



Security Analytics with Elastic

David Vázquez Pesado

Máster de Seguridad en las Tecnologías de la Información y de las Comunicaciones.

Análisis de Datos.

Pau del Canto Rodrigo

Víctor García Font

02 de Junio de 2020



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada [3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

Agradecimientos

En primer lugar, quisiera dedicarle este trabajo a mi pareja, a mis padres y a mi hermana:

A Andrea, por estar cada día apoyándome, aguantando mis malos momentos y soportando mi estrés. Has sabido estar ahí durante toda esta experiencia, y me has ayudado a sacar lo mejor de mí. Aunque ha sido muy duro y difícil de llevar, el esfuerzo ha valido la pena y se verá recompensado. Gracias por saber estar en las buenas, pero mejor aún por saber estar en las malas. Sin ti, esto no habría sido posible. Gracias por formar parte de mi vida, y por hacerla mejor cada día.

A mis padres, por insistirme cada día a que continúe con mi formación y por proporcionarme los medios y métodos para poder conseguirlo, pero sobre todo gracias por vuestra paciencia conmigo. Sin vuestro apoyo, lucha y confianza tengo claro que no llegaría hasta aquí.

Gracias por seguir enseñándome a ser mejor persona.

A mi hermana Laura, gracias por apoyarme desde la distancia, por confiar en mí, por preocuparte por mí y por tus ánimos constantes. Espero que mi experiencia te sirva de ayuda a ti también. Estoy seguro de que tú también conseguirás alcanzar las metas que te propongas. Gracias por formar parte de mi vida.

Por último, agradecer también a mis amigos, que siempre han estado ahí para escucharme y apoyarme en los momentos más difíciles, y me han sacado una sonrisa.

FICHA DEL TRABAJO FINAL

| | |
|--|---|
| Título del trabajo: | <i>Security Analytics with Elastic</i> |
| Nombre del autor: | <i>David Vázquez Pesado</i> |
| Nombre del consultor/a: | <i>Pau del Panto Rodrigo</i> |
| Nombre del PRA: | <i>Víctor García Font</i> |
| Fecha de entrega (mm/aaaa): | 06/2020 |
| Titulación: | <i>Máster universitario de Seguridad de Tecnologías de la Información y de las Comunicaciones</i> |
| Área del Trabajo Final: | <i>Trabajo Fin de Máster – Análisis de Datos</i> |
| Idioma del trabajo: | <i>Español</i> |
| Palabras clave | <i>Elastic, SIEM, Monitorización, Machine Learning.</i> |
| Resumen del Trabajo (máximo 250 palabras): <i>Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.</i> | |
| <p>Este proyecto nace para cubrir la creciente necesidad en las empresas en cuanto a la implementación de sistemas de control y monitorización de seguridad, que permitan mejorar la seguridad de los activos y de la infraestructura en los entornos empresariales.</p> <p>El proyecto se fundamenta en su totalidad en la Suite de Elastic, siendo un conjunto de herramientas que permiten un análisis sencillo de las fuentes de información existentes en los entornos empresariales, de las cuales se extrae información de seguridad, como por ejemplo DNS, Netflow, Eventos del sistema, Auditoria, etc.</p> <p>El objetivo consiste en utilizar las herramientas que proporciona Elastic para el análisis de los datos empresariales, para posteriormente integrarlos con Elastic SIEM y, mediante técnicas de Machine Learning, conseguir mejorar la prevención y detección de amenazas de seguridad.</p> <p>Esto es realmente útil para visualizar en tiempo real el estado de los activos y de la infraestructura, lo que permite mantener un control total de lo que está sucediendo en cada instante. Esto ayuda, en caso necesario, a una pronta actuación para la prevención o mitigación de una amenaza de seguridad.</p> <p>El producto final se convierte en un sistema integral de seguridad, que permite no solo controlar posibles amenazas de seguridad, sino que también proporciona información útil a través de la continua monitorización de los activos, pudiendo conocer el estado actual de la infraestructura en tiempo real.</p> | |

Abstract (in English, 250 words or less):

This project was created to cover the growing need in companies for the implementation of control and security monitoring systems, which allow for the improvement of asset and infrastructure security in business environments.

The project is based entirely on the Elastic Suite, a set of tools that allow a simple analysis of existing information sources in business environments, from which security information is extracted, such as DNS, Netflow, System Events, Audit, etc.

The objective is to use the tools provided by Elastic for the analysis of business data, to later integrate them with Elastic SIEM and, by means of Machine Learning techniques, achieve improved prevention and detection of security threats.

This is really useful to visualize in real time the state of the assets and the infrastructure, which allows to keep a total control of what is happening in every moment. This entails, if necessary, to prompt action for the prevention or mitigation of a security threat.

The final product becomes a comprehensive security system, which allows not only to control possible security threats, but also provides useful information through continuous monitoring of assets, being able to know the current status of the infrastructure in real time.

Índice

| | |
|---|----|
| 1. Introducción | 1 |
| 1.1 Contexto y justificación del Trabajo | 1 |
| 1.2 Objetivos del Trabajo | 1 |
| 1.3 Enfoque y método seguido | 2 |
| 1.4 Planificación del Trabajo | 3 |
| 1.5 Estado del Arte | 5 |
| 1.6 Breve resumen de productos obtenidos | 8 |
| 1.7 Presupuesto | 8 |
| 1.8 Viabilidad y Análisis de riesgos | 9 |
| 1.9 Breve descripción de los otros capítulos de la memoria | 10 |
| 2.0 Elastic Stack | 11 |
| 2.1 Definición de Elastic Stack | 11 |
| 2.2 Módulos de Elastic | 14 |
| 2.3 Arquitectura del proyecto | 20 |
| 2.3.1 Arquitectura de Elastic | 20 |
| 2.3.2 Arquitectura del laboratorio | 21 |
| 3.0 Elastic SIEM | 23 |
| 3.1 Definición de SIEM | 23 |
| 3.2 Integración de eventos | 25 |
| 3.2.1 Integración y normalización de eventos de FortiWeb | 26 |
| 3.2.2 Integración de Eventos del S.O | 30 |
| 3.2.3 Integración y normalización de Eventos de Cisco Umbrella | 30 |
| 3.2.4 Integración de Eventos de Palo Alto | 34 |
| 3.3 Configuración y uso de Elastic SIEM | 35 |
| 4.0 Machine Learning | 47 |
| 4.1 Definición de Machine Learning | 47 |
| 4.2 Configuración y uso de Machine Learning | 48 |
| 4.3 Detección de Anomalías | 51 |
| 5.0 Elastic Watcher | 65 |
| 5.1 Definición del plugin Watcher | 65 |
| 5.2 Configuración y uso de Watcher | 67 |
| 6.0 Hardening de Elastic | 74 |
| 7.0 Monitorización con Elastic | 76 |
| 7.1 Estado de salud de los servidores mediante métricas | 77 |
| 7.2 Controlar la disponibilidad de servicios | 78 |
| 7.3 Análisis de logs de Linux | 78 |
| 7.4 Análisis de logs de Windows | 79 |
| 7.5 Control de logins | 80 |
| 7.6 Control de usuarios | 81 |
| 7.7 Control de integridad de ficheros | 81 |
| 7.8 Control de conexiones SSH | 82 |
| 7.9 Control de procesos | 83 |
| 7.10 Flow de paquetes | 84 |
| 8.0 Conclusiones | 86 |
| 8.1 Líneas de trabajo futuro | 89 |

| | |
|--|-----|
| 9.0 Glosario | 90 |
| 10.0 Bibliografía | 92 |
| 11.0 Anexos | 94 |
| 11.1 Instalación y configuración de ElasticSearch, Kibana y Logstash | 94 |
| 11.1.1 Instalación de Elasticsearch | 94 |
| 11.1.2 Configuración de Elasticsearch | 95 |
| 11.1.3 Instalación de Kibana | 96 |
| 11.1.4 Configuración de Kibana | 97 |
| 11.1.5 Instalación de Logstash | 97 |
| 11.2 Instalación y configuración de MetricBeat | 98 |
| 11.3 Añadir logs del sistema en Linux a Elasticsearch | 101 |
| 11.4 Añadir Eventos de Windows a ElasticSearch | 102 |
| 11.5 Instalación y configuración de Heartbeat | 103 |
| 11.6 Instalar y configurar AuditBeat | 105 |
| 11.7 Establecer las comunicaciones de Elasticsearch y Kibana mediante SSL/TLS | 107 |
| 11.8 Instalar y configurar packetbeat | 109 |
| 11.9 Creación de usuarios y roles en Kibana | 110 |

Lista de figuras

| | |
|---|----|
| Ilustración 1: Tablero Trello | 3 |
| Ilustración 2: Tareas del proyecto – 1 | 4 |
| Ilustración 3: Tareas del proyecto - 2 | 4 |
| Ilustración 4: Diagrama de Gantt - 1 | 5 |
| Ilustración 5: Diagrama de Gantt - 2 | 5 |
| Ilustración 6: Funcionamiento de Logstash | 12 |
| Ilustración 7: Funcionamiento de Filebeat | 15 |
| Ilustración 8: Packetbeat flow | 16 |
| Ilustración 9: Arquitectura Elastic Stack | 20 |
| Ilustración 10: Topología de red del laboratorio | 22 |
| Ilustración 11: Arquitectura del laboratorio | 22 |
| Ilustración 12: Arquitectura del SIEM | 24 |
| Ilustración 13: Eventos de Fortiweb | 27 |
| Ilustración 14: Eventos de FortiWeb en SIEM | 29 |
| Ilustración 15: Eventos de Cisco Umbrella | 31 |
| Ilustración 16: Eventos de Cisco Umbrella en SIEM | 33 |
| Ilustración 17: Eventos de PaloAlto | 35 |
| Ilustración 18: License Management | 35 |
| Ilustración 19: SIEM índices | 36 |
| Ilustración 20: Elastic SIEM view | 36 |
| Ilustración 21: SIEM overview | 37 |
| Ilustración 22: SIEM Hosts | 37 |
| Ilustración 23: SIEM Autenticaciones | 38 |
| Ilustración 24: SIEM Uncommon Processes | 38 |
| Ilustración 25: SIEM Eventos | 39 |
| Ilustración 26: SIEM Network | 39 |
| Ilustración 27: Security Elastic | 40 |
| Ilustración 28: Definición de la regla | 42 |
| Ilustración 29: Descripción de la regla | 42 |
| Ilustración 30: Programación de la regla | 43 |
| Ilustración 31: Detección de ataque por Fuerza Bruta | 43 |
| Ilustración 32: Parada de Windows Defender | 43 |
| Ilustración 33: Detección de Malware en Windows Defender | 44 |
| Ilustración 34: Alertas detección de Malware | 44 |
| Ilustración 35: Regla de Cisco Umbrella para detectar Phising | 44 |
| Ilustración 36: Regla para detectar vulnerabilidades críticas en Fortiweb | 45 |
| Ilustración 37: Regla para detectar vulnerabilidades altas en Fortiweb | 45 |
| Ilustración 38: Señales detectadas en el SIEM | 45 |
| Ilustración 39: Timeline fuerza bruta | 46 |
| Ilustración 40: Flujo Machine Learning | 48 |
| Ilustración 41: Vista Machine Learning | 49 |
| Ilustración 42: Activación de trabajos desde el SIEM | 53 |
| Ilustración 43: Machine Learning Visualizador de Datos | 54 |
| Ilustración 44: Ejemplo de métricas Top y Distribution | 55 |
| Ilustración 45: Ejemplo de campos | 55 |
| Ilustración 46: Vista de Detección de Anomalías | 56 |

| | |
|---|----|
| Ilustración 47: Dataset Malware Cisco Umbrella | 57 |
| Ilustración 48: Job Management - Machine Learning | 58 |
| Ilustración 49: Tipos de trabajo - Machine Learning | 58 |
| Ilustración 50: Rango de tiempo del dataset - Machine Learning | 59 |
| Ilustración 51: Definición del trabajo - Machine Learning | 59 |
| Ilustración 52: Resumen del trabajo | 60 |
| Ilustración 53: Explorador de Anomalías - Machine Learning | 60 |
| Ilustración 54: Tabla Explorador de Anomalías | 61 |
| Ilustración 55: Single Metric Viewer - Machine Learning | 62 |
| Ilustración 56: Detección de anomalías - Cisco Umbrella | 63 |
| Ilustración 57: Single Metric Viewer - Cisco Umbrella | 64 |
| Ilustración 58: Detección de Anomalías – Fortiweb | 64 |
| Ilustración 59: Single Metric Viewer - Fortiweb | 64 |
| Ilustración 60: Diagrama de flujo Watcher | 66 |
| Ilustración 61: Crear Alerta Watcher | 67 |
| Ilustración 62: Definición de Watcher | 68 |
| Ilustración 63: Definición de la condición del Watcher | 68 |
| Ilustración 64: Posibilidades de notificación de una Alerta | 68 |
| Ilustración 65: Acción email Watcher | 69 |
| Ilustración 66: Notificación Watcher | 70 |
| Ilustración 67: Historial de ejecución Watcher CPU | 70 |
| Ilustración 68: Vista Watcher avanzado | 71 |
| Ilustración 69: Notificación fuerza bruta Watcher | 73 |
| Ilustración 70: Historial de ejecución Watcher Fuerza Bruta | 73 |
| Ilustración 71: Características de Seguridad por tipo de licencia | 75 |
| Ilustración 72: System Overview Dashboard | 77 |
| Ilustración 73: Host Overview Dashboard 1 | 77 |
| Ilustración 74: Host Overview Dashboard 2 | 77 |
| Ilustración 75: upTime Monitor | 78 |
| Ilustración 76: Syslog Events Dashboard Linux | 79 |
| Ilustración 77: WinlogBeat Dashboard Windows | 80 |
| Ilustración 78: Dashboard control de logins | 80 |
| Ilustración 79: Dashboard control de users | 81 |
| Ilustración 80: Dashboard control de integridad – 1 | 82 |
| Ilustración 81: Dashboard control de integridad - 2 | 82 |
| Ilustración 82: Dashboard control SSH - 1 | 83 |
| Ilustración 83: Dashboard control SSH - 2 | 83 |
| Ilustración 84: Dashboard control de procesos - 1 | 84 |
| Ilustración 85: Dashboard control de procesos – 2 | 84 |
| Ilustración 86: Flujo packetbeat – 1 | 85 |
| Ilustración 87: Flujo packetbeat – 2 | 85 |
| Ilustración 88: Flujo packetbeat – 3 | 85 |
| Ilustración 89: Instalación Elasticsearch | 94 |
| Ilustración 90: Verificación de instalación de Elasticsearch | 95 |
| Ilustración 91: Elasticsearch.yml – 1 | 95 |
| Ilustración 92: Elasticsearch.yml – 2 | 96 |
| Ilustración 93: Instalación de Kibana | 96 |
| Ilustración 94: Verificación de instalación de Kibana | 97 |
| Ilustración 95: Instalación de Logstash | 98 |
| Ilustración 96: Configuración metricbeat Ubuntu – 1 | 99 |

| | |
|--|-----|
| Ilustración 97: Configuración metricbeat Ubuntu - 2 | 99 |
| Ilustración 98: Metricbeat ubuntu verificación | 99 |
| Ilustración 99: Servicio metricbeat | 100 |
| Ilustración 100: metricbeat Windows verificación | 100 |
| Ilustración 101: Configuración filebeat – 1 | 101 |
| Ilustración 102: Configuración filebeat - 2 | 101 |
| Ilustración 103: Verificación de instalación de filebeat | 102 |
| Ilustración 104: Análisis de logs de Linux con filebeats | 102 |
| Ilustración 105: Configuración winlogbeat | 103 |
| Ilustración 106: Verificación winlogbeat | 103 |
| Ilustración 107: Configuración heartbeat – 1 | 104 |
| Ilustración 108: Configuración heartbeat – 2 | 104 |
| Ilustración 109: Configuración heartbeat - 3 | 105 |
| Ilustración 110: Configuración auditbeat - 1 | 106 |
| Ilustración 111: Configuración auditbeat - 2 | 106 |
| Ilustración 112: Certificados SSL/TLS | 107 |
| Ilustración 113: Gestión de usuarios | 110 |
| Ilustración 114: Gestión de roles | 111 |

1. Introducción

1.1 Contexto y justificación del Trabajo

En la actualidad, los ciberataques forman parte del día a día de las empresas. Los *ciberdelincuentes* se enfocan en las empresas porque saben que pueden obtener un gran beneficio. Entre los objetivos comunes de los *ciberdelincuentes*, se encuentran las filtraciones de información, el robo de datos de interés de la empresa, cuentas bancarias, la interrupción de servicios, extorsiones, etc. Esto a parte de generar consecuencias económicas para la empresa, afectaría a su reputación.

Es lógico pensar que dada la cantidad de ataques que reciben las empresas, y los riesgos que ello conlleva, cada vez se invierte más dinero en ciberseguridad para encontrar medidas que ayuden a evitar los riesgos adyacentes a un ciberataque. Por ello, es necesario que las empresas dispongan de herramientas que les permitan gestionar la información de seguridad y conocer en todo momento que está pasando en sus activos e infraestructuras.

El proyecto nace con la intención de proporcionar un sistema de monitorización en tiempo real, que proporcione medidas para detectar y prevenir cuándo una empresa está siendo atacada, y que permita a la empresa a tomar las medidas oportunas para responder y mitigar el ciberataque (Incident Response).

Esto se realizará mediante la implementación de un SIEM (Security Information and Event Management), que junto con la suite de herramientas de análisis de datos que proporciona Elastic (Elastic, Elastic, 2020) y el uso de Machine Learning, permiten formalizar una herramienta Open-Source completa para la detección y prevención de anomalías de seguridad.

1.2 Objetivos del Trabajo

Los objetivos generales del proyecto parten por estudiar las capacidades de la Elastic Suite en la detección de amenazas en entornos empresariales, en lo que se refiere a:

- Estudiar las capacidades de Elastic SIEM.
- Estudiar el funcionamiento del Machine Learning en la detección de amenazas de seguridad.
- Estudiar las capacidades de Elastic relacionadas con la analítica de datos y la utilidad que esto ofrece para la ciberseguridad.

Los objetivos específicos que se pretenden alcanzar gracias a la implementación y utilización del producto, resultante de la elaboración del proyecto, son los siguientes:

- Una herramienta que permita mantener el control de la seguridad de los activos en un entorno empresarial.
- Una herramienta que provee de un sistema de monitorización en tiempo real.
- Una herramienta que provee de un sistema de detección de amenazas en tiempo real.

- Una herramienta que provee de un sistema de prevención de amenazas de seguridad.
- Una herramienta que permita gestionar los principales riesgos de seguridad del entorno empresarial asociados a fallas de seguridad.
- Estudiar la viabilidad, en cuanto a ciberseguridad, de esta solución empresarial.

1.3 Enfoque y método seguido

El enfoque del proyecto consiste en desarrollar una herramienta integral de Seguridad informática a partir de los módulos integrantes en Elastic. El principal motivo es la centralización y unificación de los objetivos de seguridad empresariales en una única herramienta, evitando la necesidad de invertir tiempo y recursos en la integración común de múltiples herramientas.

De esta manera, se dispone de un servidor central con la herramienta Elastic que integra todos los módulos necesarios para satisfacer las necesidades de Seguridad enfocadas a la monitorización, gestión de logs y SIEM en entornos empresariales.

Para el desarrollo de este proyecto se utilizará la metodología Kanban (Kanban, 2019).

Kanban es un tipo de metodología utilizada en el desarrollo de proyectos, conocida por ser una metodología ágil. El origen de la palabra Kanban proviene de japonés, cuyo significado es el siguiente: **Kan**, significa visual, y **Ban**, hace referencia a las tarjetas. Esto es así puesto que la metodología utiliza *tarjetas* para gestionar, de manera *visual*, la realización de determinados procesos y tareas.

Las principales ventajas de esta metodología es que es muy sencilla de utilizar, actualizar y asumir. Destaca por ser una técnica de gestión de las tareas muy visual, que permite ver el estado del proyecto de manera muy clara, así como también establecer la pautar del desarrollo del trabajo de una manera efectiva y ágil.

La metodología Kanban se basa en una serie de principios:

- **Calidad garantizada.** Todo lo que se hace debe salir bien a la primera, no hay margen de error. De aquí a que en Kanban no se premie la rapidez, sino la calidad final de las tareas realizadas. Esto sucede ya que el coste de arreglar errores una vez realizada una tarea es muy elevado.
- **Reducción del desperdicio.** Kanban se basa en hacer solamente lo justo y necesario, pero el trabajo debe estar bien hecho. Esto supone la reducción de todo aquello que es superficial o secundario. A esto se le denomina *principio YAGNI*.
- **Mejora continua.** Kanban no es simplemente un método de gestión, sino también un sistema de mejora en el desarrollo de proyectos, según

los objetivos a alcanzar. Aprovechando la realización de tareas, se busca mejorar los procesos a través de un sistema de mejora continua.

- **Flexibilidad.** Lo próximo a realizar se decide del *backlog* (o tareas pendientes acumuladas), pudiendo priorizar aquellas tareas entrantes según las necesidades del momento y del proyecto (capacidad de dar respuesta a tareas imprevistas).

Para la implantación de esta metodología se hace uso de la herramienta Trello (Trello, 2020).

En Trello, se ha creado un tablero. El tablero está compuesto por las columnas “*Por hacer*”, “*En proceso*” y “*Hecho*”. Cada una de las columnas corresponde a un estado concreto del flujo de tareas, que nos servirá para saber en qué situación se encuentra cada actividad del proyecto.

Kanban se basa en el principio de desarrollo incremental, dividiendo el trabajo en distintas partes. Esto significa la tarea se divide en distintos pasos, para agilizar el proceso de producción.

Normalmente cada una de esas partes se escribe en una tarjeta y se pega en el tablero, en la fase que corresponda. Dichas tarjetas contienen la información básica para que el equipo sepa rápidamente la carga total de trabajo que supone, es decir, una descripción de la tarea con la estimación de horas.

Para la elaboración del proyecto, se crearán las tarjetas basándose en el listado de las tareas a realizar del siguiente punto, y de acuerdo a lo establecido en el plan de trabajo.

El resultado final es el siguiente:

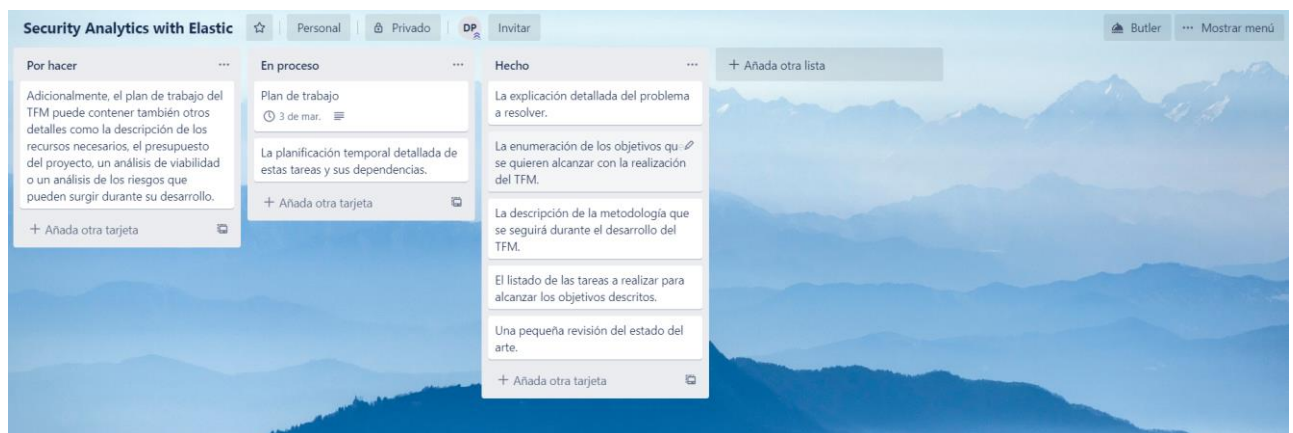


Ilustración 1: Tablero Trello

1.4 Planificación del Trabajo

A continuación, se muestra el desglose de tareas y el diagrama de Gantt del proyecto:

Tarea

| Nombre | Fecha de inicio | Fecha de fin | Duración |
|---|-----------------|--------------|----------|
| PEC-1 Plan de Trabajo | 19/02/20 | 3/03/20 | 14 |
| Enumeración objetivos | 20/02/20 | 21/02/20 | 2 |
| Explicación del problema a resolver | 19/02/20 | 21/02/20 | 3 |
| Descripción metodología | 24/02/20 | 25/02/20 | 2 |
| Listado de tareas | 24/02/20 | 25/02/20 | 2 |
| Planificación temporal | 25/02/20 | 26/02/20 | 2 |
| Estado del arte | 26/02/20 | 27/02/20 | 2 |
| Recursos y Presupuesto | 28/02/20 | 1/03/20 | 3 |
| Estudio de viabilidad y Riesgos | 28/02/20 | 1/03/20 | 3 |
| Instalación y configuración del laboratorio | 2/03/20 | 3/03/20 | 2 |
| Entrega PEC-1 | 3/03/20 | 3/03/20 | 0 |
| PEC-2 | 3/03/20 | 31/03/20 | 29 |
| Desarrollo de la memoria | 4/03/20 | 31/03/20 | 28 |
| Instalación y configuración Elastic | 3/03/20 | 5/03/20 | 3 |
| Investigar módulos de Elastic | 5/03/20 | 8/03/20 | 4 |
| Investigación out-of-the-box logs Elastic | 5/03/20 | 8/03/20 | 4 |
| Indexar logs DNS | 9/03/20 | 11/03/20 | 3 |
| Indexar logs Cisco Umbrella | 11/03/20 | 13/03/20 | 3 |
| Indexar logs PaloAlto | 13/03/20 | 15/03/20 | 3 |
| Indexar logs FortiWeb | 16/03/20 | 18/03/20 | 3 |
| Indexar logs CheckPoint | 19/03/20 | 21/03/20 | 3 |
| Indexar logs S.O | 21/03/20 | 23/03/20 | 3 |
| Parsear logs | 24/03/20 | 31/03/20 | 8 |
| Configurar Kibana | 22/03/20 | 24/03/20 | 3 |
| Creación Dashboard Kibana | 24/03/20 | 31/03/20 | 8 |
| Entrega PEC-2 | 31/03/20 | 31/03/20 | 0 |
| PEC-3 | 1/04/20 | 1/05/20 | 31 |

