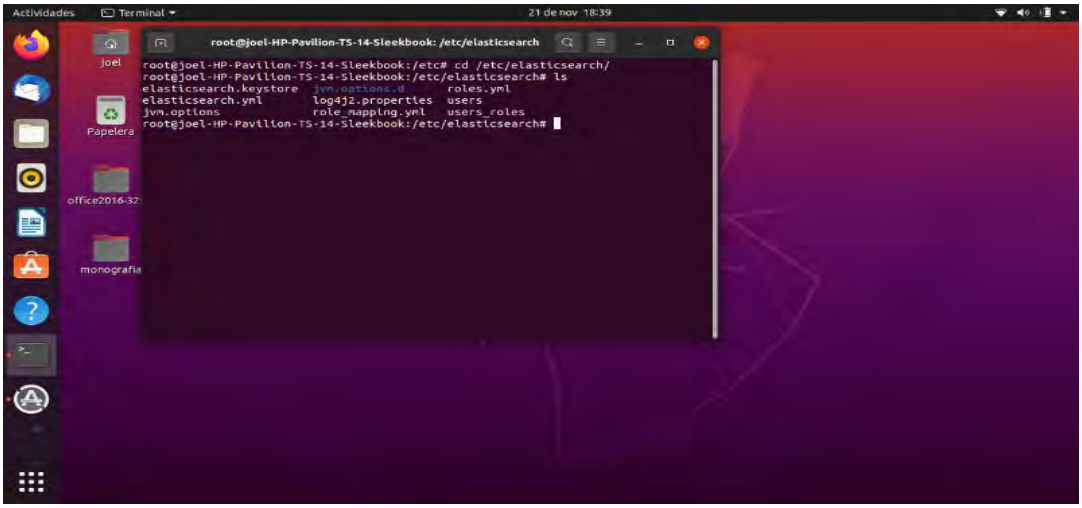


Ilustración 56: Ruta Documentos Elasticsearch

ANEXO D: INSTALACIÓN DE KIBANA

Para instalar kibana, como este pertenece al mismo Stack de ELK, ya no será necesario agregar las llaves de licencia y el repositorio, desde la terminal agregaremos el comando #apt-get install kibana.

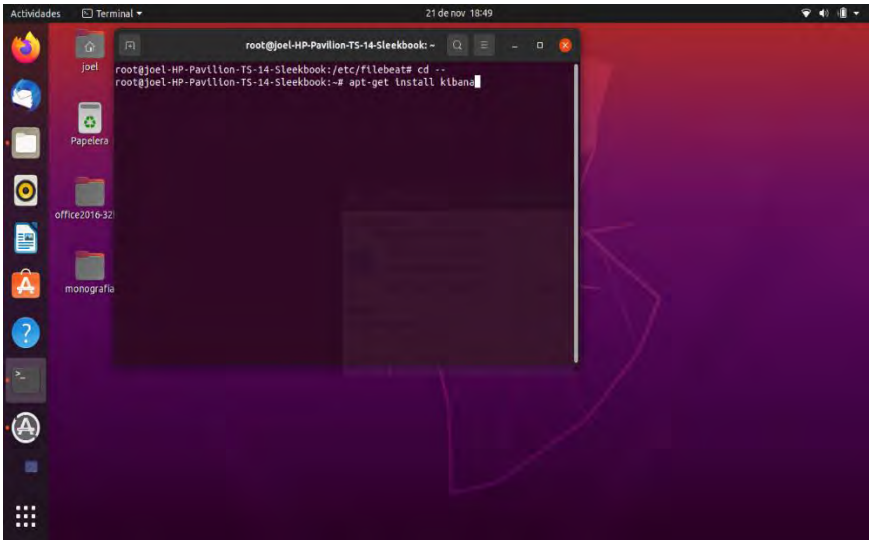


Ilustración 57: Instalación Kibana

Verificamos si se instala correctamente, ubicando la ruta de archivos de configuración con el comando: #cd /etc/kibana.