Ilustración 2: Tareas del proyecto – 1

Tarea

| Nombre | Fecha de inicio | Fecha de fin | Duración |
|------------------------------------|-----------------|--------------|----------|
| Desarrollo de la memoria | 1/04/20 | 28/04/20 | 28 |
| Análisis de log DNS | 1/04/20 | 3/04/20 | 3 |
| Estudio normalización eventos SIEM | 1/04/20 | 28/04/20 | 28 |
| Análisis de log S.O | 4/04/20 | 7/04/20 | 4 |
| Análisis de log Cisco Umbrella | 8/04/20 | 11/04/20 | 4 |
| Análisis de log FortiWeb | 12/04/20 | 15/04/20 | 4 |
| Análisis de log CheckPoint | 15/04/20 | 18/04/20 | 4 |
| Análisis de log Palo Alto | 19/04/20 | 22/04/20 | 4 |
| Configuración Watcher | 23/04/20 | 1/05/20 | 9 |
| Entrega PEC-3 | 28/04/20 | 28/04/20 | 0 |
| PEC-4 Memoria Final | 29/04/20 | 2/06/20 | 35 |
| Desarrollo de la memoria | 29/04/20 | 2/06/20 | 35 |
| Configurar SIEM | 29/04/20 | 14/05/20 | 16 |
| Investigación Machine Learning | 29/04/20 | 5/05/20 | 7 |
| Configuración Machine Learning | 5/05/20 | 14/05/20 | 10 |
| Investigación Malware | 29/04/20 | 5/05/20 | 7 |
| Instalar y detectar Malware | 5/05/20 | 15/05/20 | 11 |
| Entrega PEC-4 | 2/06/20 | 2/06/20 | 0 |
| PEC-5 Presentación en Vídeo | 3/06/20 | 9/06/20 | 7 |
| Elaboración video | 3/06/20 | 9/06/20 | 7 |
| Entrega PEC-5 | 9/06/20 | 9/06/20 | 0 |

Ilustración 3: Tareas del proyecto - 2

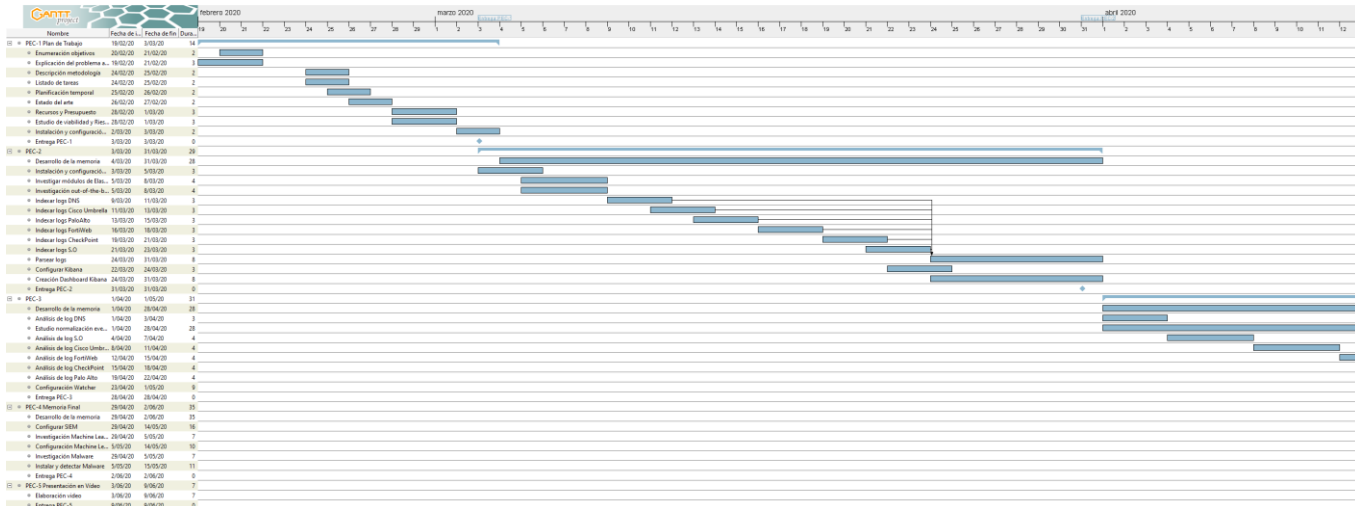


Ilustración 4: Diagrama de Gantt - 1

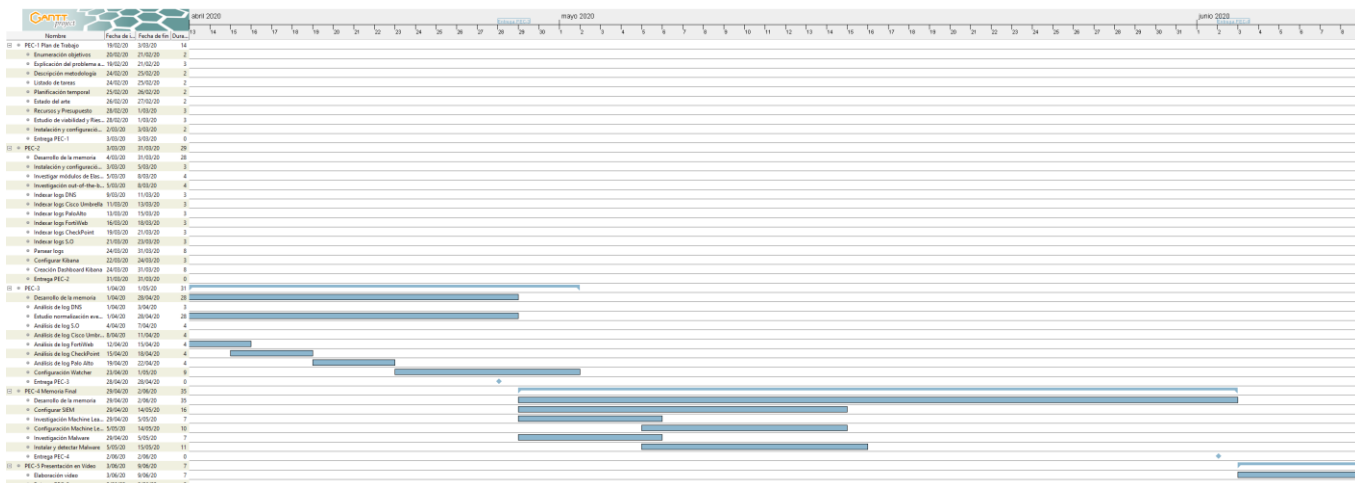


Ilustración 5: Diagrama de Gantt - 2

1.5 Estado del Arte

Elasticsearch (Elastic, Qué es Elasticsearch, 2020) es un motor de analítica y análisis distribuido y open source para todos los tipos de datos, incluidos textuales, numéricos, geospaciales, estructurados y desestructurados. Elasticsearch está desarrollado en Apache Lucene y fue presentado por primera vez en 2010 por Elasticsearch N.V. (ahora conocido como Elastic). Conocido por sus API REST simples, naturaleza distribuida, velocidad y escalabilidad, Elasticsearch es el componente principal del Elastic Stack (Elastic, Elastic Stack, 2020), un conjunto de herramientas open source para la ingesta, el enriquecimiento, el almacenamiento, el análisis y la visualización de datos. Comúnmente referido como el ELK Stack (referido a Elasticsearch, Logstash y Kibana), incluye una gran colección de agentes de envío conocidos como Beats para enviar los datos a Elasticsearch.

La velocidad y escalabilidad de Elasticsearch y su capacidad de indexar muchos tipos de contenido significan que puede usarse para una variedad de casos de uso:

- Búsqueda de aplicaciones.
- Búsqueda de sitio web.
- Búsqueda Empresarial.
- Logging y analíticas de log.
- Métricas de infraestructura y monitoreo de contenedores.
- Monitoreo de rendimiento de aplicaciones.
- Análisis y visualización de datos geoespaciales.
- Analítica de Seguridad.
- Analítica de Negocios.

Resulta interesante conocer la historia de este gran proyecto, cuyo creador es Shay Banon (Elastic, Historia de Elasticsearch, 2020).

Shay Banon empezó trabajando en su primera aproximación de lo que actualmente se conoce como Elastic en los años 2000. Ese proyecto se denominaba Compass. Mientras Shay Banon pensaba en la tercera versión de Compass, llegó a la conclusión de que habría que reescribir grandes partes de su código para crear “una solución de búsqueda escalable”. Así nace la segunda aproximación del proyecto, lo que propiamente se conoce como Elasticsearch. Este proyecto tuvo mucho éxito, y tras él comenzó a formarse una gran comunidad.

Cuando se fundó Elasticsearch Inc, había dos proyectos Open-Source que estaban comenzado su andadura.

Por un lado, Jordan Sissel estaba desarrollando Logstash, la herramienta de ingestión de datos y, por otro lado, Rashid Khan estaba trabajando en una interfaz de usuario Open-Source llamada Kibana.

Shay, Jordan y Rashid se conocían anteriormente, y juntos decidieron formar un equipo, el que resultó en la pila ELK: Elasticsearch, Logstash y Kibana Stack.

Poco después se lanzaron dos complementos comerciales: Marvel para la monitorización y Shield para la seguridad.

El equipo de Elastic continuó trabajando con el desarrollo de estos productos, trabajando en la mejora de la capacidad, la posibilidad de enviar más solicitudes de extracción, la creación de nuevos complementos y extensiones, etc. A su vez, se desarrolló Packetbeat con el objetivo de enviar datos de una red a Elastic de manera sencilla y Beats, con el objetivo de recoger métricas, registros, auditorías, etc de forma más liviana.

El crecimiento de Elastic fue drásticamente en aumento, llegando a incorporarse como un servicio AWS a través de Elastic cloud.

El lanzamiento de la versión 5.0 introdujo una experiencia de inicio más integrada, mejor probada y más fácil que nunca.

La versión 5.0 introduce todos los complementos comerciales, denominados Shield, Marvel y Watcher en ese momento, en una sola extensión llamada X-Pack. Este pack consiste en un grupo de características de seguridad, monitorización y alertas para los productos principales, y creció para incluir el Machine Learning.

Hoy en día Elastic está en la versión 7.6, y por supuesto, incluye una serie de mejoras respecto a versiones anteriores. Más concretamente, esta versión optimiza la detección automática de amenazas gracias al lanzamiento de un nuevo motor de detección de SIEM y un conjunto seleccionado de reglas de detección alineadas con la base de conocimientos MITRE ATT&CK™, aporta mejoras de rendimiento a Elasticsearch, hace que el Machine Learning supervisado esté mejor preparado para el uso con características de inferencia en la ingesta y profundiza la seguridad y observabilidad de la cloud con el lanzamiento de integraciones de datos nuevas.

Es importante e interesante profundizar en el enorme crecimiento actual en cuanto a soluciones comerciales enfocadas a la seguridad y monitorización. Además de Elastic, existen una gran diversidad de Empresas que ofrecen soluciones en forma de Software enfocado a la monitorización y gestión de logs, entre los que destacan Suricata, Grafana, Nagios, Zabbix, SolarWinds, etc, y SIEMs, entre que los que destacan QRadar, ArcSight, Splunk, OSSEC, Envolved...

Muchos de estos Software son productos de alto coste económico, por lo que el acceso a la ciberseguridad para las pequeñas empresas resulta dificultoso o prácticamente inaccesible. Esto se traduce en que las pequeñas empresas son más vulnerables, ya que una gran parte enfoca la seguridad de sus infraestructuras en sistemas de detección Endpoint y en el uso de Firewalls, no pudiendo permitirse invertir en soluciones de alto coste que permitan tener una visión en tiempo real del estado de seguridad de sus activos, que además implica un gasto en personal especializado adicional.

Además, aunque gran parte de las grandes marcas ofrecen soluciones gratuitas, estas carecen de la mayoría las características y funcionalidades Premium, siendo herramientas incompletas.

Por otro lado, las grandes empresas invierten grandes cantidades económicas en integrar distintas soluciones comerciales, con el objetivo de conseguir características propias de las soluciones del mercado más punteras. El coste de mantenimiento asociado, y el tiempo que se necesita para la integración, así como la resolución de las problemáticas comunes al proceso, caracterizan la falta de una herramienta completa que de soporte funcional a las necesidades propias de una empresa.

Elastic ofrece una solución integral Open-Source que permite centralizar la gestión de logs, la analítica de datos y el SIEM en una única herramienta. Cabe destacar que algunas de las funcionalidades que se utilizarán en este proyecto requieren de una suscripción Premium (licencia de pago), aunque es posible utilizar herramientas Open-Source desarrolladas por terceros que realizan funciones similares.

Así pues, Elastic nace como una solución atractiva e interesante por las funcionalidades implícitas que aporta, además de las facilidades y posibilidades que proporciona mediante la integración de sus distintos módulos. De esta manera, se evitan los problemas de integración que resultan de la utilización de diferentes tecnologías u herramientas, que implican un mayor impacto económico en la empresa.

Aunque es cierto que Splunk ofrece una solución similar, sus altos costes hacen que Elastic sea la solución más habitual para cubrir este tipo de servicios. Además, existe una gran comunidad por detrás de Elastic que da soporte continuo a los distintos módulos existentes, así como a la integración del producto con otras tecnologías del mercado.

1.6 Breve resumen de productos obtenidos

Los productos que se utilizarán para el desarrollo del proyecto son los siguientes:

- ElasticSearch 7.6.0
- Logstash 7.6.0
- Kibana 7.6.0
- Elastic SIEM 7.6.0
- Beats 7.6.0
- Winlogbeat 7.6.0
- MetricBeat 7.6.0
- AuditBeat 7.6.0
- FileBeat 7.6.0
- HeartBeat 7.6.0
- Packetbeat 7.6.0

Para la elaboración del laboratorio se utilizará el siguiente Software:

- Centos 7
- Windows Server 2016
- Ubuntu 16.04
- Oracle VM VirtualBox 5.2.26

1.7 Presupuesto

| Descripción | Precio | Unidades | TOTAL |
|-----------------------|--------|----------|----------------|
| Ordenador MSI GP62MVR | 1200€ | 1,00 | 1.200€ |
| Ingeniero Informático | 50€/h | 350,00 | 17.500€ |
| Windows Server 2016 | 500€ | 1,00 | 500€ |
| Licencias Software | 0€ | * | 0€ |
| | | | |
| TOTAL | | | 19.200€ |

Nota: En el desarrollo de este proyecto se utilizan funcionalidades de pago de Elastic. El costo es de 0€ porque se utiliza una versión Trial con una duración de 30 días.

1.8 Viabilidad y Análisis de riesgos

La viabilidad del proyecto es totalmente optimista dado el estado actual de la ciberseguridad en entornos empresariales. La razón se remite a la posibilidad de implementar este tipo de soluciones Open-Source con un coste de mantenimiento mínimo, y que además, analizando las características que ofrece sin necesidad de invertir capital, lo convierte en una solución totalmente viable, sobre todo en entornos empresariales donde las posibilidades de inversión en materia de ciberseguridad sean reducidas.

Por otro lado, en entornos empresariales donde se cuente con mayor capacidad económica de cara a realizar inversiones en ciberseguridad, el proyecto sigue siendo altamente viable, dado que las características que ofrece Elastic en su versión Premium se ven notablemente aumentadas, ofreciendo un servicio de mejor calidad. Cabe destacar que, pudiendo invertir más capital, también existe la posibilidad de emplear otras soluciones comerciales.

Existe algún riesgo que se debe tener en cuenta para el desarrollo del proyecto:

- **Retraso de proyecto:** Es posible que la duración del proyecto se vea afectada, dado a que es imposible determinar con certeza la estimación de la duración de las tareas, por lo que el resultado final puede verse ligeramente afectado.
- **Aumento de tareas:** Debido a que se desconoce el funcionamiento del Software utilizado, es posible que se necesite realizar implementaciones manuales, tales como scripts de parseo de logs, scripts de automatización, etc, que afecten considerablemente en la duración del proyecto y el número de tareas.
- **Cambios en la planificación de tareas:** Es posible que existan modificaciones en la planificación de tareas del proyecto. Dado que la repartición de tareas se ha hecho en base a una estimación de dificultad y tiempo, es posible que haya inconvenientes durante el desarrollo que impliquen modificaciones en el orden de realización de las tareas o en la imposibilidad de su realización, o bien en su adelanto.
- **Riesgos tecnológicos:** Debido a que se investigará la posibilidad de infectar un servidor con Malware para estudiar su detección, es posible que provoque problemas que afecten a la estabilidad y rendimiento del laboratorio.
- **Riesgo de alcance:** El alcance del proyecto puede verse afectado por la falta de funcionalidades de Elastic. Aunque gran parte de las funcionalidades son Open-Source, algunas de ellas requieren de una suscripción Premium, o incluso podrían estar en proceso de desarrollo y, por ende, incompletas o limitadas.

1.9 Breve descripción de los otros capítulos de la memoria

En los siguientes capítulos se tratarán los siguientes temas:

- **2.0 Elastic Stack.**
 - **2.1 Definición de Elastic Stack:** Se explicará qué es Elastic Stack y sus principales características y utilidades en el proyecto.
 - **2.2 Módulos de Elastic:** Se explicará detalladamente las funcionalidades de los módulos de Elastic que se utilizarán en el proyecto.
 - **2.3 Arquitectura del proyecto:** Explicación y descripción de la arquitectura utilizada en el proyecto.

- **3.0 Elastic SIEM.**
 - **3.1 Definición de SIEM:** Se explicará lo que es un SIEM y sus principales características y utilidades en el proyecto.
 - **3.2 Integración de eventos:** Se explicará el proceso de integración y análisis de eventos en el SIEM.
 - **3.3 Configuración y uso de Elastic SIEM:** Se explicará en qué consiste el proceso de configuración del SIEM y cómo se utiliza en el proyecto.

- **4.0 Machine Learning.**
 - **4.1 Definición de Machine Learning:** Se explicará en qué consiste el Machine Learning y para qué sirve su aplicación en el proyecto en la detección de amenazas.
 - **4.2 Configuración y uso de Machine Learning:** Se explicará el proceso de configuración del módulo de Machine Learning y su utilidad en el proyecto.

- **5.0 Elastic Watcher.**
 - **5.1 Definición del plugin Watcher:** Se explicará qué es Elastic Watcher y cómo funciona.
 - **5.2 Configuración y uso de Watcher:** Se explicará el proceso de configuración del plugin Watcher y su utilidad en el proyecto.

- **6.0 Hardening de Elastic:** El objetivo de este apartado es mostrar las características de Hardening disponibles en Elastic para proteger el sistema.

- **7.0 Casos de uso del proyecto:** Puesta en marcha real del proyecto, explicación de los posibles casos de uso y estudio de su funcionamiento.

2.0 Elastic Stack

Elastic Stack (Elastic, Características de Elastic Stack, 2020) es el corazón del proyecto, la pieza fundamental para que todo funcione. A continuación, se explicarán los conceptos básicos acerca de las tecnologías que se utilizarán para el desarrollo del proyecto.

2.1 Definición de Elastic Stack

Elastic Stack (ELK) es la sigla que enfoca los tres productos Open-Source de Elastic que se utilizarán para el desarrollo del actual proyecto, denominados Elasticsearch, Logstash y Kibana. Cada producto realiza una función dentro de esta pila, como se detalla a continuación:

- **Elasticsearch:** Elasticsearch (Elastic, Elasticsearch, 2020) es un motor de almacenamiento, búsqueda y análisis distribuido en tiempo real. Puede utilizarse con muchos fines, pero un contexto en el que destaca es el de indexar flujos de datos semiestructurados, como logs o paquetes de red decodificados. Se trata del corazón del Elastic Stack, su función principal es almacenar de forma centralizada los datos para que se pueda realizar y combinar muchos tipos de búsquedas: estructuradas, no estructuradas, geográficas, métricas, etc.

El funcionamiento es sencillo. Los datos sin procesar fluyen hacia Elasticsearch desde una variedad de fuentes, desde los que se incluyen logs, métricas de sistema o aplicaciones web. La ingesta de datos es el proceso mediante el cual estos datos son parseados, normalizados y enriquecidos antes de su indexación en Elasticsearch. Una vez indexados en Elasticsearch, los usuarios pueden ejecutar consultas complejas sobre sus datos y usar agregaciones para recuperar resúmenes complejos de sus datos.

Elasticsearch implementa índices invertidos con transductores de estado finito para búsquedas de texto completo, árboles de BKD para almacenar datos numéricos y geográficos, y un almacén de columnas para analíticas. Por ello, es una herramienta excelente para visualizar datos en tiempo real.

Un índice de Elasticsearch es una colección de documentos relacionados entre sí. Elasticsearch almacena datos como documentos *JSON*. Cada documento correlaciona un conjunto de claves (nombres de campos o propiedades) con sus valores correspondientes (textos, números, Booleanos, fechas, variedades de valores, geolocalizaciones u otros tipos de datos).

Durante el proceso de indexación, Elasticsearch almacena documentos y construye un índice invertido para poder buscar datos en el documento casi en tiempo real. La indexación comienza con la API de índice, a través de la cual puedes agregar o actualizar un documento JSON en un índice específico.

La escalabilidad es otro de los puntos fuertes de Elasticsearch, ya que se puede desplegar como un clúster. Gracias a esto, escala horizontalmente para manejar miles de millones de eventos por segundo, al tiempo que gestiona automáticamente la forma en la que se distribuyen los índices y las búsquedas en el clúster para realizar operaciones sin problemas de rendimiento.

- **Kibana:** Kibana (Elastic, Qué es Kibana, 2020) es una plataforma de análisis y visualización de código abierto diseñada para trabajar con Elasticsearch. Se utiliza Kibana para buscar, ver e interactuar con los datos almacenados en los índices de Elasticsearch. Puede realizar fácilmente análisis de datos avanzados y visualizar sus datos en una variedad de gráficos, tablas y mapas.

Kibana tiene dos funciones concretas, por un lado, es la interfaz gráfica que permite administrar todas las herramientas del Elastic Stack. Por otro lado, es la herramienta de visualización gráfica, se utiliza para la creación de tableros de control, mapas, informes, entre otros atractivos para visualizar información. Mantiene también la gestión de herramientas adicionales del Elastic Stack como los módulos de APM, Machine Learning, Canvas, Logging, Infraestructure y SIEM.

- **Logstash:** Logstash (Elastic, Logstash, 2020) es una útil herramienta que se integra con una amplia variedad de despliegues. Ofrece una gran selección de plugins para ayudar a analizar, enriquecer, transformar y almacenar datos de una gran variedad de fuentes. Si los datos requieren un procesamiento adicional que no está disponible en Beats, entonces se necesita utilizar Logstash en el despliegue.

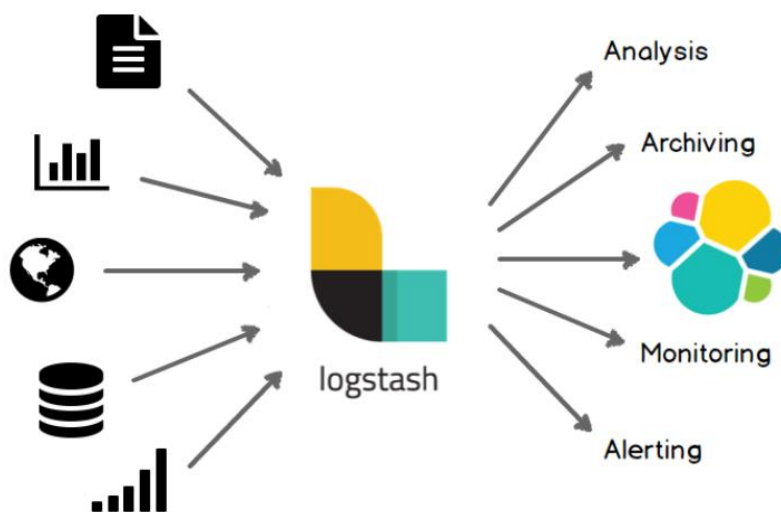


Ilustración 6: Funcionamiento de Logstash

La forma más simple de describir su funcionamiento es compararlo con un ETL (acrónimo del concepto Extract Transform, Load). Logstash ingesta, transforma y envía de forma dinámica los datos independientemente de su formato o complejidad. Para realizar su función, se utilizan tres plugins:

- **Input:** Generalmente, los datos se encuentran repartidos o en distribuidos en muchos sistemas en diversos formatos. Logstash admite una variedad de entradas que extraen eventos de una multitud de fuentes comunes, todo al mismo tiempo.
- **Filter:** A medida que los datos viajan de la fuente al almacén, los filtros de Logstash parsean cada evento, identifican los campos con nombre para crear la estructura y los transforman para que terminen en un formato común para un análisis y un valor comercial más potente.
Logstash transforma y prepara de forma dinámica tus datos independientemente de su formato o complejidad:
 - Deriva estructura a partir de datos no estructurados con grok
 - Descifra las coordenadas geográficas a partir de las direcciones IP
 - Anonimiza datos PII y excluye campos sensibles por completo
 - Facilita el procesamiento general, independientemente de la fuente de datos, el formato o el esquema.
- **Output:** Indica a donde se enviarán los datos una vez procesados. Aunque en este proyecto se utilizará Elasticsearch, no es el único punto de salida disponible.
Logstash tiene una variedad de salidas que permiten enviar los datos a donde se desee, lo que brinda la flexibilidad de desbloquear una gran cantidad de casos de uso posteriores.

Gracias a este conjunto de herramientas, se conforma la Elastic Stack. El eje fundamental del proyecto pasa por la utilización de estas herramientas para la ingesta y tratamiento de fuentes de información, y su posterior análisis para obtener información relativa a la seguridad de la infraestructura. Las razones de su utilización, entre otras, son las siguientes:

- Elasticsearch es rápido. Como Elasticsearch está desarrollado sobre Lucene, es una excelente herramienta en la búsqueda de texto completo. Conformar una plataforma de búsqueda en casi tiempo real, lo que implica que la latencia entre el momento en que se indexa un documento hasta el momento en que se puede buscar en él es muy breve. Por ello, Elasticsearch está bien preparado para casos de uso con restricciones de tiempo como analítica de seguridad y monitorización de infraestructuras.