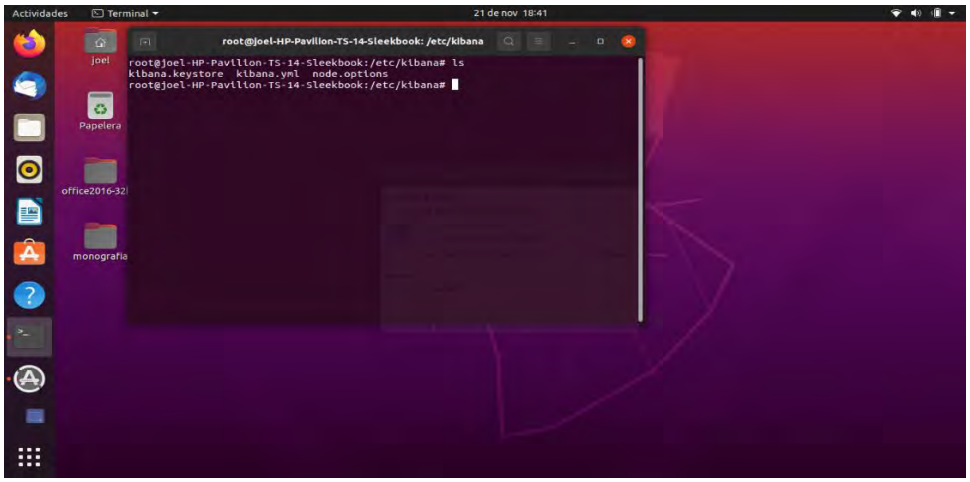


Ilustración 58: Ruta Configuración Kibana

ANEXO E: INSTALACIÓN FILEBEAT

Como mencionamos anteriormente, ya no necesitamos cargar las llaves y repositorios, instalaremos filebeat con el siguiente comando.

- Apt-get install filebeat.
- Una vez instalado corroboraremos la ruta de instalación en cd /etc/Filebeat. [59]

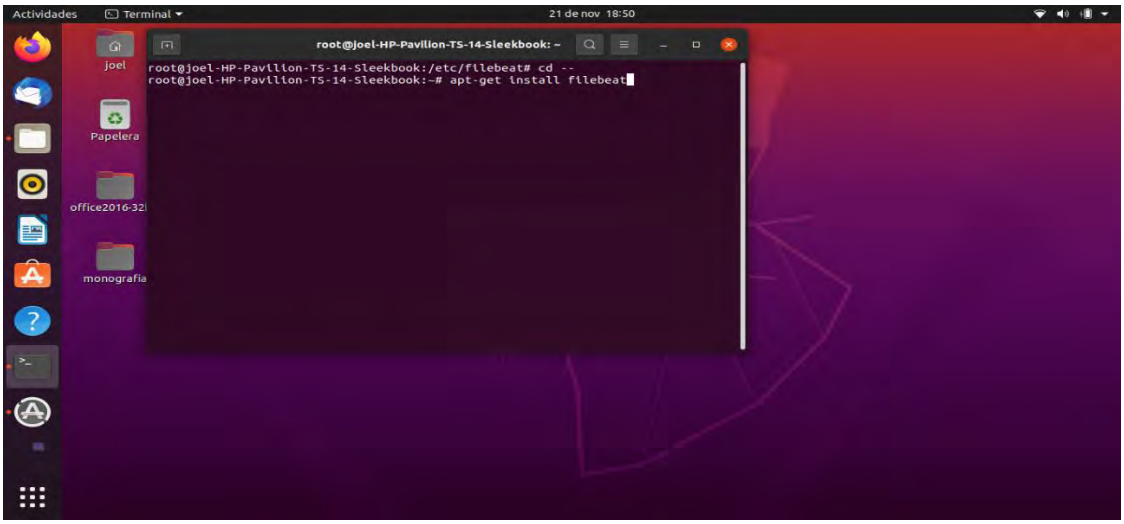


Ilustración 59: Instalando Filebeat

Ruta Filebeat : # cd /etc/filibeat

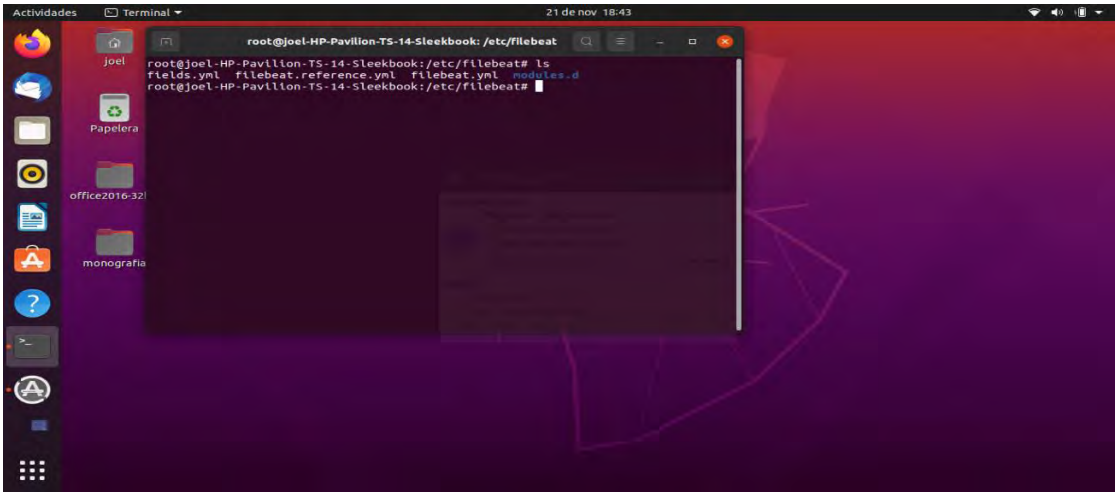


Ilustración 60: Ruta Filebeat