- Elasticsearch es una plataforma distribuida. Los documentos almacenados en Elasticsearch se distribuyen en distintos contenedores conocidos como *shards*, que están duplicados para brindar copias redundantes de los datos en caso de que falle el hardware. Su naturaleza distribuida le permite escalar horizontalmente en cuanto a grandes cantidades de servidores y gestionar volúmenes de datos de tamaño inmenso.
- Elasticsearch viene con un amplio conjunto de características (Elastic, Características de Elastic Stack, 2020). Además de su velocidad, la escalabilidad y la resistencia, Elasticsearch tiene una cantidad de características integradas que contribuyen a que el almacenamiento y la búsqueda de datos sean incluso más eficientes, como data rollup y gestión de ciclo de vida del índice. Además, incluye características específicas de seguridad, que se adaptan muy bien a los objetivos del proyecto.
- El Elastic Stack simplifica la ingesta de datos, la visualización y el reporte. La integración con Beats y Logstash facilita el proceso de datos antes de indexarlos en Elasticsearch. Y Kibana provee visualización en tiempo real de los datos de Elasticsearch así como interfaz gráfica para acceder rápidamente al monitoreo de rendimiento de aplicaciones (APM), los logs y los datos de métricas de infraestructura.

2.2 Módulos de Elastic

Elastic ofrece una serie de módulos que facilitan la ingesta de datos en Elasticsearch desde diferentes fuentes de información. El módulo referencia para la ingesta de datos es Beats.

- **Beats:** Beats (Elastic, Beats, 2020) es un conjunto de agentes que se encargan de recoger datos de hosts o servidores, con el propósito de enviarlos a ElasticSearch o a Logstash. Los tipos de agentes existentes son:
 - **Metricbeat:** Metricbeat es un agente que se utiliza en hosts de Linux, Windows y Mac para recogida de estadísticas de uso de CPU, memoria, sistema de archivos, E/S de disco y E/S de red a nivel de sistema, así como estadísticas de alto nivel para cada proceso que se ejecute en el sistema. Estos datos se envían a Elasticsearch donde permanecen indexados.
 - **Filebeat:** Filebeat es un agente que se utiliza para el envío y la centralización de datos de registro. Filebeat se encarga de monitorizar los archivos de registro o las ubicaciones que se especifiquen, recoge los eventos de registro y los envía a Elasticsearch o Logstash para su indexación. En la siguiente captura se ejemplifica su funcionamiento:

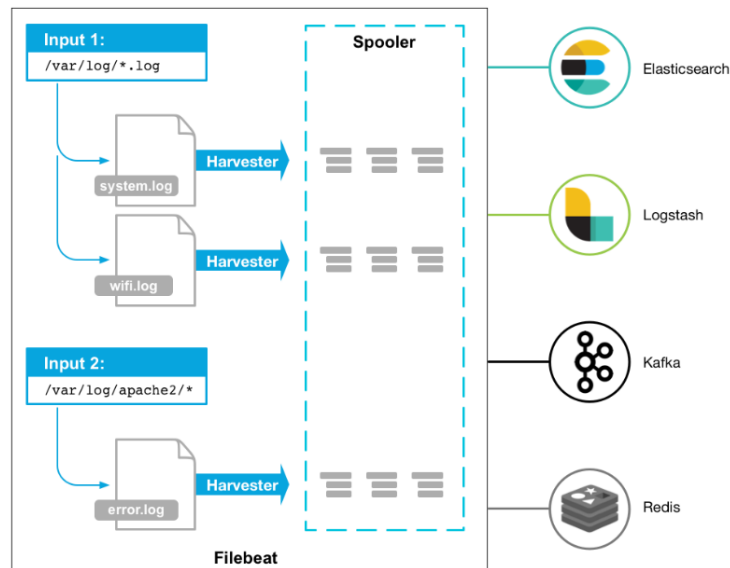


Ilustración 7: Funcionamiento de Filebeat

- **Winlogbeat:** Winlogbeat envía los registros de eventos de Windows a Elasticsearch o Logstash. Puedes instalarlo como un servicio de Windows.

Winlogbeat lee de uno o más registros de eventos usando Windows APIs, luego filtra los eventos siguiendo los criterios establecidos por el usuario, y finalmente envía los datos del evento a las salidas configuradas (Elasticsearch o Logstash). Winlogbeat vigila los registros de eventos para que los nuevos datos de eventos sean enviados a tiempo. La posición de lectura de cada registro de eventos se persigue en el disco para permitir que Winlogbeat se reanude después de los reinicios.

Winlogbeat puede capturar los datos de los eventos de cualquier registro de eventos que se ejecute en su sistema. Por ejemplo, puede capturar eventos como:

- Eventos de la aplicación.
- Eventos de hardware.
- Eventos de seguridad.
- Eventos del sistema.

- **Auditbeat:** Auditbeat es un agente que puede instalarse en servidores para auditar las actividades de los usuarios y los procesos de sus sistemas. Por ejemplo, se puede utilizar Auditbeat para recopilar y centralizar eventos de auditoría del Linux Audit Framework. También puede utilizar Auditbeat para detectar cambios en archivos críticos, como binarios y archivos de configuración, e identificar posibles violaciones de la política de seguridad.

- **Packetbeat:** Packetbeat es un analizador de paquetes de red en tiempo real que se utiliza para proporcionar un sistema de monitorización de aplicaciones y análisis de rendimiento.

Packetbeat funciona capturando el tráfico de red entre los servidores de aplicación, decodificando los protocolos de la capa de aplicación (HTTP, DNS, MySQL, etc.), correlacionando las solicitudes con las respuestas, y registrando los campos de interés para cada transacción.

Packetbeat esnifa el tráfico entre los servidores, analiza los protocolos de nivel de aplicación sobre la marcha, y correlaciona los mensajes en las transacciones.



Ilustración 8: Packetbeat flow

- **Heartbeat:** Heartbeat es un agente que se instala en un servidor remoto para comprobar periódicamente el estado de sus servicios y determinar si están disponibles. A diferencia de Metricbeat, que sólo dice si los servidores están encendidos o apagados, Heartbeat comprueba si los servicios son accesibles.

Heartbeat es útil para la seguridad del entorno, utilizándose cuando se necesita verificar que nadie desde el exterior puede acceder a los servicios de un servidor de la empresa.

Además de los módulos de ingesta de datos de Elastic, dentro de la interfaz de Kibana se proporcionan una serie de herramientas para el tratamiento y análisis de datos, a saber:

- **APM:** Elastic APM es un sistema de monitorización del rendimiento de las aplicaciones construido sobre el Elastic Stack. Permite monitorear los servicios y aplicaciones de software en tiempo real, recolectar información detallada de desempeño sobre el tiempo de respuesta de las solicitudes entrantes, consultas a la base de datos, llamadas a cachés, solicitudes HTTP externas y más. Esto hace que sea fácil localizar y solucionar rápidamente los problemas de rendimiento.

Elastic APM también recopila automáticamente los errores no manejados y las excepciones. Los errores se agrupan basándose principalmente en el stacktrace, para que pueda identificar nuevos errores a medida que aparecen y vigilar cuántas veces ocurren errores específicos.

Las métricas son otra importante fuente de información cuando se depuran los sistemas de producción. Los agentes APM recogen automáticamente métricas básicas de nivel de host y métricas específicas de los agentes.

- **Canvas:** Es una herramienta de visualización y presentación de datos que se encuentra dentro de Kibana. Con Canvas, se pueden obtener datos en vivo directamente de Elasticsearch, y combinar los datos con colores, imágenes, texto para crear presentaciones dinámicas.

Sus características principales son:

- Crear y personalizar tu espacio de trabajo con fondos, bordes, colores, fuentes y más.
 - Personalizar tu mesa de trabajo con tus propias visualizaciones, como imágenes y texto.
 - Personaliza tus datos sacándolos directamente de Elasticsearch.
 - Muestra tus datos con tablas, gráficos, monitores de progreso y más.
 - Enfoca los datos que quieres mostrar con filtros.
- **Dashboard:** Un tablero es una colección de visualizaciones, búsquedas y mapas, típicamente en tiempo real. Los cuadros de mando proporcionan una visión general de los datos y permiten profundizar en los detalles.

Los Dashboard tienen las siguientes funcionalidades:

- Agregar visualizaciones, búsquedas guardadas y mapas para el análisis de lado a lado.
 - Permite disponer de los elementos del cuadro de mando para que se muestren de manera personalizada.
 - Permite también personalizar los rangos de tiempo para mostrar sólo los datos que se deseen.
 - Inspeccionar y editar los elementos del tablero para averiguar exactamente qué tipo de datos se muestran.
- **Logs:** La monitorización de logs permite analizar los logs de la infraestructura e identificar los problemas prácticamente en tiempo real. Con esto se pueden visualizar logs de los servidores, contenedores, servicios, etc.
Permite desglosar para ver información más detallada sobre una entrada de registro individual, o puede cambiar sin problemas para ver las métricas correspondientes, la información de tiempo de actividad o los rastros de APM cuando estén disponibles. También se puede utilizar el aprendizaje automático para detectar automáticamente algunos tipos de anomalías en los registros.
 - **Uptime:** Elastic Uptime permite vigilar la disponibilidad y los tiempos de respuesta de las aplicaciones y servicios en tiempo real y detectar los problemas antes de que afecten a los usuarios.

Elastic Uptime ayuda a comprender las características del tiempo de disponibilidad y de respuesta de sus servicios y aplicaciones. Puede desplegarse tanto dentro como fuera de la red de la organización, para que se pueda analizar los problemas desde múltiples puntos de vista.

Elastic Uptime utiliza los siguientes componentes componentes: Heartbeat, Elasticsearch y Kibana.

- **Maps:** Maps permite analizar los datos geográficos a escala, con velocidad y en tiempo real. Con características como múltiples capas e índices en un mapa, trazado de documentos sin procesar, estilo dinámico del lado del cliente y búsqueda global a través de múltiples capas, es útil para entender y monitorear los datos con facilidad.

Con Elastic Maps se puede hacer lo siguiente:

- Crear mapas con múltiples capas e índices.
 - Cargar archivos GeoJSON en Elasticsearch.
 - Incrustar su mapa en Dashboards.
 - Trazar documentos individuales o usar agregados para trazar cualquier conjunto de datos, sin importar su tamaño.
 - Crear mapas coropletos.
 - Utilice un estilo basado en datos para simbolizar características de los valores de las propiedades.
 - Enfoca los datos que quieres mostrar con las búsquedas.
- **Metrics:** La aplicación Metrics permite monitorear la métrica de la infraestructura e identificar problemas en tiempo real. Se puede ver las métricas básicas de los servidores, contenedores, servicios, etc. Es posible hacer un desglose para ver métricas más detalladas, o ver los registros correspondientes, la información de tiempo de actividad o las trazas de APM cuando estén disponibles.
 - **Visualize:** Visualize (Visualizar) permite crear visualizaciones de los datos de los índices de Elasticsearch, que posteriormente se pueden añadir a los cuadros de mando para su análisis.

Las visualizaciones de Kibana se basan en las consultas de Elasticsearch. Utilizando una serie de agregaciones de Elasticsearch para extraer y procesar sus datos, permite crear gráficos que muestren las tendencias, los picos y los descensos.

- **Machine Learning:** Las propiedades del Machine Learning se utilizan para automatizar el análisis las series temporales de datos, creando líneas base del comportamiento normal de los datos e identificando patrones anómalos en dichos datos.

Elastic Machine Learning utiliza algoritmos patentados de aprendizaje automático, los cuales detectan, califican y vinculan las siguientes

circunstancias con factores de influencia estadísticamente significativos en los datos:

- Anomalías relacionadas con desviaciones temporales en los valores, recuentos o frecuencias
- Rareza estadística
- Comportamientos inusuales para un miembro de la población

Además de esto, el Machine Learning tiene un enfoque clave en la detección de anomalías. Las características de Machine Learning en la detección de anomalías utilizan una combinación a medida de diferentes técnicas como la agrupación, varios tipos de descomposición de series temporales, el modelado de distribución bayesiana y el análisis de correlación. Estos análisis proporcionan una sofisticada detección automatizada de anomalías en tiempo real para los datos de las series temporales.

Para ello, modela estadísticamente las características temporales de los datos observando el comportamiento histórico y adaptándose a los nuevos datos. El modelo representa una línea base de comportamiento normal y por lo tanto puede ser usado para determinar cuán anómalos son los nuevos eventos.

- **SIEM:** El SIEM permite el análisis de eventos de seguridad relacionados con los servidores y la red, como parte de las investigaciones de alerta o la caza interactiva de amenazas.

La aplicación SIEM se encuentra en Kibana, y proporciona un espacio de trabajo interactivo para que los equipos de seguridad seleccionen los eventos y realicen las investigaciones iniciales. Además, los trabajos de detección de anomalías de Machine Learning y el uso de reglas de detección proporcionan formas de detectar automáticamente actividades sospechosas en toda la infraestructura de servidores y estaciones de trabajo.

Además de las características comentadas anteriormente, también es importante realizar un breve inciso acerca de las utilidades de pago de Elastic Stack, las cuales se utilizarán en el este proyecto. Estas utilidades proporcionan una serie de módulos para la administración y gestión de Elastic Stack, las cuales a continuación se muestra una breve descripción:

- **Console:** Proporciona una interfaz JSON para trabajar directamente con los datos.
- **Index Patterns:** Gestiona los patrones de índice que se utilizan para recuperar los datos de Elasticsearch.
- **Monitoring:** Rastrea la salud y el rendimiento en tiempo real de tu pila elástica.
- **Rollups:** Resume y almacena los datos históricos en un índice más pequeño para su análisis futuro.

- **Saved Objects:** Importa, exporta y administra las búsquedas, visualizaciones y tableros guardados.
- **Security Settings:** Proporciona herramientas para proteger y administrar los datos mediante control de acceso basado en roles.
- **Spaces:** Organizador de tableros y otros objetos guardados en categorías significativas.
- **Watcher:** Se utiliza para detectar cambios en los datos creando, administrando y monitoreando alertas.

2.3 Arquitectura del proyecto

En este apartado se explicará la arquitectura de Elastic Stack y la arquitectura del laboratorio que se ha montado para la llevar a cabo la ejecución del proyecto.

2.3.1 Arquitectura de Elastic

La arquitectura de Elastic tiene la siguiente forma, tal y como se describió en los apartados anteriores:

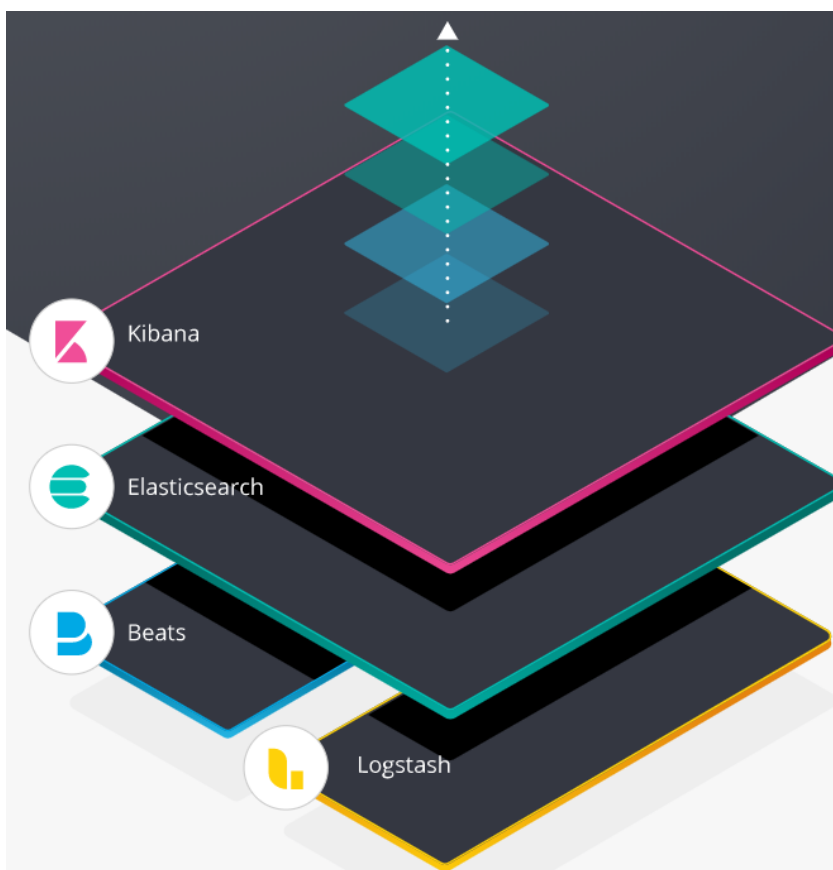


Ilustración 9: Arquitectura Elastic Stack

El flujo de los datos va de abajo para arriba. En la capa más baja están Beats y Logstash, los cuales actúan como agentes de recogida de datos. En la capa intermedia está Elasticsearch, que es la base de datos donde se almacenarán los datos recogidos por Logstash o Beats. En el nivel más alto está Kibana, que es la encargada de mostrar los datos en cuadros y gráficos para su posterior análisis.

Las comunicaciones se realizan mediante protocolo HTTP a través de los siguientes puertos:

- Elasticsearch: 9200
- Kibana: 5601
- Beats: 5044

Nota: Las comunicaciones también se pueden realizar a través del API REST.

2.3.2 Arquitectura del laboratorio

El laboratorio está compuesto por tres servidores virtuales:

- **Servidor Centos 7:** Se trata del nodo central del laboratorio. Está compuesto por los siguientes servicios:
 - Elasticsearch 7.6.0.
 - Kibana 7.6.0.
 - Logstash 7.6.0.
 - Filebeat 7.6.0.
- **Servidor Windows Server 2016:** Se trata de un host Windows Server que dispone de los siguientes servicios:
 - Servidor web en el puerto 8080.
 - Auditbeat 7.6.0.
 - Winlogbeat 7.6.0.
 - Heartbeat 7.6.0.
 - Metricbeat 7.6.0.
- **Servidor Ubuntu 16.01:** Se trata de un host Linux que dispone de los siguientes servicios:
 - Auditbeat 7.6.0.
 - Filebeat 7.6.0.
 - Heartbeat 7.6.0.
 - Metricbeat 7.6.0.

La arquitectura y topología de red del laboratorio se ve en las siguientes imágenes:

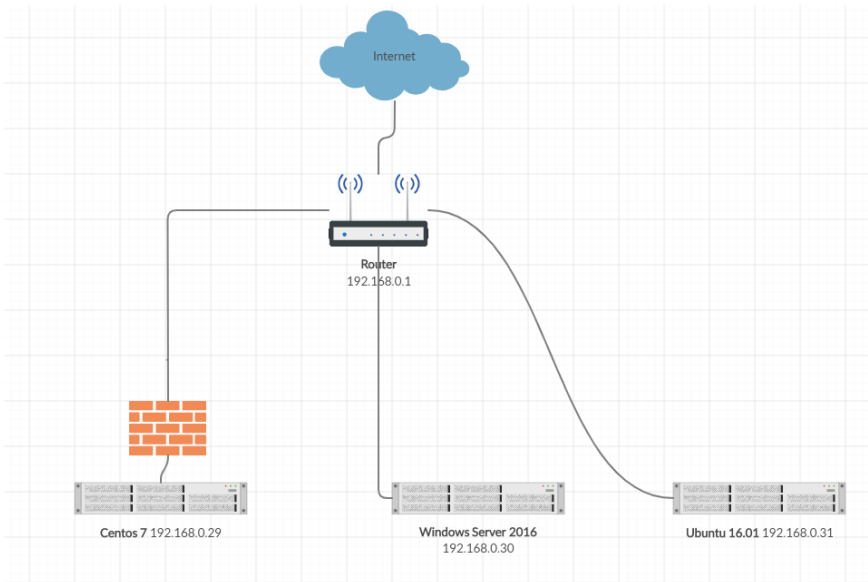


Ilustración 10: Topología de red del laboratorio

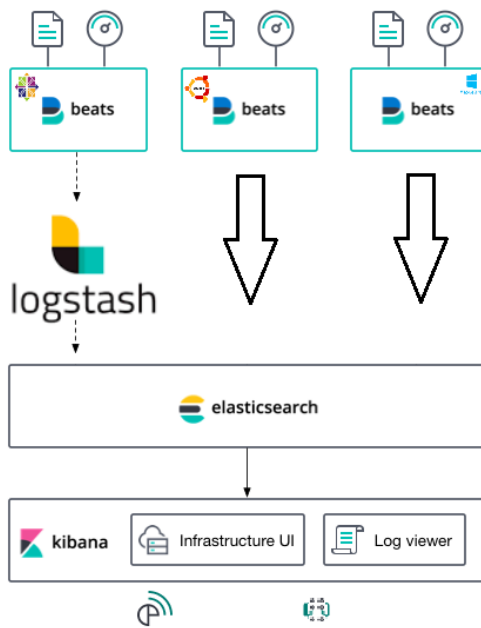


Ilustración 11: Arquitectura del laboratorio

3.0 Elastic SIEM

En esta parte de la memoria se pondrá el foco en la herramienta más destacada en cuanto a la seguridad informática de Elastic, el Elastic SIEM.

El objetivo es mostrar las características principales de un SIEM, el proceso de integración de eventos y su configuración para el posterior análisis de datos y detección de anomalías de seguridad.

3.1 Definición de SIEM

La tecnología de Gestión de Información y Eventos de Seguridad, en inglés SIEM, es una herramienta que apoya la detección de amenazas, el cumplimiento y la gestión de incidentes de seguridad mediante la recopilación y el análisis (tanto en tiempo casi real como histórico) de eventos de seguridad, así como una amplia variedad de otras fuentes de datos de eventos y contextuales. El Core de las capacidades de un SIEM provienen del amplio alcance de la recopilación y gestión de eventos de logs, la capacidad de analizar los eventos y otros datos a través de fuentes dispares, y las capacidades operacionales, tales como la gestión de incidentes, los cuadros de mando y la presentación de informes.

La tecnología SIEM nace de la combinación de las funciones de dos tecnologías: SEM (Gestión de Eventos de Seguridad) y SIM (Gestión de Información de Seguridad).

Un sistema SEM centraliza el almacenamiento y permite un análisis casi en tiempo real de lo que está sucediendo en la gestión de la seguridad, detectando patrones anormales de accesibilidad y dando mayor visibilidad a los sistemas de seguridad.

Por otro lado, un sistema SIM recopila los datos a largo plazo en una base de datos central para posteriormente analizarlo mediante tendencias y patrones de conducta. Este sistema proporciona informes automatizados para el cumplimiento y la generación de informes centralizados.

Con la unión de SIM y SEM surgen los sistemas SIEM, los cuales proporcionan una identificación, análisis y recuperación más rápida de los eventos de seguridad.

El objetivo de un sistema SIEM es dotar a las empresas con la capacidad de:

- Centralizar y controlar las amenazas potenciales.
- Determinar qué amenazas son falsos positivos.
- Recopilar información acerca de las amenazas para colaborar en la respuesta a incidentes.
- Documentar, en un registro de auditoría, los eventos detectados y cómo fueron resueltos.
- Cumplir con las regulaciones de la industria en un formato de reporte sencillo.

Por esto, los SIEM se han convertido en sistemas indispensables para la seguridad informática de la empresa. Sin duda, lo que se consigue utilizando

un SIEM no es solo una mejor gestión del tiempo de trabajo y una mayor facilidad para desarrollar las labores del equipo de seguridad, sino que acorta los tiempos de actuación, aspecto fundamental para una empresa en caso de una amenaza urgente. Esto es posible mediante un análisis centralizado de datos de seguridad, obtenidos desde múltiples sistemas, que incluyen aplicaciones antivirus, firewalls y soluciones de prevención de intrusiones.

La posibilidad de contar con datos de seguridad que fluyen en una vista centralizada de la infraestructura es efectiva solo si esos datos pueden ser estandarizados. Esto implica que todos los datos deben tener un formato común, siendo necesario para que el sistema SIEM pueda ejecutar su análisis y correlación de eventos.

En el nivel más básico, el funcionamiento de un sistema SIEM puede estar basado en reglas estadísticas o en el empleo de un motor de correlación para establecer relaciones entre entradas de registro de eventos.

Una vez centralizados los datos, el SIEM analiza los datos y busca los comportamientos anómalos.

En un nivel más avanzado, se emplean técnicas de detección de anomalías con base en Machine Learning. Elastic utiliza el aprendizaje automático no supervisado, el cual es útil para encontrar patrones en los datos.

La arquitectura de Elastic SIEM se resume en la siguiente imagen:

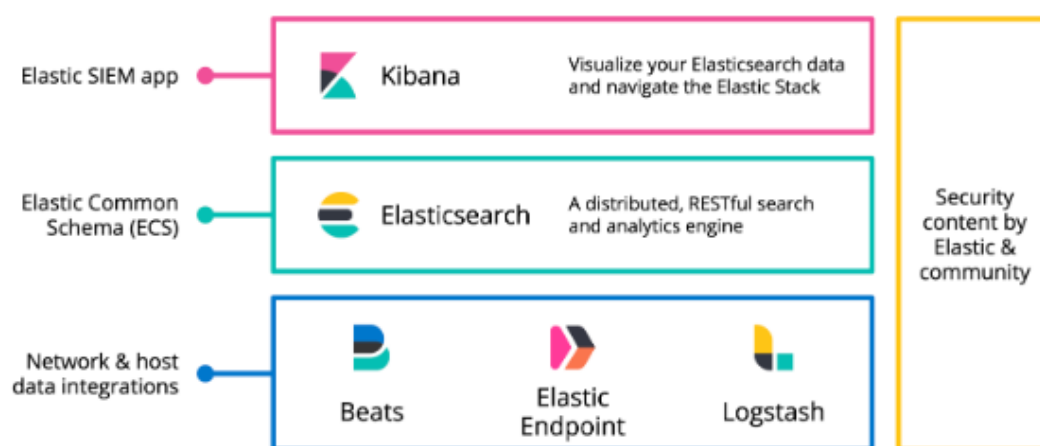


Ilustración 12: Arquitectura del SIEM

El SIEM recibe eventos y datos desde los agentes de entrada. Los registros son recogidos de múltiples fuentes de información, tales como la red, aplicaciones del sistema, sistemas de seguridad, logs del sistema operativo, etc.

Los logs recogidos son normalizados y estandarizados para poder ser almacenados e interpretados por el SIEM.

Con esto, el SIEM proporciona monitorización en tiempo real de la infraestructura, la detección de amenazas y una rápida respuesta a las brechas potenciales.

3.2 Integración de eventos

En esta parte de la memoria se procederá a explicar el proceso de integración de eventos en el SIEM, el cual parte por indexar eventos en Elasticsearch y normalizarlos mediante ECS (Elastic, Elastic Common Schema (ECS) | ECS Guide 1.5, 2020). El Elastic Common Schema (ECS) es una especificación de código abierto, desarrollada con el apoyo de la comunidad de usuarios de Elastic. El ECS define un conjunto común de campos que se utilizarán cuando se almacenen datos de eventos en Elasticsearch, como registros y métricas.

ECS especifica los nombres de los campos y los tipos de datos de Elasticsearch para cada campo, y proporciona descripciones y ejemplos de uso. El ECS también agrupa los campos en niveles de ECS, que se utilizan para señalar cuánto se espera que esté presente un campo.

El objetivo de ECS es permitir a los usuarios de Elasticsearch la normalización sus datos de eventos, para que puedan analizar, visualizar y correlacionar mejor los datos representados en sus eventos. ECS permite normalizar una amplia variedad de eventos, abarcando:

- **Fuentes de eventos:** Si la fuente de su evento es un producto de Elastic, un producto de terceros, o una aplicación personalizada construida por su organización.
- **Arquitecturas de ingesta:** Si la ruta de ingesta de sus eventos incluye procesadores Beats, Logstash, nodo de ingesta Elasticsearch, o ninguno de estos.
- **Consumidores:** Ya sea consumido por API, consultas de Kibana, tableros, aplicaciones u otros medios.

A continuación, se citan los eventos (logs) que se utilizarán en el proyecto:

- **Logs de Fortinet:** FortiWeb es lo que se conoce como un WAF (Web Application Firewall). Adopta un enfoque integral para proteger las aplicaciones web, incluyendo la reputación de IP, la protección DDoS, la validación de protocolos, las firmas de ataque de aplicaciones, la mitigación de bots y más para defender su aplicación contra una amplia gama de amenazas, incluidas las del OWASP Top 10.
- **Logs del Sistema Operativo:** Logs provenientes de eventos de Windows y Linux relacionados con el Sistema Operativo en cuestión.
- **Logs de Cisco Umbrella:** Servicio que recoge el tráfico DNS y limpia aquellas consultas dirigidas a dominios peligrosos o aquellas que formen parte de listas negras. Además, categoriza todas las resoluciones.
- **Logs de PaloAlto:** Firewall NextGen (Next Generation) que está provisto de IPS y AntiMalware.

Nota: Como en el momento de la realización del presente proyecto Elastic SIEM se encontraba en fase BETA, y el objetivo principal de esta parte es el de mostrar el proceso de estandarización de un log y el trabajo que ello conlleva,

tan sólo se realizará el proceso de estandarización a ECS con los logs de Cisco Umbrella y de Fortiweb. Los demás logs se integrarán en Elasticsearch sin estandarizar para ver el proceso de ingesta de logs desde distintas fuentes de información.

3.2.1 Integración y normalización de eventos de FortiWeb

Para la integración de los eventos de FortiWeb se utilizará Logstash. Para ello, se ha definido un fichero de configuración con el siguiente código:

```
input {
  file {
    path => "/home/user/Descargas/attack_download.log"
    start_position => "beginning"
  }
}

filter {
  kv {
    source => "message"
  }
  mutate {
    add_field => { "timestamp" => "%{date} %{time}" }
  }
  date {
    match => ["timestamp", "yyyy-MM-dd HH:mm:ss"] }
  }
}

output {
  elasticsearch {
    index => "fortiweb"
    hosts => ["localhost:9200"]}
}
```

Lo que hace este código es filtrar el fichero de log y extraer los campos y sus valores, para poder trabajar con ellos. Una vez extraídos, se crea un índice denominado *fortinet* y se almacenan los datos.

Para indexar manualmente los datos, se utiliza el siguiente comando:

```
/usr/share/logstash/bin/logstash --path.settings
/etc/logstash -f /etc/logstash/conf.d/fortiweb.conf
```

Una vez ejecutado el comando, se comprueba si se ha creado el índice en Elasticsearch desde el Index Management de Kibana.

Para poder visualizar los datos, hay que crear un patrón de índice en Kibana. Para ello, desde el Index Management de Kibana se accede a “*Create Index pattern*”. A continuación:

- Se introduce el nombre deseado.
- Se introduce el valor *timestamp* como Time Filter.

Para verificar que se están recibiendo datos, desde el Discover de Kibana se pueden visualizar los datos entrantes:

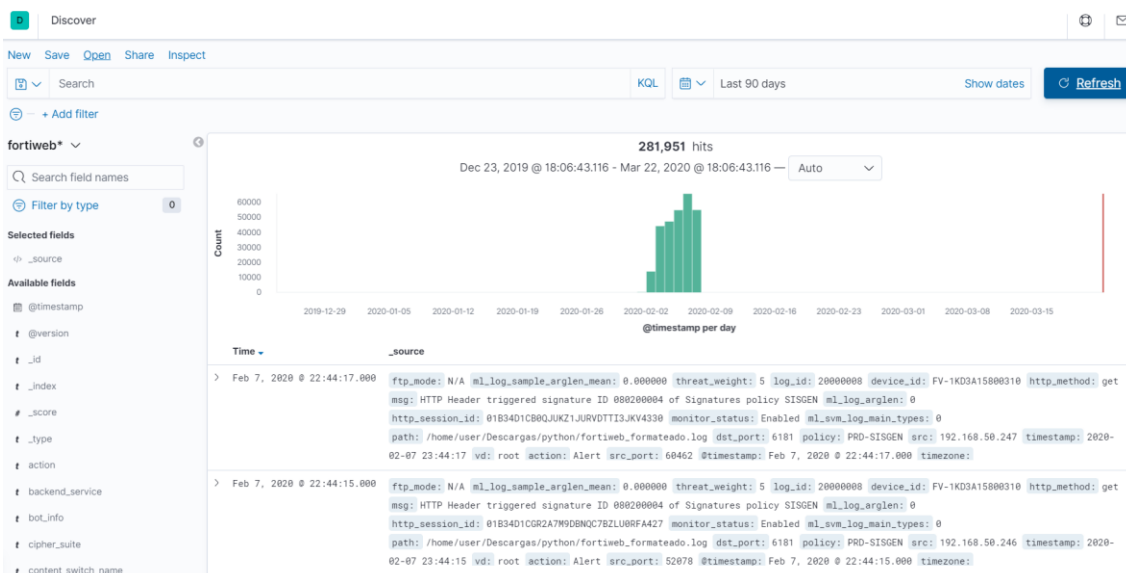


Ilustración 13: Eventos de Fortiweb

El proceso de normalización de eventos es más complejo. Este proceso parte por renombrar los campos originales que utiliza Fortiweb y emplear los nombres de campo que corresponden con el ECS para que puedan ser detectados. Este proceso requiere de un paso previo de investigación para determinar la relación entre los campos originales y los que cumplen el estándar ECS.

Para realizar este proceso, hay que modificar el archivo de configuración de Logstash anterior y modificar los campos originales. El archivo de configuración debe quedar de la siguiente manera:

```

input {
  file {
    path => "/home/user/Descargas/attack_download.log"
    start_position => "beginning"
  }
}
filter {
  kv {
    source => "message"
  }
  mutate {
    rename => { "[service]" => "[network][protocol]" }
    rename => { "[action]" => "[event][action]" }
    rename => { "[cipher_suite]" => "[tls][cipher]" }
    rename => { "[device_id]" => "[observer][serial_number]" }
    rename => { "[dst]" => "[destination][ip]" }
    rename => { "[dst_port]" => "[destination][port]" }
    rename => { "[http_agent]" => "[user_agent][original]" }
    rename => { "[http_host]" => "[url][domain]" }
    rename => { "[http_method]" => "[http][request][method]" }
    rename => { "[http_refer]" => "[http][request][referrer]" }
    rename => { "[http_request_bytes]" => "[http][request][bytes]" }
    rename => { "[http_response_bytes]" => "[http][response][bytes]" }
    rename => { "[http_retcode]" => "[http][response][status_code]" }
    rename => { "[http_url]" => "[url][path]" }
    rename => { "[http_version]" => "[http][version]" }
    rename => { "[log_id]" => "[event][code]" }
    rename => { "[main_type]" => "[event][category]" }
    rename => { "[msg]" => "[message]" }
    rename => { "[msg_id]" => "[event][sequence]" }
    rename => { "[policy]" => "[rule][name]" }
    rename => { "[proto]" => "[network][transport]" }
    rename => { "[signature_subclass]" => "[vulnerability][category]" }
    rename => { "[owasp_top10]" => "[vulnerability][description]" }
    rename => { "[signature_cve_id]" => "[vulnerability][id]" }
    rename => { "[src]" => "[source][ip]" }
    rename => { "[src_port]" => "[source][port]" }
    rename => { "[monitor_status]" => "[event][outcome]" }
    rename => { "[threat_level]" => "[vulnerability][severity]" }
    rename => { "[type]" => "[event][dataset]" }
    rename => { "[user_name]" => "[user][name]" }
    rename => { "[sub_type]" => "[event][category]" }
    add_field => { "timestamp" => "%{date} %{time}" }
  }
  date {
    match => ["timestamp", "yyyy-MM-dd HH:mm:ss"]
  }
  output {
    elasticsearch {
      index => "forti_ecs"
      hosts => ["localhost:9200"]}
  }
}

```

El código funciona de la misma manera que el anterior. La diferencia es que este código incluye las sentencias necesarias para renombrar los campos y normalizarlos a ECS.

Además, se crea un índice denominado *forti_ecs*.

Para indexar manualmente los datos, se utiliza el siguiente comando:

```
/usr/share/logstash/bin/logstash --path.settings  
/etc/logstash -f /etc/logstash/conf.d/fortiweb2.conf
```

Para poder visualizar los datos, hay que crear un patrón de índice en Kibana. Para ello, desde el Index Management de Kibana se accede a “*Create Index pattern*”. A continuación:

- Se introduce el nombre deseado.
- Se introduce el valor @timestamp como Time Filter.

Se puede verificar desde el SIEM la correcta ingesta de los eventos de forma normalizada desde el apartado Eventos:

| @timestamp ↓ | url.path | event.category | vulnerability.severity | message | event.dataset |
|------------------------------|-------------------------------|------------------------|------------------------|-------------------------------|---------------|
| > Feb 7, 2020 @ 23:44:17.000 | /sisgen-web/rest/alertas/n... | Information Disclosure | Low | HTTP Header triggered sign... | attack |
| > Feb 7, 2020 @ 23:44:15.000 | /sisgen-web/rest/administ... | Information Disclosure | Low | HTTP Header triggered sign... | attack |
| > Feb 7, 2020 @ 23:44:14.000 | /sisgen-web/rest/alertas/n... | Information Disclosure | Low | HTTP Header triggered sign... | attack |
| > Feb 7, 2020 @ 23:44:14.000 | /sisgen-web/rest/alertas/n... | Information Disclosure | Low | HTTP Header triggered sign... | attack |
| > Feb 7, 2020 @ 23:44:12.000 | /sisgen-web/rest/alertas/n... | Information Disclosure | Low | HTTP Header triggered sign... | attack |
| > Feb 7, 2020 @ 23:44:12.000 | /sisgen-web/rest/alertas/n... | Information Disclosure | Low | HTTP Header triggered sign... | attack |
| > Feb 7, 2020 @ 23:44:11.000 | /sisgen-web/LOGIN/ | Information Disclosure | Low | HTTP Header triggered sign... | attack |
| > Feb 7, 2020 @ 23:44:11.000 | /sisgen-web/rest/alertas/n... | Information Disclosure | Low | HTTP Header triggered sign... | attack |
| > Feb 7, 2020 @ 23:44:11.000 | /sisgen-web/rest/usuarios... | Information Disclosure | Low | HTTP Header triggered sign... | attack |
| > Feb 7, 2020 @ 23:44:09.000 | /sisgen-web/rest/alertas/n... | Information Disclosure | Low | HTTP Header triggered sign... | attack |
| > Feb 7, 2020 @ 23:44:05.000 | /sisgen-web/rest/alertas/n... | Information Disclosure | Low | HTTP Header triggered sign... | attack |
| > Feb 7, 2020 @ 23:44:05.000 | /sisgen-web/rest/alertas/n... | Information Disclosure | Low | HTTP Header triggered sign... | attack |
| > Feb 7, 2020 @ 23:44:04.000 | /sisgen-web/rest/alertas/n... | Information Disclosure | Low | HTTP Header triggered sign... | attack |
| > Feb 7, 2020 @ 23:44:04.000 | /sisgen-web/rest/alertas/n... | Information Disclosure | Low | HTTP Header triggered sign... | attack |

Ilustración 14: Eventos de FortiWeb en SIEM

3.2.2 Integración de Eventos del S.O

La integración de eventos del S.O se realiza mediante los módulos *Filebeat*, *Winlogbeat* y *Auditbeat*.

En los siguientes anexos se puede ver el proceso de integración de estos eventos en Elasticsearch:

- [11.4 Añadir logs del sistema en Linux a Elasticsearch](#)
- [11.5 Añadir Eventos de Windows a ElasticSearch](#)
- [11.8 Instalar y configurar AuditBeat](#)

3.2.3 Integración y normalización de Eventos de Cisco Umbrella

Para la integración de los eventos de Cisco Umbrella se utilizará Logstash. Para ello, se ha definido un fichero de configuración con el siguiente código:

```
input {
  file {
    path => "/home/user/Descargas/umbrella_crlf.log"
    start_position => "beginning"
  }
}

filter {
  json {
    source => "message"
  }
  date {
    match => ["Timestamp", "yyyy-MM-dd HH:mm:ss"] }
}

output {
  elasticsearch {
    index => "umbrella"
    hosts => ["localhost:9200"]
  }
}
```

Lo que hace este código es filtrar el fichero de log y extraer los campos y sus valores, para poder trabajar con ellos. Como se trata de un JSON, se utiliza el filtro *json* de Logstash para conseguirlo. Una vez extraídos, se crea un índice denominado *umbrella* y se almacenan los datos.

Para indexar manualmente los datos, se utiliza el siguiente comando:


```
/usr/share/logstash/bin/logstash --path.settings
/etc/logstash -f /etc/logstash/conf.d/umbrella.conf
```

Para poder visualizar los datos, hay que crear un patrón de índice en Kibana. Para ello, desde el Index Management de Kibana se accede a “*Create Index pattern*”. A continuación:

- Se introduce el nombre deseado.
- Se introduce el valor @timestamp como Time Filter.

Para verificar que se están recibiendo datos, desde el Discover de Kibana se pueden visualizar los datos entrantes:

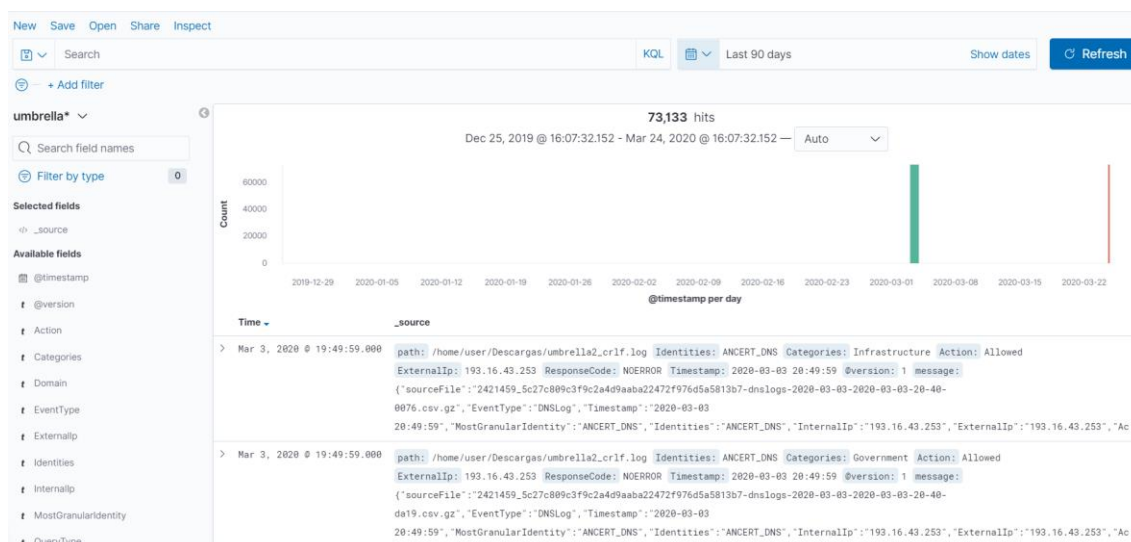


Ilustración 15: Eventos de Cisco Umbrella

El proceso de normalización de eventos de Cisco Umbrella, parte por renombrar los campos originales que utiliza y emplear los nombres de campo que corresponden con el ECS (Elastic, ECS DNS | ECS Guide 1.5, 2020). Al igual que en el caso de Fortiweb, este paso requiere de un paso previo de investigación para determinar la relación de los campos con el estándar ECS.

Para realizar este proceso, hay que modificar el archivo de configuración de Logstash anterior y modificar los campos originales. El archivo de configuración debe quedar de la siguiente manera:

```

input {
  file {
    path => "/home/user/Descargas/umbrella_crlf.log"
    start_position => "beginning"

  }
}

filter {
  json {
    source => "message"
  }
  mutate {
    rename => { "EventType" => "[dns][type]" }
    rename => { "ResponseCode" => "[dns][repose_code]" }
    rename => { "InternalIp" => "[source][ip]" }
    rename => { "Action" => "[event][action]" }
    rename => { "Identities" => "[event][module]" }
    rename => { "Domain" => "[dns][question][name]" }
    rename => { "Categories" => "[dns][answers][type]" }
    rename => { "QueryType" => "[dns][op_code]" }
    rename => { "sourceFile" => "[file][name]" }
    remove_field => ["ExternalIp", "MostGranularIdentity"]
  }
}

date {
  match => ["Timestamp", "yyyy-MM-dd HH:mm:ss" ]
}

output {
  elasticsearch {
    index => "umbrella_ecs"
    hosts => ["localhost:9200"]
  }
}

```

Este fichero funciona de la misma manera que su anterior. El único cambio es que una vez extraídos los campos, se renombran siguiendo los estándares del ECS (Elastic Common Schema) y se crea un nuevo índice denominado *umbrella_ecs* donde se almacenarán los datos.

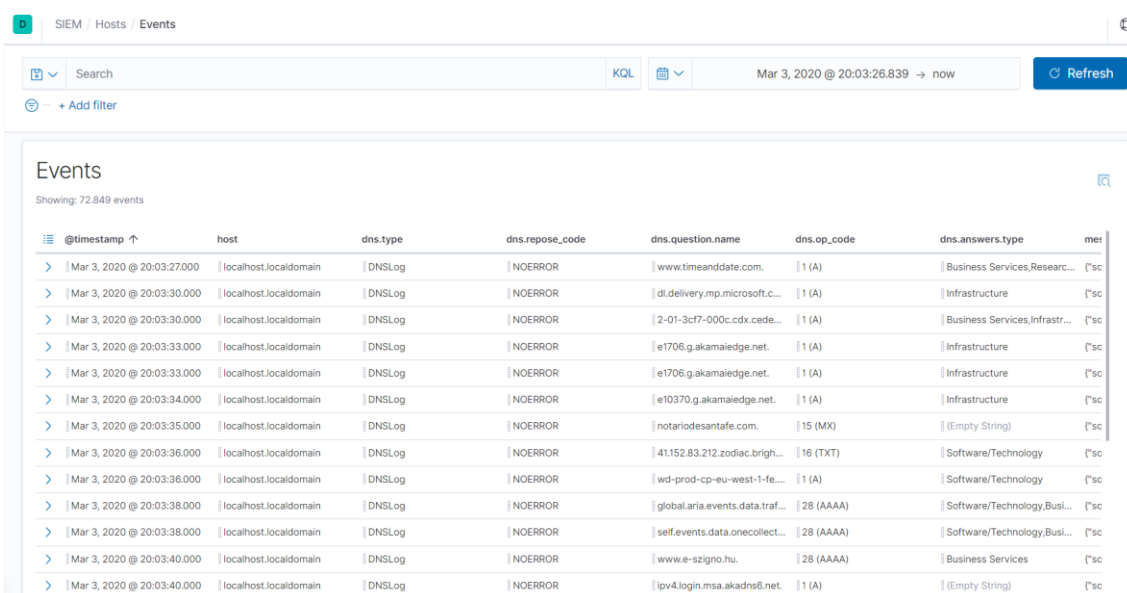
Para indexar manualmente los datos, se utiliza el siguiente comando:

```
/usr/share/logstash/bin/logstash --path.settings  
/etc/logstash -f /etc/logstash/conf.d/umbrella2.conf
```

Para poder visualizar los datos, hay que crear un patrón de índice en Kibana. Para ello, desde el Index Management de Kibana se accede a “*Create Index pattern*”. A continuación:

- Se introduce el nombre deseado.
- Se introduce el valor `@timestamp` como Time Filter.

Se puede verificar desde el SIEM la correcta ingesta de los eventos de forma normalizada desde el apartado Eventos:



The screenshot shows the SIEM interface with the 'Events' section active. The search bar contains 'KQL' and the time range is set to 'Mar 3, 2020 @ 20:03:26.839 -> now'. Below the search bar, there are 13 event entries in a table format. Each entry includes a timestamp, host, dns.type, dns.repose_code, dns.question.name, dns.op_code, dns.answers.type, and a message.

| @timestamp ↑ | host | dns.type | dns.repose_code | dns.question.name | dns.op_code | dns.answers.type | me |
|----------------------------|-----------------------|----------|-----------------|---------------------------------|-------------|-------------------------------|------|
| Mar 3, 2020 @ 20:03:27.000 | localhost.localdomain | DNSLog | NOERROR | www.timeanddate.com. | 1 (A) | Business Services,Researc... | ["sc |
| Mar 3, 2020 @ 20:03:30.000 | localhost.localdomain | DNSLog | NOERROR | dl.delivery.mp.microsoft.c... | 1 (A) | Infrastructure | ["sc |
| Mar 3, 2020 @ 20:03:30.000 | localhost.localdomain | DNSLog | NOERROR | 2-01-3cf7-000c.cdx.cede... | 1 (A) | Business Services,Infrastr... | ["sc |
| Mar 3, 2020 @ 20:03:33.000 | localhost.localdomain | DNSLog | NOERROR | e1706.g.akamaiedge.net. | 1 (A) | Infrastructure | ["sc |
| Mar 3, 2020 @ 20:03:33.000 | localhost.localdomain | DNSLog | NOERROR | e1706.g.akamaiedge.net. | 1 (A) | Infrastructure | ["sc |
| Mar 3, 2020 @ 20:03:34.000 | localhost.localdomain | DNSLog | NOERROR | e10370.g.akamaiedge.net. | 1 (A) | Infrastructure | ["sc |
| Mar 3, 2020 @ 20:03:35.000 | localhost.localdomain | DNSLog | NOERROR | notariodesantafe.com. | 15 (MX) | (Empty String) | ["sc |
| Mar 3, 2020 @ 20:03:36.000 | localhost.localdomain | DNSLog | NOERROR | 41152.83.212.zodiac.brigh... | 16 (TXT) | Software/Technology | ["sc |
| Mar 3, 2020 @ 20:03:36.000 | localhost.localdomain | DNSLog | NOERROR | wd-prod-cp-eu-west-1-fe... | 1 (A) | Software/Technology | ["sc |
| Mar 3, 2020 @ 20:03:38.000 | localhost.localdomain | DNSLog | NOERROR | global.aria.events.data.traf... | 28 (AAAA) | Software/Technology,Busi... | ["sc |
| Mar 3, 2020 @ 20:03:38.000 | localhost.localdomain | DNSLog | NOERROR | self.events.data.onecollect... | 28 (AAAA) | Software/Technology,Busi... | ["sc |
| Mar 3, 2020 @ 20:03:40.000 | localhost.localdomain | DNSLog | NOERROR | www.e-szigno.hu. | 28 (AAAA) | Business Services | ["sc |
| Mar 3, 2020 @ 20:03:40.000 | localhost.localdomain | DNSLog | NOERROR | ipv4.login.msa.akadns6.net. | 1 (A) | (Empty String) | ["sc |

Ilustración 16: Eventos de Cisco Umbrella en SIEM

3.2.4 Integración de Eventos de Palo Alto

Para la integración de los eventos de Palo Alto se utilizará Logstash. Para ello, se ha definido un fichero de configuración con el siguiente código:

```
input {
  file {
    path => "/home/user/Descargas/paloalto.log"
    start_position => "beginning"
  }
}
filter{
  kv {
    source => "message"
    field_split => "|"
  }
  date {
    match => [ "ReceiveTime", "yyyy/MM/dd HH:mm:ss" ]
    timezone => "America/Chicago"
  }
}
output
{
  elasticsearch { hosts => ["localhost:9200"]
  index => "paloalto"
  }
}
```

El código anterior filtra los campos y valores mediante *kv*, extrayendo las claves y valores separadas por el carácter "|". Una vez extraídos, se crea un index denominado *paloalto* y se almacenan los datos.

Para indexar manualmente los datos, se utiliza el siguiente comando:

```
/usr/share/logstash/bin/logstash --path.settings
/etc/logstash -f /etc/logstash/conf.d/paloalto.conf
```

Para poder visualizar los datos, hay que crear un patrón de índice en Kibana. Para ello, desde el Index Management de Kibana se accede a "*Create Index pattern*". A continuación:

- Se introduce el nombre deseado.
- Se introduce el valor *ReceiveTime* como Time Filter.

Para verificar que se están recibiendo datos, desde el Discover de Kibana se pueden visualizar los datos entrantes:



Ilustración 17: Eventos de PaloAlto

3.3 Configuración y uso de Elastic SIEM

El Elastic SIEM forma parte del conjunto de módulos Open-Source que implementa Elastic, por lo que disponemos de las características completas del SIEM en la Licencia Basic de Elastic, sin necesidad de disponer de una licencia de pago.

De todos modos, para poder utilizar las características de Elastic SIEM junto a la detección de anomalías con Machine Learning, es preciso activar la licencia Platino.

Para activar la Trial de la licencia Platino, es necesario entrar en los ajustes de Kibana, y una vez ahí entrar en el apartado License Management.

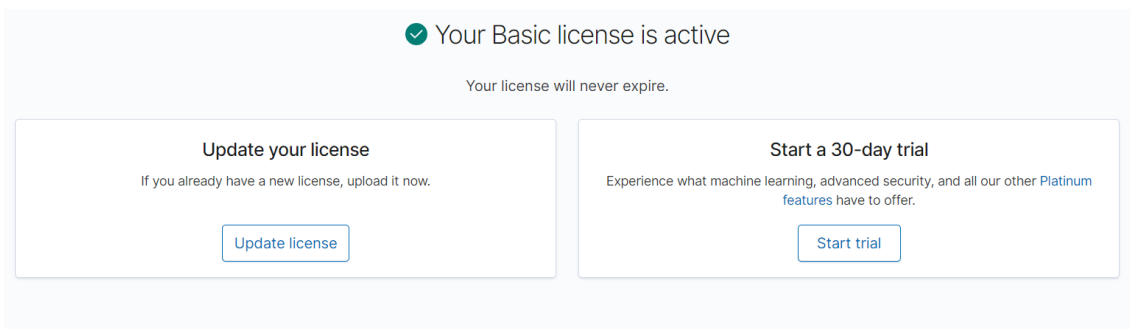


Ilustración 18: License Management

Una vez activo, hay que configurar los índices que se utilizarán en el SIEM para visualizar los datos. Esto se hace desde los *Ajustes Avanzados* dentro del menú *Ajustes* de Kibana. En los ajustes avanzados, se busca el apartado destinado al SIEM que incluye los índices que serán empleados:

Elasticsearch indices

Comma-delimited list of Elasticsearch indices from which the SIEM app collects events.

Default:
`apm-*-transaction*, auditbeat-*, endgame-*, filebeat-*, packetbeat-*, winlogbeat-*`

siem:defaultIndex

[Reset to default](#)

Ilustración 19: SIEM índices

Para poder visualizar los eventos de Cisco Umbrella y Fortiweb en el SIEM, hay que añadir los índices `umbrella_ecs` y `forti_ecs` a dicha lista. El acceso al SIEM se realiza desde Kibana. La vista es la siguiente:

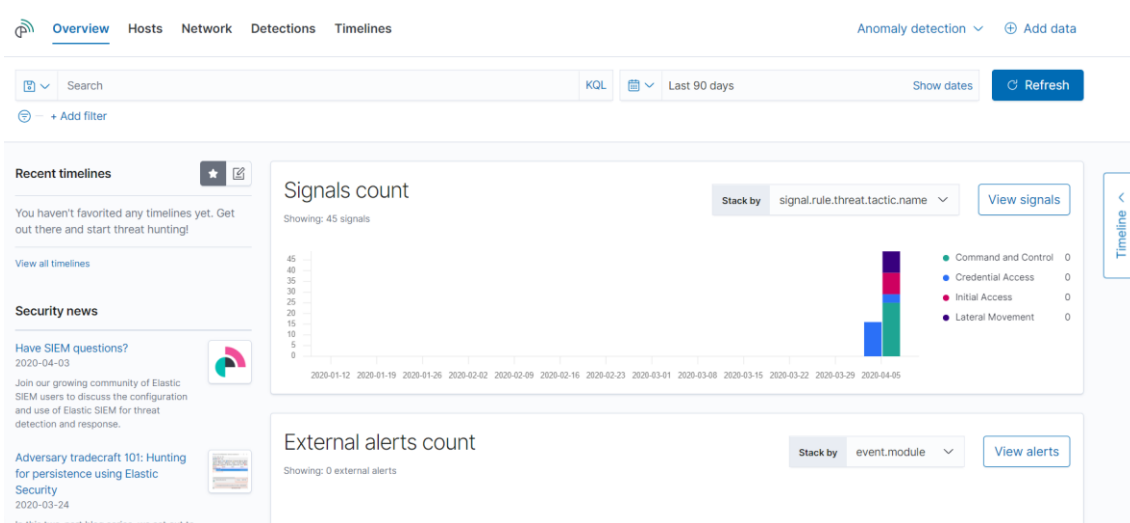


Ilustración 20: Elastic SIEM view

El SIEM está dividido por las siguientes áreas: *Overview*, *Hosts*, *Network*, *Detections* y *Timelines*. Cada área muestra distinta información:

- **Overview:** La vista resumen proporciona una visión de alto nivel de los eventos de seguridad disponibles para el análisis.

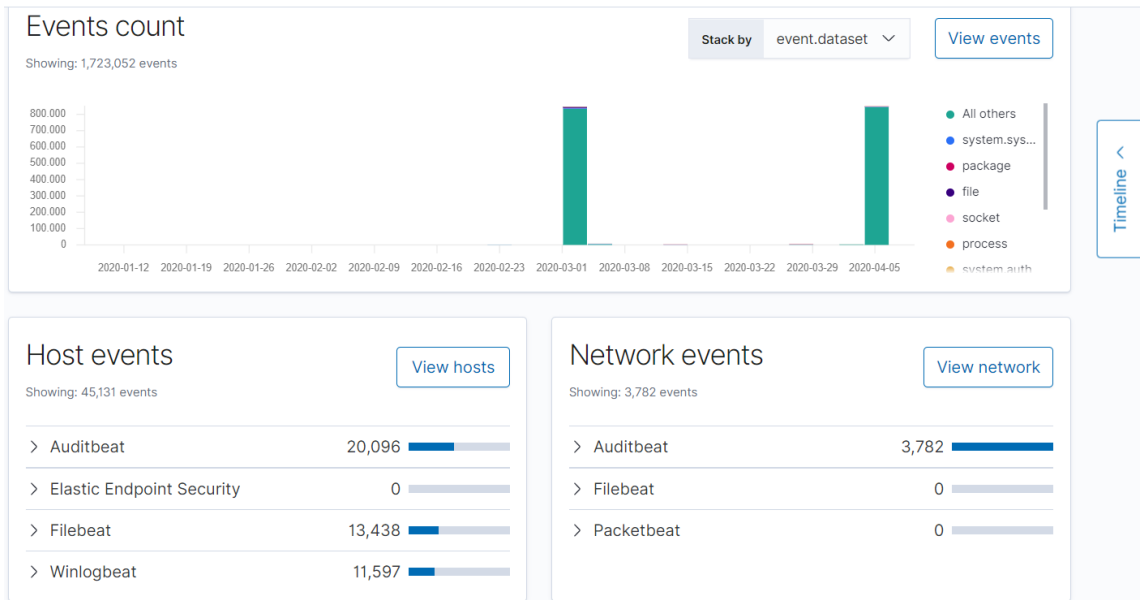


Ilustración 21: SIEM overview

- **Hosts:** La vista de Hosts proporciona métricas clave sobre los eventos de seguridad relacionados con los Hosts, y un conjunto de tablas de datos que le permiten interactuar con el Visor de Eventos en el Timeline.

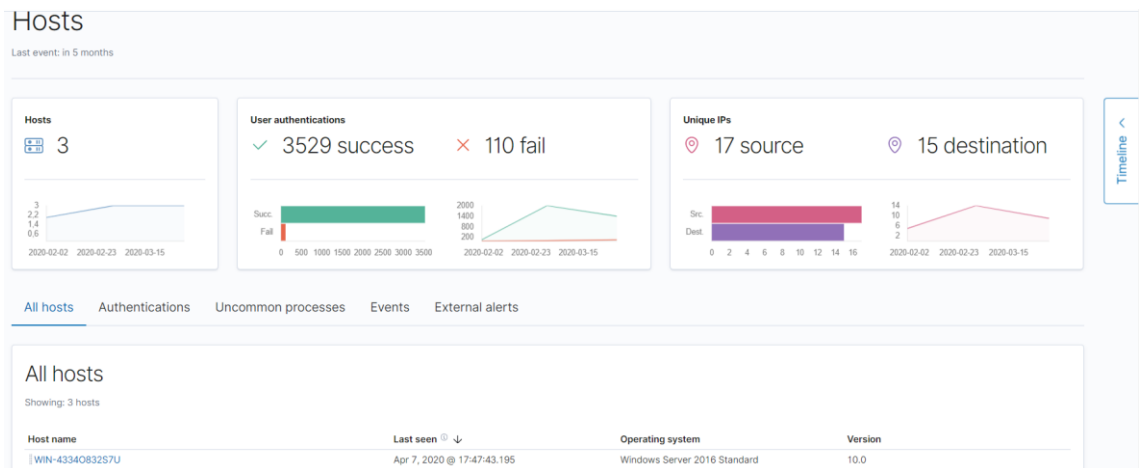


Ilustración 22: SIEM Hosts

Dentro de esta vista, se proporcionan métricas clave de seguridad como Autenticaciones, Procesos no comunes, Eventos y Alertas Externas. Este tipo de métricas proporcionan una vista panorámica del estado actual de la infraestructura monitorizada.

Las *autenticaciones* muestran los intentos de autenticación exitosos y/o fallidos, además de una tabla resumen de los últimos eventos de autenticación de los sistemas monitorizados.

Authentications



Authentications

Showing: 18 users

| User | Successes | Failures | Last success | Last successful source | Last successful destina... | Last failure | Last failed source | Last failed destination |
|-------------------|-----------|----------|----------------------------|--------------------------|----------------------------------|--------------|--------------------|-------------------------|
| WIN-4334083257U\$ | 3139 | 0 | Apr 7, 2020 @ 17:47:24.142 | fe80-2cdd-4830-fd11-3c6d | WIN-4334083257U.estudiant.es.uoc | --- | --- | --- |
| SYSTEM | 220 | 0 | Apr 7, 2020 @ 13:33:01.672 | --- | WIN-4334083257U.estudiant.es.uoc | --- | --- | --- |

Ilustración 23: SIEM Autenticaciones

Los *procesos no comunes* muestran aquellos procesos en ejecución de los sistemas monitorizados que pueden presentar un elemento de riesgo o sospecha, con el objetivo de prevenir procesos maliciosos que den indicio de que el sistema está comprometido. Se muestran aquellos procesos que no son comunes en los sistemas operativos, por lo que la gran mayoría podrían ser legítimos.

Uncommon processes

Showing: 121 processes

| Process name | Hosts | Instances | Host names | Last command | Last user |
|--|-------|-----------|-----------------|---|---------------------------|
| 80.0.3987.132.80.0.3987.122_chrome_updater.exe | 1 | 1 | WIN-4334083257U | C:\Program Files (x86)\Google\Update\Install\{E6C02D59-9B47-4EE2-91D6-037317E0E15B}\80.0.3987.132.80.0.3987.122_chrome_updater.exe +3 More | NT AUTHORITY\SYSTEM |
| 80.0.3987.149.80.0.3987.132_chrome_updater.exe | 1 | 1 | WIN-4334083257U | C:\Program Files (x86)\Google\Update\Install\{8A16027F-56DF-4F20-9D56-DDC8D21CBCC3}\80.0.3987.149.80.0.3987.132_chrome_updater.exe +3 More | NT AUTHORITY\SYSTEM |
| AM_Base_Patch1.exe | 1 | 1 | WIN-4334083257U | C:\Windows\SoftwareDistribution\Download\Install\AM_Base_Patch1.exe | NT AUTHORITY\SYSTEM |
| ApplicationFrameHost.exe | 1 | 1 | WIN-4334083257U | C:\Windows\system32\ApplicationFrameHost.exe +1 More | ESTUDIANTES\Administrador |
| DeviceCensus.exe | 1 | 1 | WIN-4334083257U | C:\Windows\system32\devicecensus.exe | NT AUTHORITY\SYSTEM |
| DismHost.exe | 1 | 1 | WIN-4334083257U | C:\Users\ADMINI~1\WIN\AppData\Local\Temp\80548341-1-6582-4D3E-BAB6-3EB739DA6FF9\dismhost.exe +1 More | ESTUDIANTES\Administrador |
| MusNotificationUx.exe | 1 | 1 | WIN-4334083257U | %systemroot%\system32\MusNotificationUX.exe +2 More | ESTUDIANTES\Administrador |
| RuntimeBroker.exe | 1 | 1 | WIN-4334083257U | C:\Windows\System32\RuntimeBroker.exe +1 More | ESTUDIANTES\Administrador |
| SHClient.exe | 1 | 1 | WIN-4334083257U | C:\Windows\System32\shclient.exe | NT AUTHORITY\SYSTEM |

Ilustración 24: SIEM Uncommon Processes

En el área *eventos* se tiene una vista sintetizada de los eventos del sistema, lo que permite tener una visión general de aquellas acciones que se están llevando a cabo.



Ilustración 25: SIEM Eventos

Por último las *Alertas Externas*, que en este proyecto no serán utilizadas puesto que no disponemos de elementos externos que proporcionen Alertas, tales como el Elastic Endpoint.

- Network:** La vista de la red proporciona métricas clave de la actividad de la red y un mapa interactivo, facilita el enriquecimiento del tiempo de investigación y proporciona tablas de eventos de la red que permiten la interacción con el Timeline. Se permite hacer una diferenciación del tráfico, pudiendo filtrar aquellos flujos que sean peticiones DNS, HTTP, TLS o Alertas Externas.

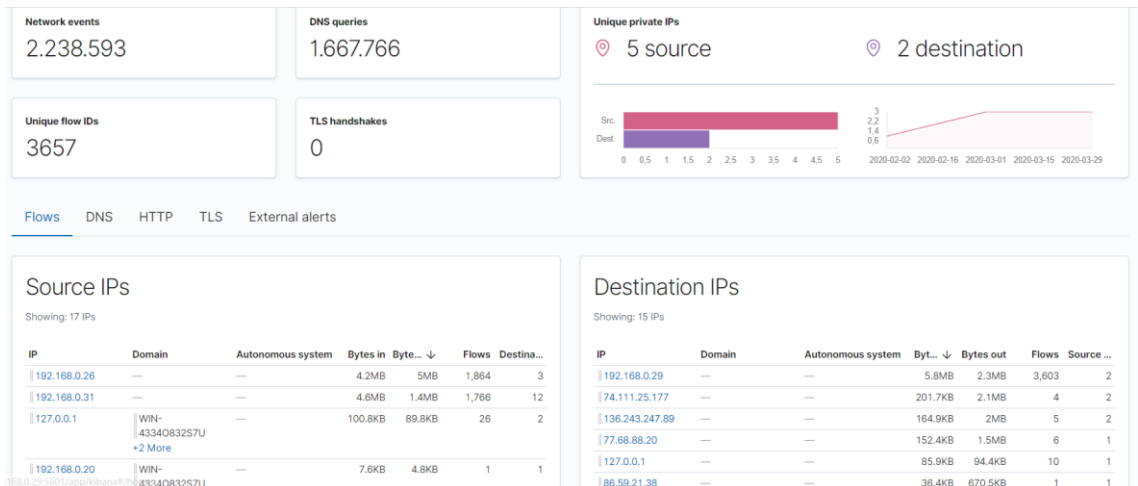


Ilustración 26: SIEM Network

- Detections:** La vista de Detecciones proporciona una visión general de todas las señales creadas por las reglas de detección de señales. También es el lugar donde se pueden habilitar las reglas preconstruidas y crear nuevas reglas.

Para poder utilizar la funcionalidad *Detections* del SIEM (Elastic, Elastic Guide 7.6 | Detection Engine Overview, 2020) hay que realizar una serie de configuraciones previas en la plataforma:

- Debe haber comunicaciones HTTPS entre Kibana y Elasticsearch ([Configuración de las comunicaciones mediante SSL/TLS](#)).
- Se debe establecer la configuración de `xpack.security.enabled` en `true` en el archivo de configuración `elasticsearch.yml`.
- Se debe añadir la configuración `xpack.encryptedSavedObjects.encryptionKey` con cualquier valor alfanumérico de al menos 32 caracteres en el archivo de configuración `kibana.yml`.

Además, para ver, crear o modificar las señales y las reglas de detección, se debe tener un usuario con los siguientes privilegios:

- El usuario debe tener privilegios de lectura en el *Kibana Space*, propiamente en la característica SIEM.
- El usuario debe tener privilegios sobre el clúster `manage_api_key`.
- El usuario debe tener permisos de lectura y escritura para el índice deseado.

Los usuarios se pueden crear desde los Ajustes de la interfaz de Kibana, en el apartado Security.

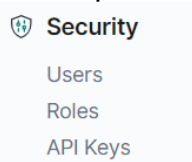


Ilustración 27: Security Elastic

Desde Security, se tiene acceso a la gestión RBAC (Role-based access control) de la plataforma. El proceso de creación de usuario es muy sencillo, se realiza desde el apartado Users. Por otro lado, la gestión de los roles se hace desde el apartado Roles.

Una vez configurado el SIEM, se puede empezar a crear reglas de Detección para visualizar en el SIEM los eventos sospechosos que están ocurriendo en la infraestructura monitorizada (Elastic, Elastic Guide 7.6 | Rules Create, 2020). Elastic tiene una serie de reglas pre-construidas, que dan soporte a la detección de anomalías de diversas índoles. Las más curiosas para la detección de Malware y otros ataques en Windows son:

- **DNS Activity to the Internet:** Esta regla detecta cuando un cliente de la red interna envía el tráfico DNS directamente a Internet. Este es un comportamiento atípico para una red administrada, y puede ser indicativo de malware, exfiltración, comando y control, o, simplemente, mala configuración.

- **Adding Hidden File Attribute via Attrib:** Es común que el Malware trate de añadir el atributo "oculto" a ciertos archivos para ocultarlos del usuario en un intento de evadir la detección.
- **Local Scheduled Task Commands:** Una tarea programada puede ser utilizada por un atacante para establecer la persistencia, moverse lateralmente y/o aumentar los privilegios.
- **Unusual Network Connection via RunDLL32:** Identifica instancias inusuales de rundll32.exe haciendo conexiones de red de salida. Esto puede indicar presencia de un Malware y permite identificar DLLs maliciosas.
- **PsExec Network Connection:** Identifica el uso de la herramienta de SysInternals PsExec.exe haciendo una conexión de red. Esto podría ser una indicación de movimiento lateral.
- **Unusual Parent-Child Relationship:** Identifica los programas de Windows que se ejecutan a partir de procesos parentales inesperados. Esto podría indicar un enmascaramiento u otra actividad extraña en un sistema.
- **User Account Creation:** Identifica los intentos de crear nuevos usuarios locales. A veces lo hacen los atacantes para aumentar el acceso a un sistema o dominio.
- **System Shells via Services:** Los servicios de Windows normalmente se ejecutan como SYSTEM y pueden ser utilizados como una oportunidad de escalada de privilegios. El malware o los probadores de penetración pueden ejecutar un shell como un servicio para obtener los permisos del sistema.
- **Whoami Process Activity:** Identifica el uso de whoami.exe que muestra información de usuario, grupo y privilegios del usuario que está actualmente conectado al sistema local.
- **Disable Windows Firewall Rules via Netsh:** Identifica el uso del netsh.exe para desactivar o debilitar el firewall local. Los atacantes utilizarán esta herramienta de línea de comandos para desactivar el firewall durante la resolución de problemas o para permitir la movilidad de la red.

Estos son algunas de las reglas más útiles para el proyecto, aunque existen más. Dentro de este conjunto de reglas pre-construidas, existen reglas específicas para sistemas operativos Linux y para Windows. También es preciso destacar que existe una gran cantidad de reglas específicas para detectar Malware, pero requieren disponer del Elastic Endpoint para su uso.

Por otro lado, existe también la posibilidad de crear nuestras propias reglas. A continuación se explicará el proceso de creación de una regla para detectar ataques de fuerza bruta al Login en Windows:

- Desde el SIEM, en Detections se accede a al Manage signal detection rules.
- Se hace click en New Rule para empezar el proceso de creación de reglas. Este proceso se realiza por fases.
- La primera fase es la de definición. En esta fase, se utilizan filtros y los campos de consulta para crear los criterios utilizados para la detección de señales. En este caso, se necesita detectar los eventos de seguridad de Windows cuyo id sea 4625.

Definition

Index patterns

apm-*-transaction* × auditbeat-* × endgame-* × filebeat-* × packetbeat-* × winlogbeat-* × prueba_ciscou* ×

forti_ecs* ×

Enter the pattern of Elasticsearch indices where you would like this rule to run. By default, these will include index patterns defined in SIEM advanced settings.

Custom query Import query from saved timeline

winlog.event_id : "4625" KQL

+ Add filter

Ilustración 28: Definición de la regla

- La segunda fase es la de descripción. Aquí se establece la información relacionada con la alerta. Se indica la severidad, el Risk Score y una breve descripción de lo que hace la rule. Además, se pueden añadir tácticas de MITRE ATT&CK™.

About

Description

Se realiza un ataque de password cracking al login

Severity

● High

Risk score

73

Investigate detections using this timeline template

Default blank timeline

MITRE ATT&CK™

Credential Access (TA0006)

└ Brute Force (T1110)

Tags

Windows

Ilustración 29: Descripción de la regla

- Por último, se indica la periodicidad con la que se verificará la detección:

3 Schedule rule

Runs every

5 Minutes

Rules run periodically and detect signals within the specified time frame.

Additional look-back time Optional

1 Minutes

Adds time to the look-back period to prevent missed signals.

Create rule without activating it Create & activate rule

Ilustración 30: Programación de la regla

Se puede comprobar que la regla esté funcionando realizando como prueba un ataque de fuerza bruta en el Servidor Windows Server, ya sea de forma manual o mediante el uso de herramientas que lo automaticen:

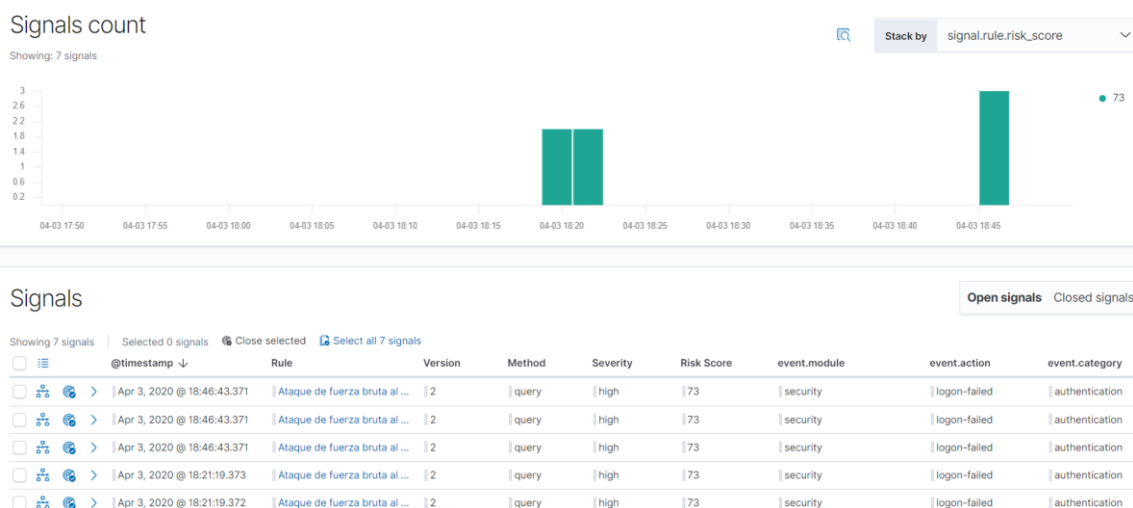


Ilustración 31: Detección de ataque por Fuerza Bruta

Otra regla útil que se puede crear sería para detectar cuando se para la protección en tiempo real de Windows Defender, o cuando Windows Defender detecta un Malware. El resumen de las reglas quedaría de la siguiente manera:

| | | |
|--|---|--|
| Definition Index patterns winlogbeat-* Custom query winlog.event_id: "5001" | About Description Esta regla detecta si la protección en tiempo real de Windows Defender se ve detenida. Severity Critical Risk score 100 Investigate detections using this timeline template Default blank timeline Tags Windows | Schedule Runs every 5m Additional look-back time 1m |
|--|---|--|

Ilustración 32: Parada de Windows Defender

Definition

Index patterns
winlogbeat-*

Custom query
winlog.event_id: "1116" or winlog.event_id: "1117" or winlog.event_id: "1015"

About

Description
Esta regla alerta si Windows Defender detecta Malware o comportamientos sospechosos en el servidor.

Severity
● Medium

Risk score
65

Investigate detections using this timeline template
Default blank timeline

Tags
Windows

Schedule

Runs every
1m

Additional look-back time
1m

Ilustración 33: Detección de Malware en Windows Defender

Signals

Showing 7 signals | Selected 0 signals | Close selected | Select all 7 signals

| | @timestamp | Rule | Version | Method | Severity | Risk Score | event.module | event.action | event.category | host.name | user.name | source.ip |
|--|-----------------------------|-------------------------------|---------|--------|----------|------------|--------------|--------------|----------------|----------------|---------------|-----------|
| | Apr 15, 2020 @ 21:00:20.704 | Malware detectado en WL... | 2 | query | medium | 65 | --- | --- | --- | WIN-4334083... | --- | --- |
| | Apr 15, 2020 @ 20:59:41.356 | Parada de Protección en T... | 2 | query | critical | 100 | --- | --- | --- | WIN-4334083... | --- | --- |
| | Apr 14, 2020 @ 21:56:34.846 | Ataque de fuerza bruta al ... | 2 | query | high | 73 | security | logon-failed | authentication | WIN-4334083... | Administrador | 127.0.0.1 |
| | Apr 14, 2020 @ 21:56:34.846 | Ataque de fuerza bruta al ... | 2 | query | high | 73 | security | logon-failed | authentication | WIN-4334083... | Administrador | 127.0.0.1 |
| | Apr 14, 2020 @ 21:56:34.846 | Ataque de fuerza bruta al ... | 2 | query | high | 73 | security | logon-failed | authentication | WIN-4334083... | Administrador | 127.0.0.1 |
| | Apr 14, 2020 @ 21:56:34.845 | Ataque de fuerza bruta al ... | 2 | query | high | 73 | security | logon-failed | authentication | WIN-4334083... | Administrador | 127.0.0.1 |
| | Apr 14, 2020 @ 21:56:34.845 | Ataque de fuerza bruta al ... | 2 | query | high | 73 | security | logon-failed | authentication | WIN-4334083... | Administrador | 127.0.0.1 |

Ilustración 34: Alertas detección de Malware

Para que funcionen estas detecciones, es preciso añadir a la configuración de winlogbeat las siguientes líneas:

```
winlogbeat.event_logs:
  - name:Microsoft-Windows-Windows Defender/Operational
    include_xml: true
```

Volcándonos en otro tipo de logs, resulta interesante también crear alertas para detectar consultas a DNS maliciosas, las cuales conduzcan a páginas con Malware o sean intentos de Phishing. Para detectar este tipo de anomalías se utilizarán los eventos de Cisco Umbrella, ya que filtra todas las peticiones DNS de la red.

El proceso de creación de la regla es el mismo que el anterior. A continuación se mostrará el resumen de la configuración de la regla:

Definition

Index patterns
apm-* transaction* auditbeat-*
endgame-* filebeat-* packetbeat-*
winlogbeat-* prueba_ciscou* forti_ecs*

Custom query
event.action: Blocked and (dns.answers.type: Malware or dns.answers.type: Web Spam, Phishing, Illegal Activities)

About

Description
Esta regla muestra las peticiones DNS con destinos maliciosos o intentos de Phishing

Severity
● Medium

Risk score
50

Investigate detections using this timeline template
Default blank timeline

Tags
Windows

Schedule

Runs every
5m

Additional look-back time
1m

Ilustración 35: Regla de Cisco Umbrella para detectar Phishing

Además de intentos de Phishing, se podrían crear reglas para detectar consultas con destinos no deseados, tales como páginas webs con contenidos pornográficos, páginas de apuestas y casinos, videojuegos, etc.

Por otro lado, aprovechando que disponemos de un dispositivo Fortinet en la infraestructura, resulta muy interesante crear reglas para aprovechar las capacidades de un WAF a la hora de detectar ataques en tiempo real. Para ello, vamos a crear dos reglas que detecten vulnerabilidades con severidad alta y crítica con el objetivo de detectar aquellas que sean altamente peligrosas. El proceso de creación de la regla es el mismo que el explicado anteriormente, por lo que tan solo se mostrará el resumen de las reglas:

Ilustración 36: Regla para detectar vulnerabilidades críticas en Fortiweb

Ilustración 37: Regla para detectar vulnerabilidades altas en Fortiweb

Finalmente, tras la creación de las reglas oportunas, se tiene un panel con las diferentes señales que el SIEM va detectando sobre la infraestructura:

Signals Open signals Closed signals

Showing 40 signals | Selected 0 signals | Close selected | Select all 40 signals

| <input type="checkbox"/> | <input type="checkbox"/> | @timestamp ↓ | Rule | Version | Method | Severity | Risk Score | event.module | event.action | event.category |
|--------------------------|--------------------------|----------------------------|--------------------------------|---------|--------|----------|------------|--------------|--------------|-----------------|
| <input type="checkbox"/> | <input type="checkbox"/> | Apr 7, 2020 @ 14:12:03.653 | SSH (Secure Shell) to the I... | 1 | query | low | 21 | system | network_flow | network_traffic |
| <input type="checkbox"/> | <input type="checkbox"/> | Apr 7, 2020 @ 14:12:03.653 | SSH (Secure Shell) to the I... | 1 | query | low | 21 | system | network_flow | network_traffic |
| <input type="checkbox"/> | <input type="checkbox"/> | Apr 7, 2020 @ 14:11:45.543 | DNS Activity to the Internet | 1 | query | medium | 47 | system | network_flow | network_traffic |
| <input type="checkbox"/> | <input type="checkbox"/> | Apr 7, 2020 @ 14:11:45.543 | DNS Activity to the Internet | 1 | query | medium | 47 | system | network_flow | network_traffic |
| <input type="checkbox"/> | <input type="checkbox"/> | Apr 7, 2020 @ 14:07:08.762 | SSH (Secure Shell) from th... | 1 | query | medium | 47 | system | network_flow | network_traffic |
| <input type="checkbox"/> | <input type="checkbox"/> | Apr 7, 2020 @ 14:07:02.626 | SSH (Secure Shell) to the I... | 1 | query | low | 21 | system | network_flow | network_traffic |
| <input type="checkbox"/> | <input type="checkbox"/> | Apr 7, 2020 @ 13:41:42.572 | DNS Activity to the Internet | 1 | query | medium | 47 | system | network_flow | network_traffic |
| <input type="checkbox"/> | <input type="checkbox"/> | Apr 7, 2020 @ 13:41:42.572 | DNS Activity to the Internet | 1 | query | medium | 47 | system | network_flow | network_traffic |
| <input type="checkbox"/> | <input type="checkbox"/> | Apr 7, 2020 @ 13:41:42.572 | DNS Activity to the Internet | 1 | query | medium | 47 | system | network_flow | network_traffic |
| <input type="checkbox"/> | <input type="checkbox"/> | Apr 7, 2020 @ 13:36:42.291 | DNS Activity to the Internet | 1 | query | medium | 47 | system | network_flow | network_traffic |
| <input type="checkbox"/> | <input type="checkbox"/> | Apr 7, 2020 @ 13:35:32.676 | Ataque de fuerza bruta al ... | 2 | query | high | 73 | security | logon-failed | authentication |

Ilustración 38: Señales detectadas en el SIEM

El motor de detección, por defecto, cuando crea una regla retrocede de 5 minutos en 5 minutos buscando señales, por lo que es un sistema

enfocado a detección en tiempo real y no en base a históricos. Esto es fundamental para recoger información de lo que está sucediendo en la infraestructura.

Aparte de la utilidad en cuanto a monitorización, las detecciones se utilizan como evidencias para realizar investigaciones.

Las señales tienen dos estados, abiertas o cerradas. Por defecto, todas las señales de la tabla aparecen como abiertas, aunque se pueden cerrar para indicar que no se necesita realizar ninguna investigación.

Mediante el SIEM se puede filtrar por el estado de las señales, pudiendo ver en todo momento aquellas que estén abiertas o cerradas respectivamente en un momento dado.

- **Timelines:** Las líneas de tiempo se utilizan como espacio de trabajo para investigaciones de alerta o threat hunting. Se pueden añadir datos de múltiples índices a una línea de tiempo, lo que permite investigar amenazas complejas, como por ejemplo, el movimiento lateral de un malware a través de los hosts de la red.

Por tanto, el Timeline se utiliza para realizar el proceso de investigación de una señal. Desde la tabla de señales, se pueden enviar directamente al Timeline clicando en el icono "View in timeline".

Para utilizar Timeline, se pueden arrastrar objetos de interés al Visor de eventos para crear exactamente el filtro de consulta que se necesita para llegar al fondo de una alerta. Se pueden arrastrar elementos desde los widgets de tabla dentro de las páginas de Hosts y de Network, o incluso desde la propia Timeline.

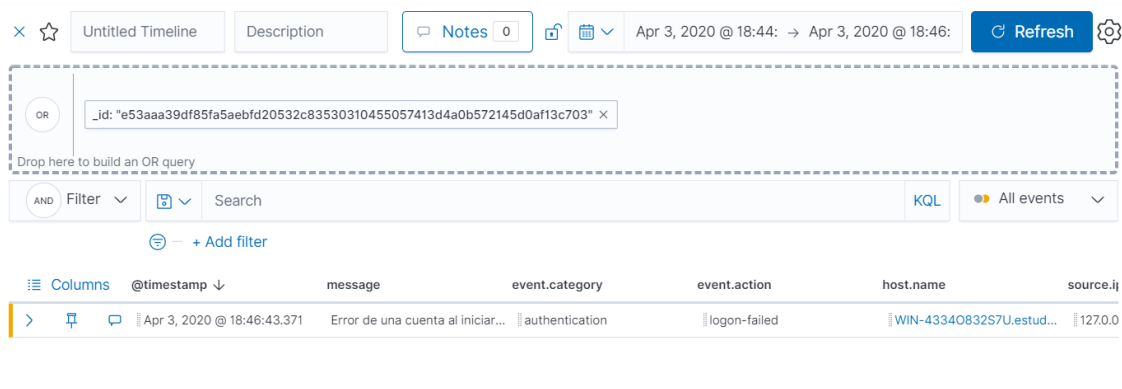


Ilustración 39: Timeline fuerza bruta

Los eventos de seguridad en Timeline se presentan en una vista renderizada e intuitiva, que permite una comprensión rápida de lo que está sucediendo.

Resumiendo, el Timeline proporciona las siguientes funcionalidades:

- Agregar, eliminar, reordenar o cambiar el tamaño de las columnas de la línea de tiempo.
- Guardar, abrir y enumerar líneas de tiempo.

- Agregar notas a eventos individuales.
- Agregar notas de investigación para toda la línea de tiempo.
- Anclar eventos a la línea de tiempo para tener persistencia.

4.0 Machine Learning

4.1 Definición de Machine Learning

El Machine Learning, o aprendizaje automático, es una parte de la inteligencia artificial utilizada para que las máquinas aprendan por sí mismas, sin necesidad de ser expresamente programadas para ello. Mediante el uso de algoritmos, las máquinas aprenden a identificar patrones, y con ello pueden realizar predicciones, convirtiéndose en sistemas inteligentes y autónomos.

Existen varios tipos de algoritmos de Machine Learning:

- **Aprendizaje por refuerzo:** Se produce cuando el algoritmo aprende por medio de prueba y error, a través de su propia experiencia, hasta alcanzar la mejor manera de completar la tarea.
Este tipo de aprendizaje tiene tres componentes principales: el agente (el que aprende o toma decisiones), el entorno (todo con lo que interactúa el agente) y las acciones (lo que el agente puede hacer). El objetivo es que el agente elija acciones que logren la cantidad de recompensas esperadas en cierta cantidad de tiempo. El agente logrará la meta mucho más rápido si aplica una buena política, de tal forma que el objetivo en el aprendizaje con refuerzo es aprender la mejor política.
- **Aprendizaje supervisado:** Los algoritmos cuentan con una fase de aprendizaje previa, basada en un sistema de etiquetas asociadas a unos datos que les permiten tomar decisiones o hacer predicciones.
El algoritmo de aprendizaje recibe un conjunto de entradas junto con los resultados correctos correspondientes, y el algoritmo aprende comparando el resultado real con los resultados correctos para encontrar errores. Luego modifica el modelo en consecuencia. A través de métodos como la clasificación, regresión, predicción y aumento de gradiente, el aprendizaje supervisado utiliza patrones para predecir los valores de la etiqueta en datos no etiquetados adicionales.
- **Aprendizaje semi-supervisado:** El algoritmo utiliza datos etiquetados y no etiquetados para el entrenamiento. Este tipo de aprendizaje se puede utilizar con métodos como la clasificación, regresión y predicción.
El aprendizaje semi-supervisado es de utilidad cuando el costo asociado con el etiquetado es demasiado alto para permitir un proceso de entrenamiento completamente etiquetado.
- **Aprendizaje no supervisado:** Los algoritmos no identifican patrones en bases de datos etiquetadas, sino que buscan similitudes. En este caso,

los algoritmos no están programados para detectar un tipo específico de datos, sino que se enfrentan de forma bruta a los datos, con el objetivo de encontrar patrones que permitan organizarlos de alguna manera.

A medida que los conjuntos de datos aumentan en tamaño y complejidad, el esfuerzo humano necesario para inspeccionar los tableros o mantener las reglas para detectar problemas de infraestructura, ataques cibernéticos o problemas en los hosts se vuelve difícil.

Para disminuir la dificultad y aumentar la eficacia de las prácticas seguidas, se utilizará el Machine Learning. El objetivo del proyecto es utilizarlo para mejorar la detección de anomalías y valores atípicos.

La función de detección de anomalías junto a Machine Learning de Elastic modela automáticamente el comportamiento normal de los datos de las series temporales, las tendencias de aprendizaje, la periodicidad, etc. Esto se hace en tiempo real, lo que permite identificar anomalías, agilizar el análisis un incidente y reducir los falsos positivos.

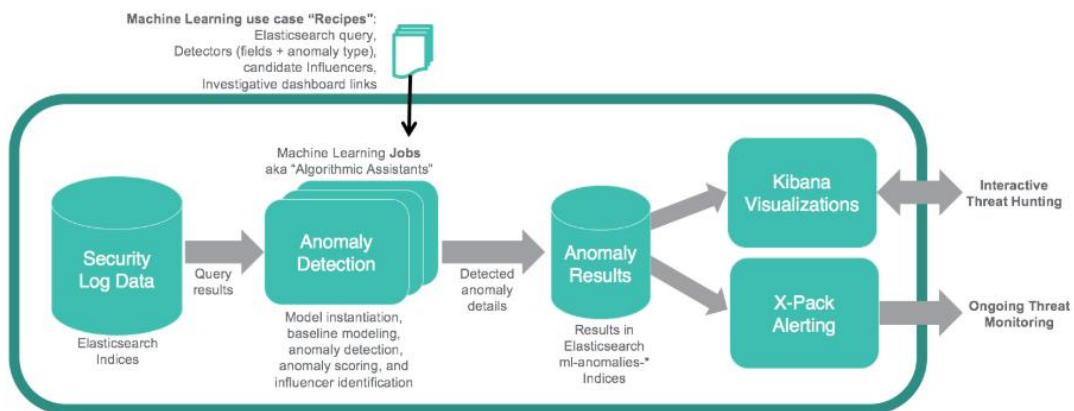


Ilustración 40: Flujo Machine Learning

4.2 Configuración y uso de Machine Learning

El paso previo para poder utilizar el módulo de Machine Learning, tal y como se explicó anteriormente, es el de activar la licencia Premium de Elastic.

Una vez activada la licencia, se puede acceder a las características de Machine Learning desde el menú de Kibana.

Overview Anomaly Detection Data Frame Analytics Data Visualizer

Getting started

Welcome to Machine Learning. Get started by reviewing our documentation or creating a new job. For more information about machine learning in the Elastic stack please see here. We recommend using Elasticsearch's transforms to create feature indices for analytics jobs.

Feedback

If you have input or suggestions regarding your experience with Machine Learning please feel free to submit feedback online.

Anomaly Detection

Active ML Nodes: 1 Total jobs: 16 Open jobs: 6 Closed jobs: 0 Active datafeeds: 6

| Group ID ↑ | Max anomaly score | Jobs in group | Latest timestamp | Docs processed | Actions |
|----------------|-------------------|---------------|------------------|----------------|---------|
| authentication | △ | 1 | | 0 | Explore |
| authentication | △ | 2 | | 0 | Explore |
| authentication | △ | 5 | | 0 | Explore |
| process | △ | 9 | | 0 | Explore |
| user | △ | 16 | | 0 | Explore |
| user | △ | 5 | | 0 | Explore |
| authentication | △ | 10 | | 0 | Explore |

Rows per page: 10

Refresh Manage jobs

Ilustración 41: Vista Machine Learning

El Machine Learning funciona mediante el uso de *trabajos*. Los trabajos de Machine Learning contienen la información de configuración y los metadatos necesarios para realizar una tarea analítica. También contienen los resultados de la tarea analítica.

El SIEM viene con *trabajos* de detección de anomalías de Machine Learning preconstruidos (Elastic, Machine Learning trabajos preconstruidos | Elastic Guide 7.6.0, 2020), que sirven para detectar automáticamente anomalías de los host y de la red. También existe la posibilidad de crear *trabajos* personalizados.

Los trabajos pre-construidos que se pueden utilizar para la detección de Malware son los siguientes:

- **windows_anomalous_network_activity_ecs:** Identifica los procesos del sistema operativo que no suelen utilizar la red pero que tienen una actividad de red inesperada, lo que puede indicar actividad de mando y control, movimiento lateral, persistencia o exfiltración de datos.

Un proceso con actividad inusual en la red puede denotar explotación o inyección de procesos, donde el proceso se utiliza para ejecutar mecanismos de persistencia que permiten a un actor malicioso el acceso o control remoto del host, la exfiltración de datos y la ejecución de aplicaciones de red no autorizadas.

- **windows_anomalous_process_all_hosts_ecs:** Busca procesos raros que se ejecuten en múltiples hosts de una infraestructura o red.

Esto reduce la detección de falsos positivos, ya que los procesos de mantenimiento automatizados, de forma general, sólo se ejecutan en una sola máquina.

- **windows_anomalous_user_name_ecs:** Busca la actividad de los usuarios que no están normalmente activos, lo que puede indicar cambios no autorizados, actividad de usuarios no autorizados, movimiento lateral y credenciales comprometidas.

En las organizaciones, no se suelen crear nuevos nombres de usuario aparte de tipos específicos de actividades del sistema, como la creación

de nuevas cuentas para nuevos empleados. Estas cuentas de usuario se convierten rápidamente en activas y rutinarias.

Los eventos de usuarios inactivos pueden apuntar a una actividad sospechosa.

Los nombres de usuario inusuales también pueden indicar pivoting, donde las credenciales comprometidas se utilizan para intentar moverse lateralmente de un host a otro.

- **rare_process_by_host_windows_ecs:** Identifica procesos raros que no suelen ejecutarse en hosts individuales, lo que puede indicar la ejecución de servicios no autorizados, programas malignos o mecanismos de persistencia.

Los procesos se consideran raros cuando sólo se ejecutan ocasionalmente en comparación con otros procesos que se ejecutan en el host.

- **suspicious_login_activity_ecs:** Identifica un número inusualmente alto de intentos de autenticación, lo que puede dar lugar a un ataque de fuerza bruta (password cracking).
- **windows_anomalous_path_activity_ecs:** Identifica procesos iniciados a partir de carpetas atípicas del sistema de archivos, lo que podría indicar mecanismos de ejecución o persistencia de malware.

En los entornos corporativos Windows, la instalación de software se suele gestionar de forma centralizada, y no es habitual que los programas se ejecuten desde directorios de usuario o temporales. Los procesos ejecutados desde estas ubicaciones pueden indicar que un usuario descargó software directamente de Internet o que un script malicioso/macro ejecutó malware.

- **windows_anomalous_process_creation:** Identifica relaciones inusuales de procesos padre-hijo que pueden indicar la ejecución de malware o mecanismos de persistencia.

Los guiones maliciosos suelen recurrir a otras aplicaciones y procesos como parte de su carga de explotación. La vigilancia y la identificación de relaciones de procesos anómalos es una excelente forma de detectar malware nuevo y emergente que aún no es reconocido por los antivirus.

- **windows_anomalous_script:** Busca scripts de PowerShell con características de datos inusuales, como la ofuscación, que puede ser una característica de los bloques de texto de los scripts malignos de PowerShell.
- **windows_anomalous_service:** Busca servicios inusuales de Windows que puedan indicar la ejecución de servicios no autorizados, malware o mecanismos de persistencia.

En los entornos corporativos de Windows, los hosts no suelen ejecutar muchos servicios raros o únicos. Este *trabajo* ayuda a detectar malware y mecanismos de persistencia que se han instalado y ejecutado como un servicio.

- **windows_rare_user_runas_event:** Busca cambios inusuales de contexto de usuario utilizando el comando *runas* o técnicas similares, que pueden indicar la toma de cuenta o la escalada de privilegios utilizando cuentas comprometidas.
- **windows_rare_user_type10_remote_login:** Busca inicios de sesión inusuales del protocolo de escritorio remoto (RDP), que pueden indicar la toma de cuenta o la persistencia de credenciales usando cuentas comprometidas. Los ataques RDP, como BlueKeep o Wannacry, también tienden a usar nombres de usuario inusuales.
- **packetbeat_dns_tunneling:** El tunelado del DNS puede utilizarse para la actividad de Command&Control, la persistencia o la exfiltración de datos.
Este trabajo busca un número inusualmente grande de consultas DNS para un solo dominio DNS de nivel superior, que a menudo se utiliza para el tunelado DNS.
- **packetbeat_rare_dns_question, packetbeat_rare_server_domain y packetbeat_rare_urls:** La búsqueda de consultas DNS raras e inusuales pueden indicar que existe actividad dirigida a dominios maliciosos o sospechosos. Esto puede deberse al acceso inicial, a la persistencia, a la actividad de Command&Control o a la exfiltración.
- **packetbeat_rare_user_agent:** Busca agentes de usuario (User Agent) raros e inusuales que pueden indicar que la actividad de navegación por la web se está realizando mediante un proceso inusual, que no es un navegador web. Esto puede deberse a la persistencia, la actividad de Command&Control o la exfiltración. Los agentes de usuario poco comunes que llegan de fuentes remotas a destinos locales suelen ser el resultado de escáneres, bots y web scrappers.

Estos trabajos analizan hasta dos semanas de datos históricos desde el momento que se habilitan. Además, desde su activación analizan los datos en tiempo real.

4.3 Detección de Anomalías

Una percepción errónea bastante común acerca del machine Learning en la ciberseguridad es pensar que funciona como una “*caja mágica*” de algoritmos a la que le pasas los datos y empiezan a producir resultados asombrosos.

La mejor forma para entender el Machine Learning en el contexto de la ciberseguridad sería verlo como un arsenal de "asistentes algorítmicos", que ayudan al equipo de seguridad a automatizar las tareas de análisis de datos,

buscando anomalías y patrones potencialmente anómalos, pero siempre bajo la supervisión de personas con conocimientos en ciberseguridad.

El objetivo de utilizar Machine Learning en la seguridad es el descubrimiento de comportamientos de ataque elementales, que pueden ser difíciles de detectar por otros medios. Los comportamientos de ataque elementales incluyen actividades como el tunneling del DNS, la exfiltración de datos web, la ejecución de procesos endpoint sospechosos, etc.

Dentro del Machine Learning, la Detección de Anomalías es la técnica de identificación de eventos u observaciones poco frecuentes que pueden levantar sospechas al ser estadísticamente diferentes del resto de las observaciones. Este comportamiento anómalo se traduce típicamente en algún tipo de problema como los comentados anteriormente.

Una anomalía se puede clasificar en tres categorías:

- **Anomalía de punto:** Una sola instancia de datos es anómala si está demasiado lejos del resto. Por ejemplo, un proceso inusual del sistema operativo podría indicar un potencial ciberataque.
- **Anomalía contextual:** Una observación es una anomalía contextual si es una anomalía debido al contexto de la observación. Es decir, la anomalía es específica del contexto. Este tipo de anomalía es común en los datos de series temporales. Por ejemplo, que una web reciba muchas visitas en determinadas fechas puede ser normal, pero puede ser extraño si se reciben un día normal.
- **Anomalía colectiva:** Un conjunto de instancias de datos ayudan a encontrar una anomalía. Por ejemplo, alguien está tratando de copiar datos de una máquina remota a un host local de forma inesperada, una anomalía que sería marcada como un potencial ciberataque.

Para empezar con la parte de Detección de anomalías orientada al SIEM, es preciso activar aquellos trabajos de Machine Learning que se consideren oportunos. Para ello, se accede a la sección de Detección de Anomalías del SIEM.

Estos trabajos vienen pre-configurados para detectar ataques como los que se han comentado en el punto anterior de la memoria. Además de esta posibilidad, si se considera necesario se pueden crear trabajos personalizados, como haremos más adelante.

ANOMALY DETECTION SETTINGS

Run any of the Machine Learning jobs below to view anomalous events throughout the SIEM application. We've provided a few common detection jobs to get you started. If you wish to add your own custom jobs, simply create and tag them with "SIEM" from the [Machine Learning](#) application for inclusion here.

Groups Elastic jobs Custom jobs

Showing: 23 jobs

| Job name | Groups | Run job |
|---|---------------------------|-------------------------------------|
| linux_anomalous_user_name_ecs SIEM Auditbeat: Rare and unusual users that are not normally active may indicate unauthorized changes or activity by an unauthorized user which may be credentialled access or lateral movement (beta) | auditbeat process siem | <input type="checkbox"/> |
| packetbeat_dns_tunneling SIEM Packetbeat: Looks for unusual DNS activity that could indicate command-and-control or data exfiltration activity (beta) | packetbeat siem web | <input checked="" type="checkbox"/> |
| packetbeat_rare_dns_question SIEM Packetbeat: Looks for unusual DNS activity that could indicate command-and-control activity (beta) | packetbeat siem web | <input checked="" type="checkbox"/> |
| packetbeat_rare_server_domain SIEM Packetbeat: Looks for unusual HTTP or TLS destination domain activity that could indicate execution, persistence, command-and-control or data exfiltration activity (beta) | packetbeat siem web | <input checked="" type="checkbox"/> |
| packetbeat_rare_urls SIEM Packetbeat: Looks for unusual web browsing URL activity that could indicate execution, persistence, command-and-control or data exfiltration activity (beta) | packetbeat siem web | <input checked="" type="checkbox"/> |

< 1 2 3 4 5 >

Anomaly detection

Show dates

reat.tactic.name

event.module

Ilustración 42: Activación de trabajos desde el SIEM

Para ver las *Anomalías* detectadas por los trabajos de Machine Learning desde el SIEM, hay que acceder al área *Hosts*. En esta área, una vez activadas las características de Machine Learning, aparecerá un nuevo widget de anomalías que incluye aquellas que se vayan detectando.

Aparte de las anomalías de datos en el SIEM, Kibana proporciona una sección enfocada exclusivamente al Machine Learning y la analítica de datos. Estas características se pueden ver desde el menú de Kibana.

Antes de entrar de lleno en la detección de anomalías, y con el objetivo de obtener los mejores resultados de la analítica, es necesario comprender los datos que disponemos, hacer un breve estudio. Es preciso conocer los tipos de datos y el rango y distribución de valores. Para ello, se utilizará el Visualizador de datos, que permite explorar los campos de los datos:

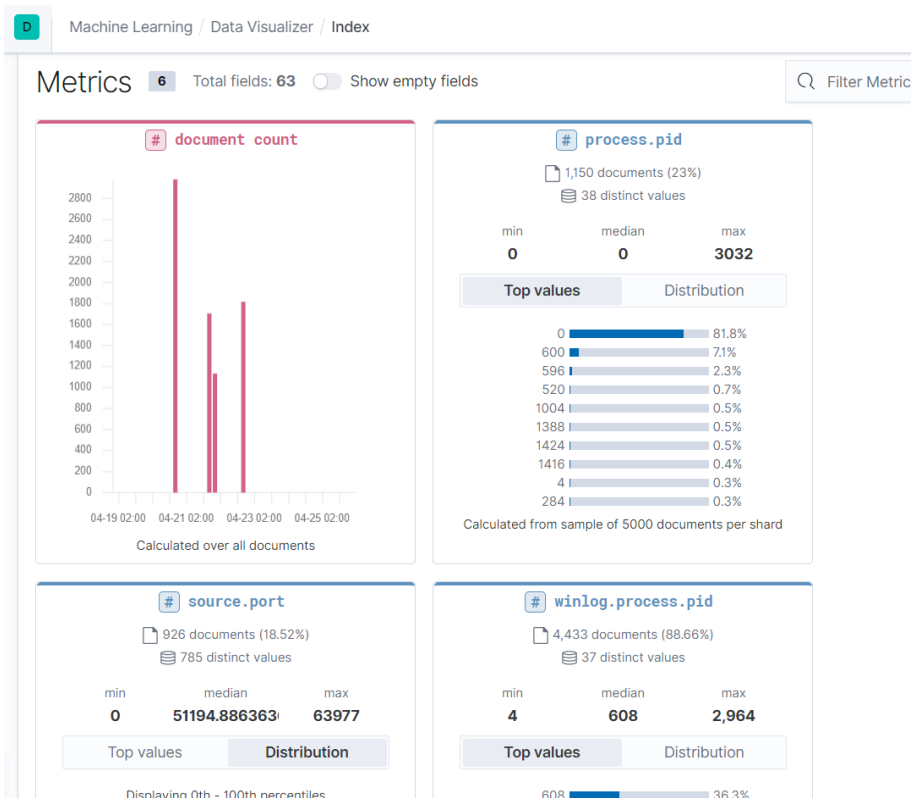


Ilustración 43: Machine Learning Visualizador de Datos

La primera sección contiene los tipos de datos numéricos (métricas). La segunda sección contiene tipos de datos no numéricos (como palabras clave, texto, fecha, boolean, ip).

Para cada métrica, el Visualizador de datos indica cuántos documentos contienen el campo en el período de tiempo seleccionado. También proporciona información sobre los valores mínimos, medianos y máximos, el número de valores distintos y su distribución. Se puede usar el gráfico de distribución para tener una mejor idea de cómo se agrupan los valores en los datos. Alternativamente, se puede ver el top de valores para los campos de métricas.

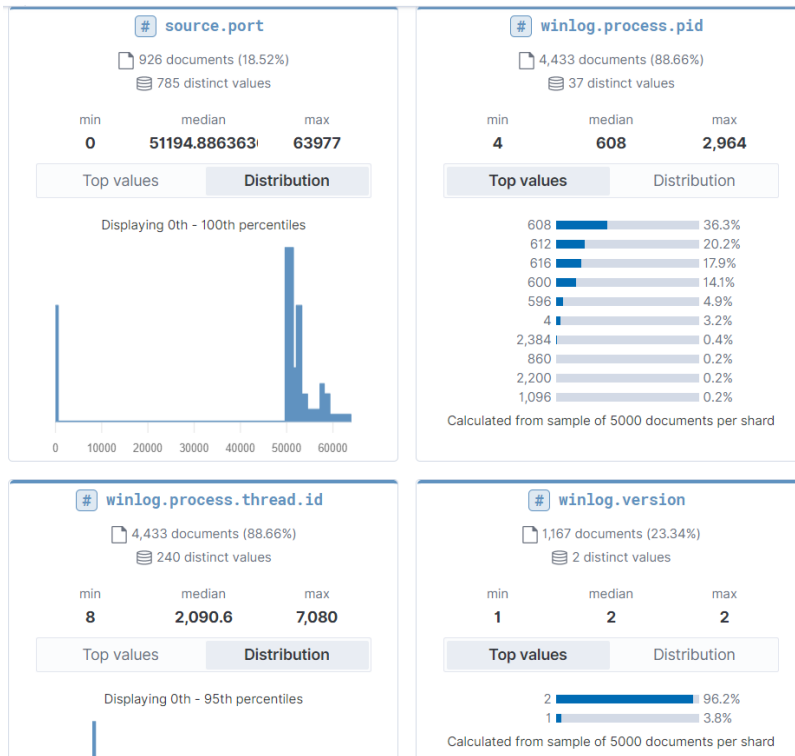


Ilustración 44: Ejemplo de métricas Top y Distribution

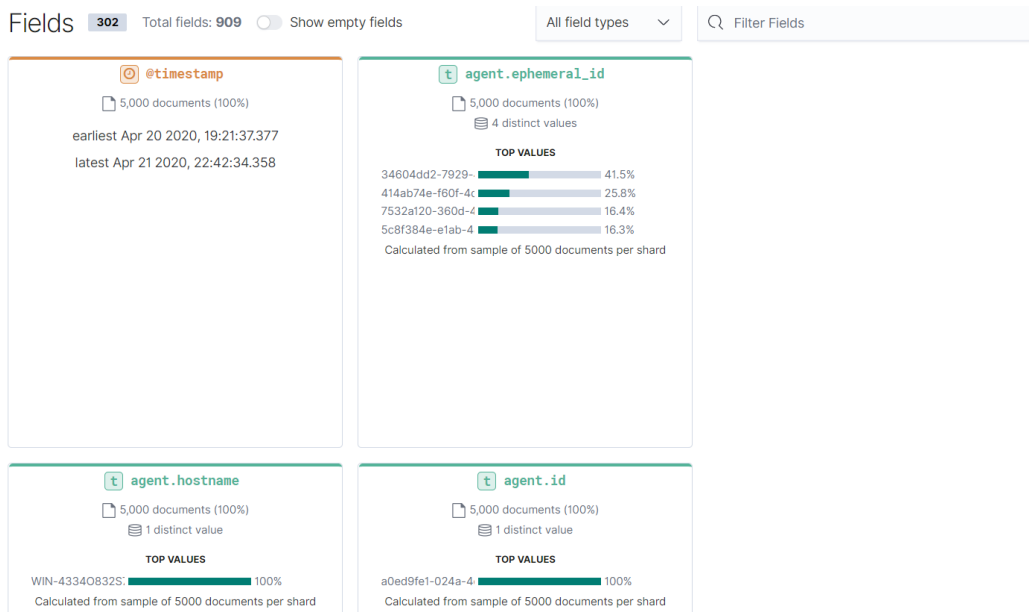


Ilustración 45: Ejemplo de campos

Una vez familiarizado con los datos, se puede comenzar con los trabajos de detección de anomalías.

Para hacer un análisis de la detección de anomalías, hay que dirigirse al menú de Detección de Anomalías en desde el área de Machine Learning de Kibana.

| ID | Description | Processed records | Memory status | Job state | Datafeed state | Latest timestamp | Actions |
|-------------------------------|---|-------------------|---------------|-----------|----------------|---------------------|---------|
| packetbeat_dns_tunneling | SIEM Packetbeat: Looks for unusual DNS activity that could indicate command-and-control or data exfiltration activity (beta) | 134 | ok | opened | started | 2020-04-21 23:43:22 | 🔍 📄 ⋮ |
| packetbeat_rare_dns_question | SIEM Packetbeat: Looks for unusual DNS activity that could indicate command-and-control activity (beta) | 134 | ok | opened | started | 2020-04-21 23:43:22 | 🔍 📄 ⋮ |
| packetbeat_rare_server_domain | SIEM Packetbeat: Looks for unusual HTTP or TLS destination domain activity that could indicate execution, persistence, command-and-control or data exfiltration activity (beta) | 0 | ok | opened | stopped | | 🔍 📄 ⋮ |

Ilustración 46: Vista de Detección de Anomalías

Aquí se muestran las configuraciones de los *trabajos* y los *suministros de datos* (*datafeeds*).

Un suministro de datos recupera datos de series temporales de Elasticsearch y se los proporciona a un trabajo de detección de anomalías para su análisis.

Las funciones de Machine Learning, utilizan el concepto de un *cubo* para dividir las series de tiempo en lotes para su procesamiento.

El intervalo del cubo es parte de la información de configuración de un trabajo de detección de anomalías. Define el intervalo de tiempo que se utiliza para resumir y modelar los datos, lo que suele abarcar entre 5 minutos y 1 hora dependiendo de las características de los datos.

La duración del cubo tiene dos propósitos. El primero, dictar el período de tiempo para buscar características anómalas en los datos, y el segundo determinar la rapidez con la que se pueden detectar las anomalías.

La elección de un cubo más corto permite detectar las anomalías más rápidamente. Sin embargo, existe el riesgo de ser demasiado sensible a las variaciones naturales o al ruido en los datos de entrada. Elegir un intervalo demasiado largo puede significar que las anomalías interesantes se promedien.

Cada trabajo de detección de anomalías contiene uno o más detectores, que definen el tipo de análisis que ocurre (por ejemplo, funciones analíticas máximas, promedio o raras) y los campos que se analizan.

Una vez que se inician los suministros de datos y los trabajos de detección de anomalías han procesado algunos datos, se pueden ver los resultados en Kibana.

Para procesar estos resultados, las funciones de Machine Learning analizan el flujo de entrada de datos, modelan su comportamiento y realizan análisis basados en los detectores de cada *trabajo*. Cuando un evento ocurre fuera del modelo establecido, ese evento es identificado como una anomalía.

Para probar las características de Machine Learning de Elastic enfocadas a la detección de anomalías, se utilizará un fragmento del log de Cisco Umbrella. El objetivo de esta primera prueba será identificar las anomalías referentes a consultas DNS que tengan relación con Malware.

En esta prueba, se creará un trabajo de Machine Learning configurado para analizar un fragmento de dataset que incluye todas las peticiones DNS relacionadas con el Malware detectadas por Cisco Umbrella.

El primer paso antes de empezar con el Machine Learning es obtener el muestreo de los datos relacionados con el Malware. Para ello, se realiza una consulta sobre el log de Cisco Umbrella y se guardan los resultados. La consulta que determina si el tráfico DNS ha sido categorizado como Malware es la siguiente:

```
dns.answers.type:"Malware"
```

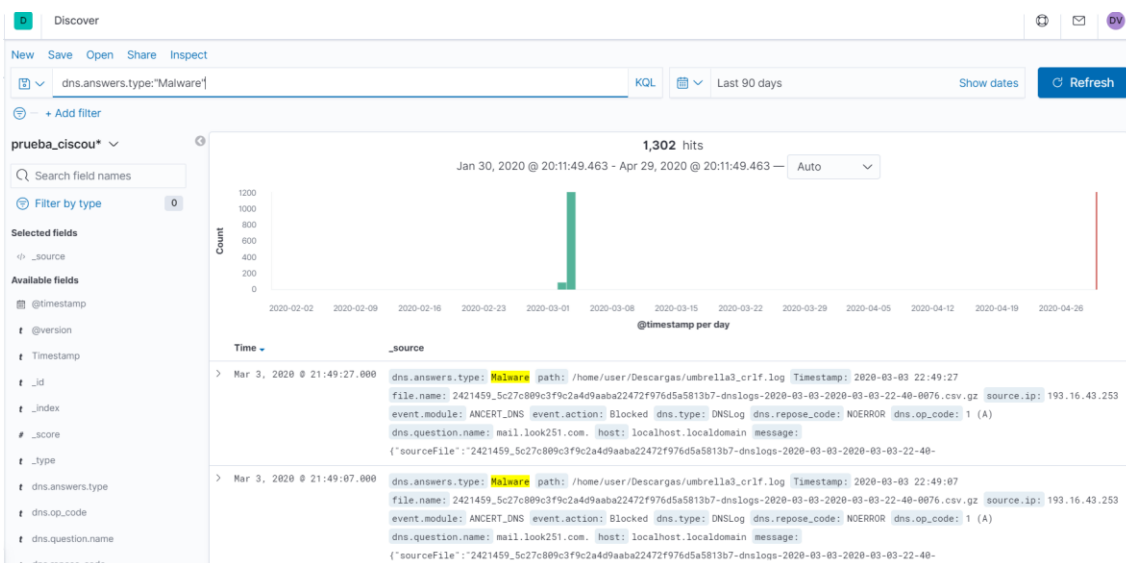


Ilustración 47: Dataset Malware Cisco Umbrella

Desde el panel, se pincha en Save para guardar los resultados.

Una vez realizado, el siguiente paso es crear un trabajo de Machine Learning que analice la muestra. Para ello, hay que acceder al área de Job Management, que se puede encontrar en **Machine Learning → Anomaly Detection → Job Management**.

Machine Learning / Anomaly Detection / Job Management

Job Management Anomaly Explorer Single Metric Viewer Settings

Anomaly detection jobs

Active ML Nodes: 0 Total jobs: 11 Open jobs: 0 Closed jobs: 11 Active datafeeds: 0 Refresh Create new job

Search...

Opened Closed Failed Started Stopped Group

| ID | Description | Processed records | Memory status | Job state | Datafeed state | Latest timestamp | Actions |
|------|-------------------------------------|-------------------|---------------|-----------|----------------|---------------------|---------|
| 0001 | | 95 | ok | closed | stopped | 2020-03-03 22:44:59 | |
| 0002 | | 95 | ok | closed | stopped | 2020-03-03 22:44:59 | |
| 0003 | | 826,860 | ok | closed | stopped | 2020-03-03 22:49:59 | |
| 0004 | | 826,860 | ok | closed | stopped | 2020-03-03 22:49:59 | |
| 0006 | | 574 | ok | closed | stopped | 2020-02-07 23:29:59 | |
| 0007 | Detecta anomalías en Cisco Umbrella | 826,860 | ok | closed | stopped | 2020-03-03 22:49:59 | |
| 0008 | | 826,860 | ok | closed | stopped | 2020-03-03 22:49:59 | |
| 0009 | | 286 | ok | closed | stopped | 2020-03-03 22:44:59 | |
| 0010 | | 143 | ok | closed | stopped | 2020-03-03 22:39:59 | |

Ilustración 48: Job Management - Machine Learning

Desde esta vista, se accede a crear un nuevo trabajo. Cabe destacar que también se pueden crear trabajos personalizados desde la sección de *Anomaly Detection* del SIEM.

Los pasos a seguir para la creación del trabajo son los siguientes:

- **Seleccionar el índice o consulta:** En este punto se selecciona el dataset con el que se va a trabajar. En este caso seleccionaremos la consulta que se ha guardado anteriormente.
- **Seleccionar el tipo de trabajo:** En este punto hay que elegir el tipo de trabajo que se desea crear. Existen trabajos *Single Metric*, *Multi Metric*, *Population*, *Advanced* y *Categorization*.

Create a job from the saved search Malware_DNS

Use a wizard

Use one of the wizards to create a machine learning job to find anomalies in your data.

Single metric
Detect anomalies in a single time series.

Multi metric
Detect anomalies in multiple metrics by splitting a time series by a categorical field.

Population
Detect activity that is unusual compared to the behavior of the population.

Advanced
Use the full range of options to create a job for more advanced use cases.

Categorization
Group log messages into categories and detect anomalies within them.

Learn more about your data

If you're not sure what type of job to create, first explore the fields and metrics in your data.

Data Visualizer
Learn more about the characteristics of your data.

Ilustración 49: Tipos de trabajo - Machine Learning

Cada tipo de trabajo tiene sus características, tal y como definen la imagen anterior. Un trabajo *Single Metric* contiene un solo detector. Un detector define el tipo de análisis que se producirá (por ejemplo, funciones analíticas máximas, medias o raras) y los campos que se analizarán. Para este caso en particular se utilizará el tipo *Multi Metric*, ya que vamos a trabajar sobre diversas métricas.

- **Seleccionar el rango de tiempo:** Se selecciona el rango de tiempo sobre el que deseamos trabajar. En este caso, se cogerá el rango completo.

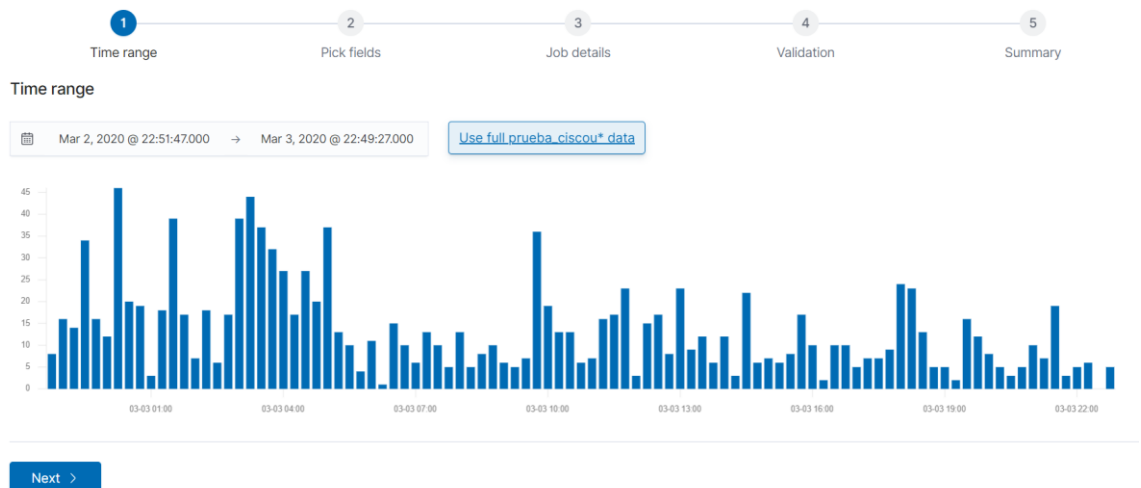


Ilustración 50: Rango de tiempo del dataset - Machine Learning

- **Selección de campos:** Se seleccionarán los campos sobre los que se trabajarán. Se utilizará la función de *Count (Event rate)* para rastrear la tasa de ocurrencia de este tipo de mensajes a lo largo del tiempo. Y emplea el campo *dns.answers.type* como el "campo dividido". El Bucket span se estima mediante el botón *Estimate bucket span*.

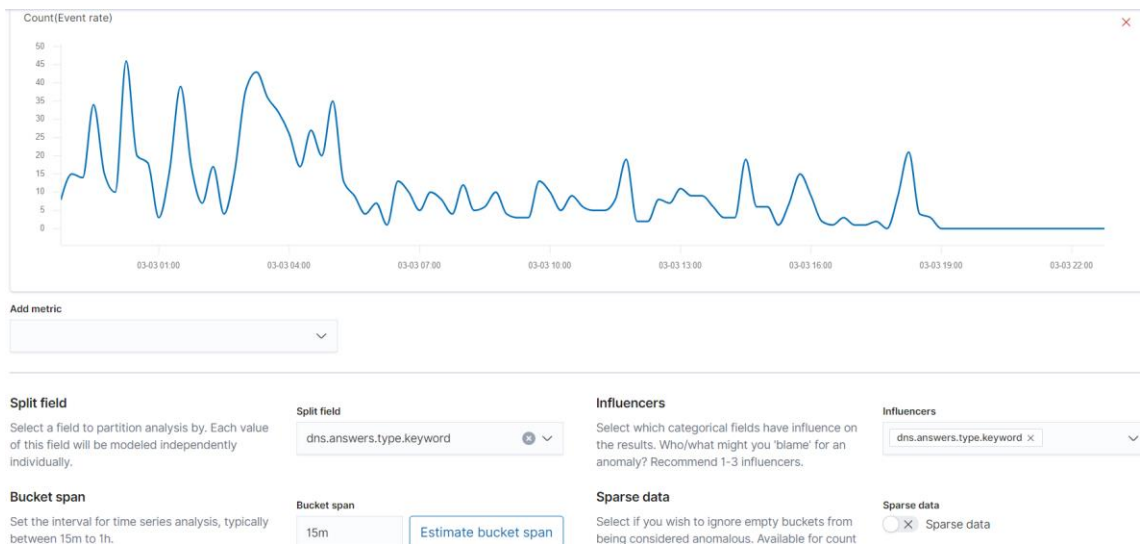


Ilustración 51: Definición del trabajo - Machine Learning

- **Detalles del trabajo:** El siguiente paso incluye los detalles del trabajo, en el que el usuario debe introducir el ID que desea para el trabajo, introducir una descripción, etc.
- **Validación del trabajo:** El Wizard de creación de trabajos valida que la configuración realizada sea correcta.
- **Resumen del trabajo:** Es el paso final, incluye un resumen del trabajo.

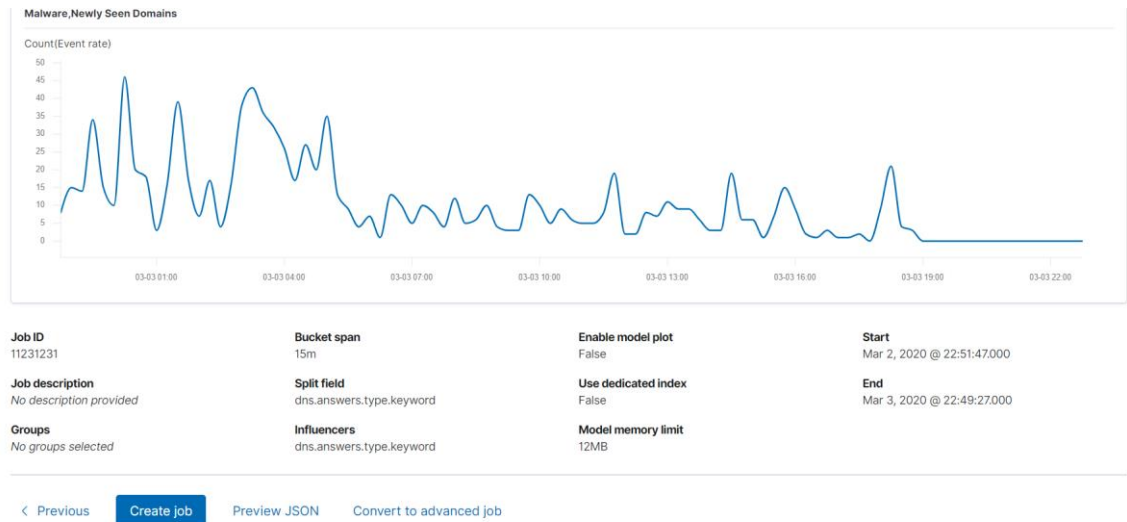


Ilustración 52: Resumen del trabajo

Una vez creado el trabajo, se pueden examinar los resultados del análisis. Kibana proporciona dos herramientas para examinar los resultados de los trabajos de detección de anomalías, por un lado el *Anomaly Explorer* y por otro el *Single Metric Viewer*. A continuación se verán las características de cada uno:

Anomaly Explorer

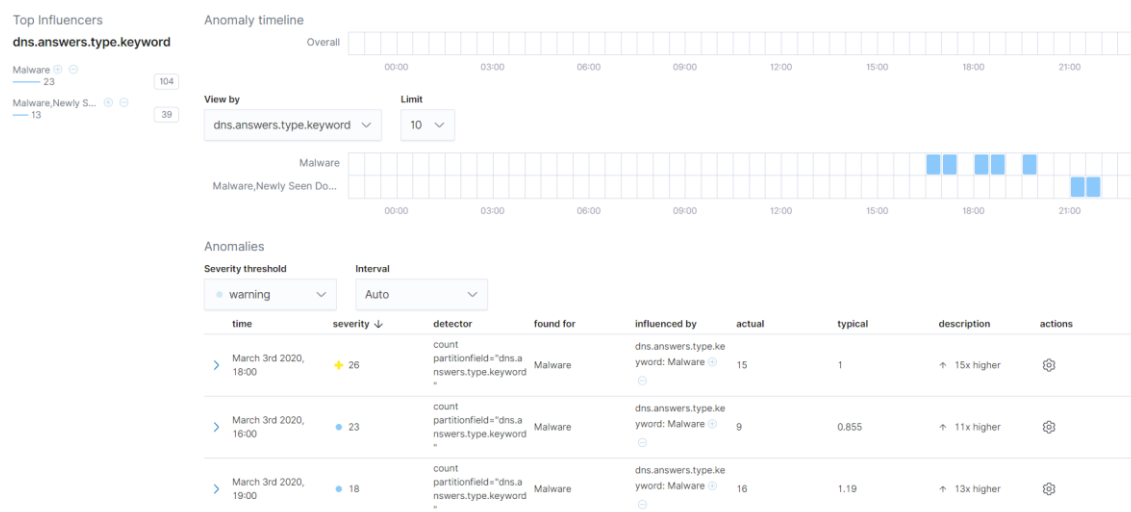


Ilustración 53: Explorador de Anomalías - Machine Learning

En este escenario, el análisis está segmentado de tal manera que se tienen líneas de base completamente diferentes para cada valor distinto del campo de partición. Al observar los patrones temporales sobre una base de cada entidad, se podrían detectar cosas que de otra manera podrían estar ocultas en la vista global.

En el lado izquierdo del Explorador de Anomalías, hay una lista de los principales *influencers* para todas las anomalías detectadas en ese mismo período de tiempo. La lista incluye las puntuaciones máximas de las anomalías, que en este caso se agregan para cada *influencer*, para cada cubo, en todos los detectores. También hay una suma total de las puntuaciones de las anomalías para cada *influencer*. Se utiliza esta lista para ayudar a reducir los factores contribuyentes y enfocarse en las entidades más anómalas.

Se pueden ver las horas exactas en las que se produjeron las anomalías. Si hay múltiples detectores o métricas en el trabajo, se puede ver cuál captó la anomalía.

Debajo de los gráficos, hay una tabla que proporciona más información, como los valores típicos y reales y los *influencers* que contribuyeron a la anomalía.

| time | severity ↓ | detector | found for | influenced by | actual | typical | description | actions |
|---|------------|--|-----------|-----------------------------------|--------|---------|--------------|---------|
| March 3rd 2020, 18:00 | + 26 | count partitionfield="dns.answers.type.keyword" | Malware | dns.answers.type.keyword: Malware | 15 | 1 | ↑ 15x higher | |
| <p>Description minor anomaly in count partitionfield="dns.answers.type.keyword" found for dns.answers.type.keyword Malware</p> <p>Details on highest severity anomaly</p> <p>dns.answers.type.keyword Malware</p> <p>time March 3rd 2020, 18:00:00 to March 3rd 2020, 18:15:00</p> <p>function count</p> <p>actual 15</p> <p>typical 1</p> <p>job ID 11231231</p> <p>multi-bucket impact high</p> <p>probability 0.012497176565399538</p> <p>Influencers</p> <p>dns.answers.type.keyword Malware</p> | | | | | | | | |
| March 3rd 2020, 16:00 | ● 23 | count partitionfield="dns.answers.type.keyword" | Malware | dns.answers.type.keyword: Malware | 9 | 0.855 | ↑ 11x higher | |

Ilustración 54: Tabla Explorador de Anomalías

Las puntuaciones de las anomalías que se ven en cada sección del Explorador de Anomalías pueden diferir ligeramente. Esta disparidad ocurre porque para cada trabajo hay resultados de cubo, resultados de *influencer* y resultados de registro. Las puntuaciones de anomalías se generan para cada tipo de resultado. La línea de tiempo de anomalías utiliza las puntuaciones de anomalías a nivel de cubo.

La lista de los principales *influencers* utiliza las puntuaciones de las anomalías a nivel de *influencer*.

La lista de anomalías utiliza las puntuaciones de las anomalías a nivel de registro.

Single Metric Viewer

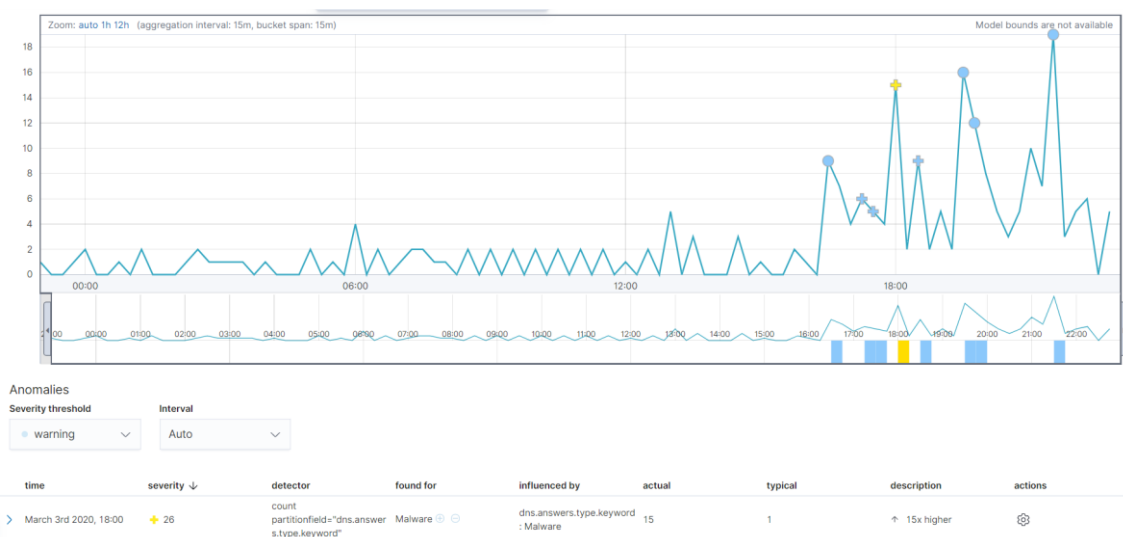


Ilustración 55: Single Metric Viewer - Machine Learning

Esta vista contiene un gráfico que representa los valores reales y previstos a lo largo del tiempo. Sólo puede mostrar una única serie temporal.

La línea azul del gráfico representa los valores de datos reales. El área entre los límites superior e inferior son los valores más probables para el modelo. Si un valor está fuera de esta área, entonces se puede decir que es anómalo. Si apareciese un área sombreada en azul, esta representaría los límites de los valores esperados.

Si analizamos el selector de tiempo desde el principio hasta el final de los datos, se puede ver cómo el modelo mejora a medida que procesa más datos. Al principio, el rango esperado de valores es bastante amplio y el modelo no está captando la periodicidad de los datos. Pero aprende rápidamente y comienza a reflejar los patrones en sus datos.

Cualquier punto de datos fuera del rango que fue predicho por el modelo se marca como anomalía. A fin de proporcionar una visión sensata de los resultados, se calcula una puntuación de anomalía para cada intervalo de tiempo del cubo. La puntuación de la anomalía es un valor de 0 a 100, que indica la importancia de la anomalía en comparación con las anomalías observadas anteriormente. Los valores altamente anómalos se muestran en rojo y los valores de puntuación baja se indican en azul. Un intervalo con una puntuación alta de anomalía es significativo y requiere investigación.

Para cada anomalía, se pueden ver detalles clave como el tiempo, los valores reales y esperados ("típicos"), y su probabilidad en la sección de Anomalías del visor, de la misma forma que en la tabla de la **ilustración 54**.

Por defecto, la tabla contiene todas las anomalías que tienen una severidad de "warning" o superior en la sección seleccionada de la línea de tiempo. Si por ejemplo sólo interesan las anomalías críticas, se puede cambiar el umbral de gravedad de esta tabla.

Para finalizar, una vez identificadas las anomalías, el siguiente paso debería ser el de tratar de determinar el contexto de las situaciones vistas. Sería bueno hacerse preguntas, como por ejemplo:

- ¿hay otros factores que contribuyan al problema?
- ¿Se limitan las anomalías a determinados servidores o equipos?
- ¿Se limitan las anomalías a determinadas horas/días?

Analíticas adicionales

Para complementar el resultado de la prueba anterior de análisis de Machine Learning, se han creado dos trabajos más de Machine Learning que analizan los datasets completos de Cisco Umbrella y de Fortinet, con el objetivo de visualizar de forma breve las anomalías resultantes tras el análisis en estos conjuntos de datos.

A continuación, se muestran ilustraciones en base a las herramientas de análisis utilizadas (Anomaly Explorer y Single Metric Viewer):



Ilustración 56: Detección de anomalías - Cisco Umbrella

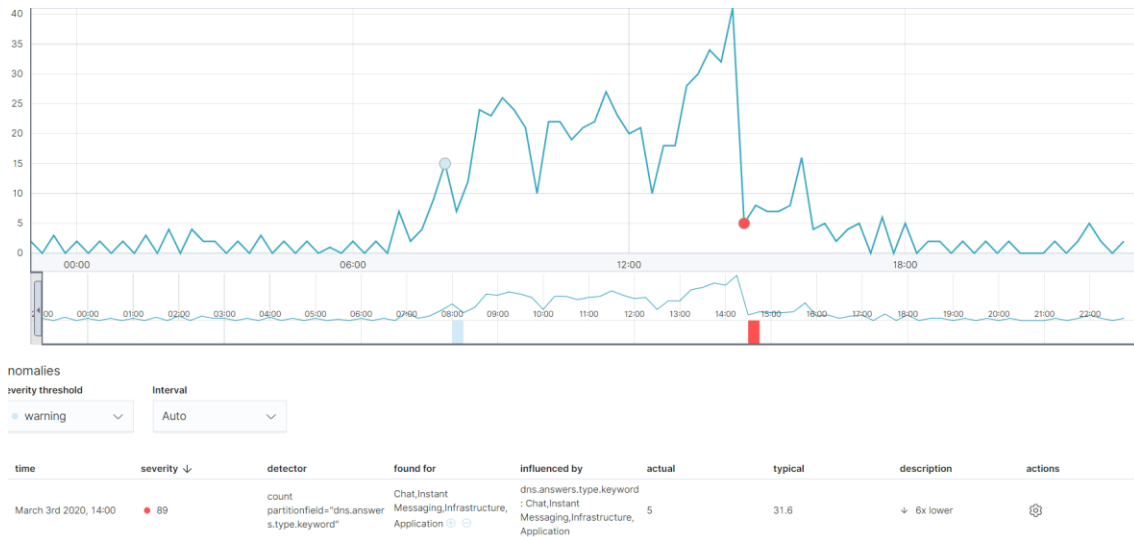


Ilustración 57: Single Metric Viewer - Cisco Umbrella

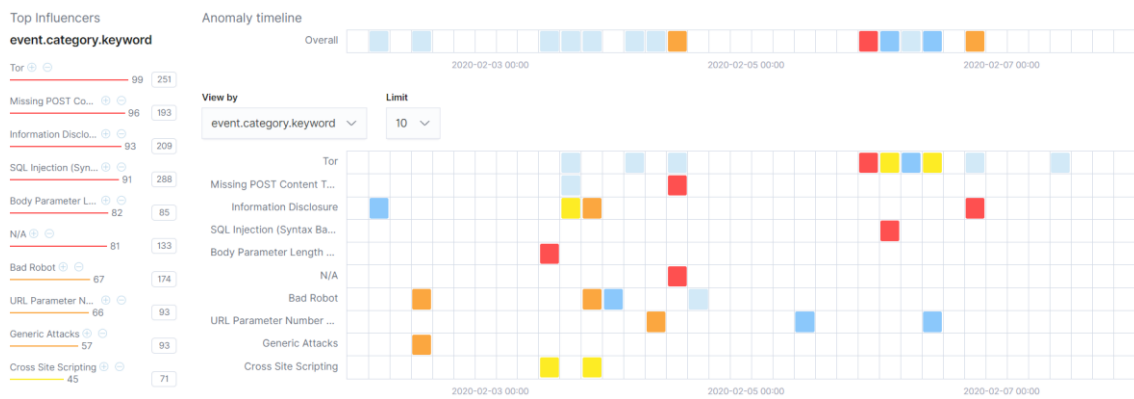


Ilustración 58: Detección de Anomalías – Fortiweb

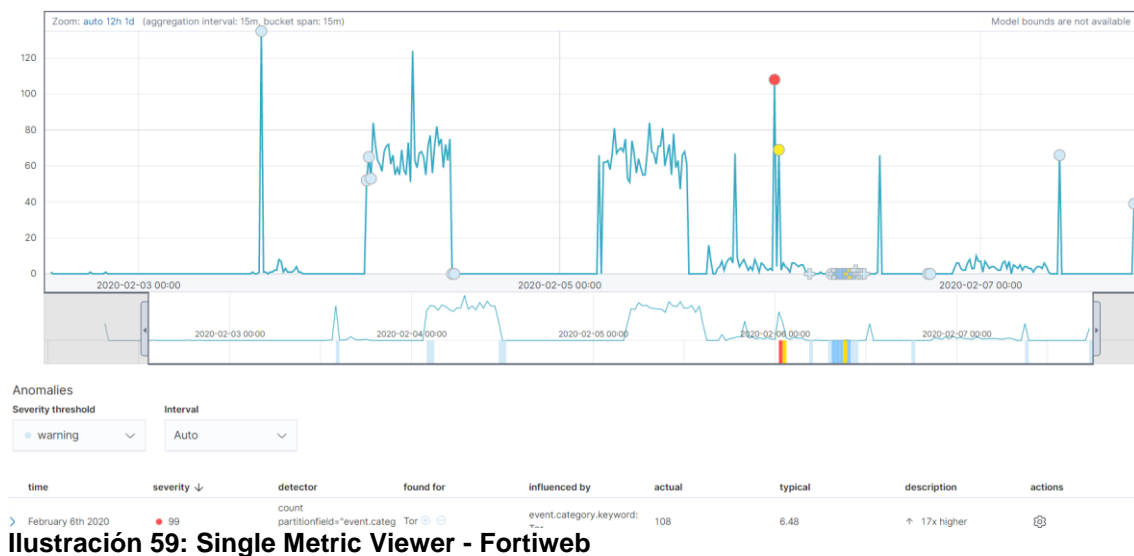


Ilustración 59: Single Metric Viewer - Fortiweb

Por un lado, en los resultados del análisis del dataset de Cisco Umbrella (Ilustración 56), se puede ver que se detectan anomalías en las resoluciones DNS relacionadas con:

- Viajes.
- Chats o Aplicaciones de mensajería instantánea.
- Noticias/Deportes.

Por otro lado, en los resultados del análisis del dataset de Fortiweb (Ilustración 58), se pueden ver las anomalías relacionadas con las amenazas de la web. Las anomalías de seguridad encontradas más relevantes son:

- Tráfico Tor.
- Content Type POST ausente.
- Divulgación de información.
- Ataques SQL Injection.
- Violación de la longitud del parámetro del cuerpo.

Para concluir la parte de detección de anomalías con Machine Learning, es preciso comentar una serie de cualidades del Machine Learning en Elastic:

- El Machine Learning procesa los datos en orden cronológico, y sólo analiza los datos una vez. Por tanto, si se hace que el Machine Learning aprenda sobre los datos históricos, pero luego se le pide que se ejecute de forma continua (en tiempo real), el algoritmo no vuelve a mirar los datos antiguos, si no que observará los nuevos datos entrantes.
- La detección de anomalías de Elastic no es supervisada, lo que significa que al algoritmo no se le indican los patrones de detección, sino que simplemente se detectan los cambios en el comportamiento de los datos en una cierta dimensión. El algoritmo de Machine Learning encontrará desviaciones en el valor o la tasa de los datos a lo largo del tiempo, detectando así las anomalías.

5.0 Elastic Watcher

5.1 Definición del plugin Watcher

Elastic Watcher es un plugin utilizado para generar alertas en tiempo real. Esta característica de Elastic es de pago, por lo que solo se puede utilizar si se dispone de una licencia de pago.

El principal objetivo de Watcher es crear acciones basadas en condiciones, que se evalúan en base a consultas periódicas a los datos. Si los resultados de las consultas son sospechosos, Watcher generará alertas y notificaciones.

Watcher permite elegir entre muchas opciones de alerta, con integraciones incorporadas para correo electrónico, PagerDuty, Slack y HipChat. También incluye una potente salida de webhook para integrarlo con una infraestructura de monitoreo existente o con cualquier sistema de terceros.

Para determinadas acciones que son más difíciles de definir con reglas y umbrales, es posible combinar las alertas con características de Machine Learning sin supervisión para encontrar cosas inusuales.

Así pues, se trata de una característica sobresaliente para la analítica de seguridad, ya que su enfoque es perfectamente válido para detectar incidentes de seguridad a través del análisis de los datos. Por ejemplo, se puede utilizar para observar los registros de aplicaciones para detectar interrupciones en el funcionamiento, para auditar los logs de acceso para detectar amenazas a la seguridad o para detectar señales de un ataque a través de los logs de red (SIEM), aunque existen múltiples posibilidades.

El funcionamiento de un Watcher se resume en el siguiente diagrama:

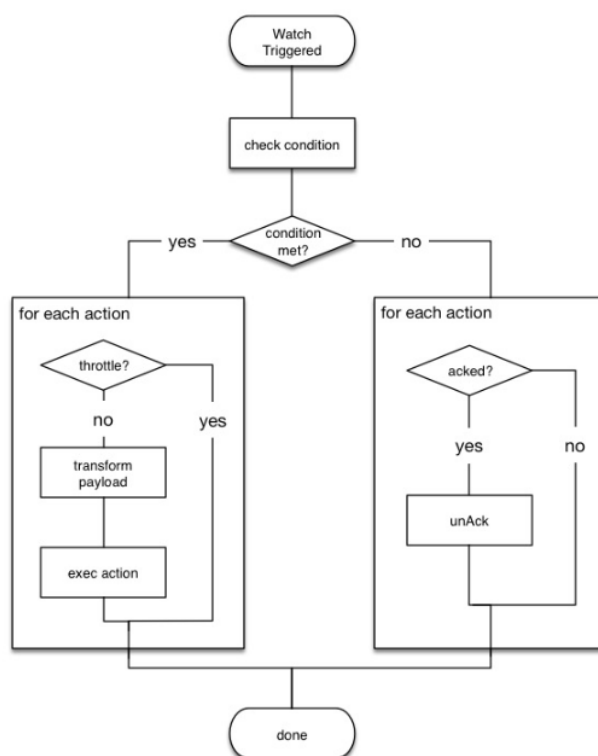


Ilustración 60: Diagrama de flujo Watcher

Durante este proceso, el Watcher hace lo siguiente:

- Carga los datos de entrada como carga útil en el contexto de ejecución de la vigilancia. Esto hace que los datos estén disponibles para todos los pasos posteriores del proceso de ejecución. Este paso es controlado por la entrada del Watcher.
- Evalúa la condición del Watcher para determinar si se debe continuar procesando el Watcher o no. Si se cumple la condición (se evalúa a verdadero), el procesamiento avanza al siguiente paso. Si no se cumple, es decir, se evalúa como falso, la ejecución del Watcher se detiene.
- Aplica transformaciones a la carga útil del Watcher, si es necesario.

- Ejecuta las acciones del Watcher siempre que se cumpla la condición y el Watcher no esté “estrangulado”. Durante la ejecución del Watcher, una vez que se cumple la condición, se decide por cada acción configurada si debe ser estrangulado. El propósito principal de la estrangulación de la acción es evitar demasiadas ejecuciones de la misma acción para el mismo Watcher.

5.2 Configuración y uso de Watcher.

Para empezar a utilizar Watcher (Elastic, Watcher | Kibana Guide 7.6.0, 2020), hay que dirigirse a **Management** → **Elasticsearch** → **Watcher** desde Kibana.

Nota: Si las características de seguridad de Elasticsearch están habilitadas, es necesario tener los privilegios `manage_watcher` o `monitor_watcher` para usar Watcher en Kibana.

Desde aquí, se pueden realizar las siguientes acciones:

- Crear un Watcher simple de umbral.
- Ver el historial de un Watcher y el estado de las acciones.
- Desactivar y borrar un Watcher.
- Crear un Watcher avanzado usando la sintaxis de la API (JSON).

Una alerta de umbral es el Watcher más común que se suele crear, y también el más simple. Esta alerta comprueba periódicamente cuando los datos están por encima, por debajo, son iguales o están entre un cierto umbral dentro de un intervalo de tiempo determinado.

A continuación se muestra el proceso de creación de un Watcher de umbral para generar una alerta. Esta alerta se activará cuando el uso total máximo de CPU en una máquina supere un determinado porcentaje. Se utilizará Metricbeat para recoger métricas de los sistemas y servicios.

Los pasos son los siguientes:

- Hay que acceder a Crear y luego seleccionar Crear alerta de umbral (threshold alert).



You don't have any watches yet

Watch for changes or anomalies in your data and take action if needed. [Learn more.](#)

Create ▾

Ilustración 61: Crear Alerta Watcher

- A continuación se introduce el nombre que se desee para la Alerta, los índices que van a utilizar las consultas, el campo, y la programación de la tarea.

Create threshold alert

Send an alert when your specified condition is met. Your watch will run every 1 minute.

Name
Alerta uso de CPU superior al 25%

Indices to query
metricbeat-* x

Time field
@timestamp

Run watch every
1 minute

Use * to broaden your query.

Ilustración 62: Definición de Watcher

- El siguiente paso es añadir una *condición*. La condición evalúa los datos que cargados en el Watcher y determina si se requiere alguna acción. La condición será cuando los valores de CPU estén por encima del umbral del 25%.

Match the following condition

WHEN max() OF system.process.cpu.total.norm.pct OVER all documents IS ABOVE 0.25 FOR THE LAST 5 minutes

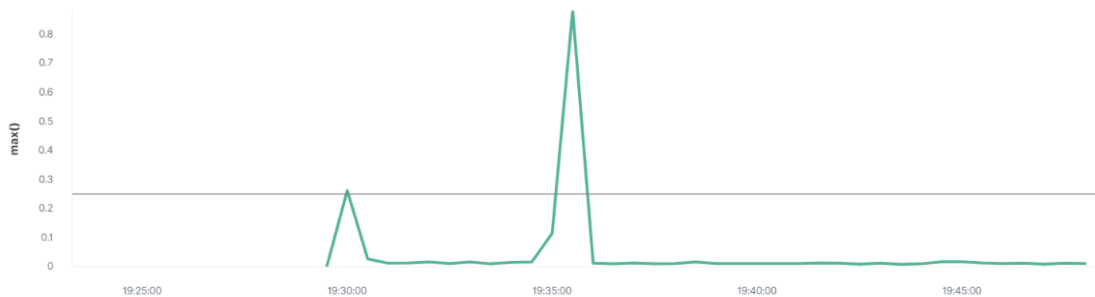


Ilustración 63: Definición de la condición del Watcher

- El último paso es la *acción*. La acción se desencadena cuando se cumple la condición del Watcher. Existen las siguientes formas de notificación de la alerta:

- Email**
Send an email from your server.
- Logging**
Add an item to the logs.
- Slack**
Send a message to a Slack user or channel.
- Webhook**
Send a request to a web service.
- Index**
Index data into Elasticsearch.
- PagerDuty**
Create an event in PagerDuty.
- Jira**
Create an issue in Atlassian's Jira Software.

Ilustración 64: Posibilidades de notificación de una Alerta

Para este ejemplo, se configurara una acción de correo electrónico. Previamente, es necesario establecer la configuración desde un servidor SMTP. En este caso, se utilizará el de Google, y para ello es necesario introducir las siguientes líneas en el fichero *elasticsearch.yml*:

```
xpack.notification.email.account:
  gmail_account:
    profile: gmail
    smtp:
      auth: true
      starttls.enable: true
      host: smtp.gmail.com
      port: 587
      user: <username>
```

Y además, almacenar la contraseña en la keystore de elasticsearch a través del siguiente comando:

```
bin/elasticsearch-keystore add
xpack.notification.email.account.gmail_account.smtp.secure_password
```

Una vez hecho esto, ya se pueden rellenar los campos del formulario para enviar el email.

✕ Email

To email address
davidvazquezp@uoc.edu ✕

Subject (optional)
Watch {{{ctx.metadata.name}}}

Body
Carga de CPU superior al 25%

Ilustración 65: Acción email Watcher

- Finalmente, se crea la alerta.

Una vez creada y activada la alerta (por defecto se activa tras su creación), cuando se dispare la condición programada, se recibirá la notificación por email.



Watcher Alert <noreply@watcheralert.found.io>

to me ▾

Carga de CPU superior al 25%

Ilustración 66: Notificación Watcher

Las alertas pueden tener los siguientes estados:

- **Firing:** El Watcher se dispara y realiza activamente las acciones asociadas.
- **Error:** El Watcher no funciona correctamente.
- **OK:** El Watcher no está disparando activamente pero funciona correctamente.
- **Disabled:** El Watcher no se disparará bajo ninguna circunstancia.

Current status for 'Alerta uso de CPU superior al 25%'

[Deactivate](#)

[Delete](#)

[Execution history](#)

[Action statuses](#)

Last one hour ▾

| Trigger time | State | Comment |
|---------------------------|----------|---------|
| 2020-05-06T22:45:53+02:00 | ✓ OK | |
| 2020-05-06T22:44:53+02:00 | ▶ Firing | |

Rows per page: 10 ▾ [<](#) [1](#) [>](#)

Ilustración 67: Historial de ejecución Watcher CPU

Además de las alertas de umbral, que son los más simples, se pueden crear Watchers avanzados (Elastic, Watcher Avanzado | Elastic Guide 7.6.0, 2020) que incluyan consultas más complejas.

A continuación, se creará y se mostrará el código de un Watcher avanzado que sirve para detectar intentos de login fallidos en el Windows Server 2016, que puedan indicar la presencia de un ataque por fuerza bruta a la contraseña.

La alerta funciona de la siguiente manera: si se detectan más de 5 logins fallidos en un intervalo de 5 minutos se enviará una notificación por email que incluye un mensaje con el usuario afectado y el log de seguridad adjunto.

Para crear este Watcher, hay que acceder a Crear y seleccionar Crear Watcher avanzado. A continuación se debe introducir un nombre y el código JSON del Watcher.

Create advanced watch

[Edit](#) [Simulate](#)

Name (optional)

ID

Watch JSON (API syntax)

```
1 {
2   "trigger": {
3     "schedule": {
4       "interval": "30m"
5     }
6   },
7   "input": {
8     "search": {
9       "request": {
10        "body": {
11          "size": 0,
12          "query": {
13            "match_all": {}
14          }
15        },
16        "indices": [
17          ""
18        ]
19      }
20    }
21  },
22  "condition": {
23    "compare": {
24      "ctx.payload.hits.total": {
25        "gte": 10
26      }
27    }
28  },
29  "actions": {
30    "my-logging-action": {
31      "logging": {
32        "text": "There are {{ctx.payload.hits.total}} documents in your index. Threshold is 10."
33      }
34    }
35  }
36 }
```

Ilustración 68: Vista Watcher avanzado

Un Watcher tiene la siguiente estructura:

- **Trigger:** Determina cuándo se comprueba el Watcher. Un Watcher debe tener un disparador.
- **Input:** Carga los datos en la carga útil del Watcher. Si no se especifica ninguna entrada, se carga una carga útil vacía.
- **Condition:** Controla si las acciones de vigilancia se ejecutan. Si no se especifica ninguna condición, la condición es por defecto siempre.
- **Transform:** Procesa la carga útil del Watcher para prepararla para las acciones. Se pueden definir transformaciones a nivel de la vigilancia o definir transformaciones específicas de la acción. Esto es opcional.
- **Actions:** Especifica lo que sucede cuando se cumple la condición del Watcher.

Para el caso práctico comentado, el JSON que se debe introducir tendrá la siguiente forma:

```

{
  "trigger": {
    "schedule": {
      "interval": "5m"
    }
  },
  "input": {
    "search": {
      "request": {
        "search_type": "query_then_fetch",
        "indices": [
          "winlogbeat-*"
        ],
        "rest_total_hits_as_int": true,
        "body": {
          "query": {
            "bool": {
              "filter": [
                {
                  "range": {
                    "@timestamp": {
                      "from": "now-5m",
                      "to": "now"
                    }
                  }
                },
                {
                  "match": {
                    "winlog.event_id": "4625"
                  }
                }
              ]
            }
          }
        },
        "aggs": {
          "users": {
            "terms": {
              "field": "winlog.event_data.TargetUserName.keyword",
              "size": 10
            }
          }
        }
      }
    },
    "condition": {
      "compare": {
        "ctx.payload.hits.total": {
          "gte": 5
        }
      }
    },
    "actions": {
      "my-logging-action": {
        "logging": {
          "level": "info",
          "text": "Watcher Notification Encountered {{ctx.payload.hits.total}} failed logon from user {{#ctx.payload.hits.hits.0}}[_source.winlog.event_data.TargetUserName]}{/ctx.payload.hits.hits.0}."
        },
        "email_administrator": {
          "email": {
            "profile": "standard",
            "attachments": {
              "attached_data": {
                "data": {
                  "format": "json"
                }
              }
            },
            "priority": "high",
            "to": [
              "davidvazquezp@uoc.edu"
            ],
            "subject": "Encountered {{ctx.payload.hits.total.value}} errors",
            "body": {
              "text": "Encountered {{ctx.payload.hits.total.value}} failed logon from user {{#ctx.payload.hits.hits.0}}[_source.winlog.event_data.TargetUserName]}{/ctx.payload.hits.hits.0}."
            }
          }
        }
      }
    }
  }
}

```

Una vez creado el Watcher, se realizará una prueba para verificar su funcionamiento. Tras realizar varios intentos fallidos de conexión consecutivos, se recibe la siguiente notificación, que verifica el correcto funcionamiento del Watcher:

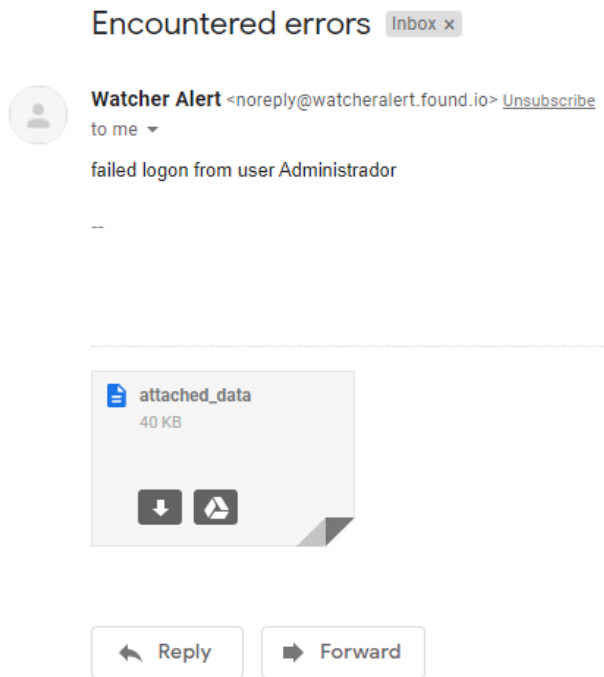


Ilustración 69: Notificación fuerza bruta Watcher

Current status for 'Alerta intento de login fallido'

[Execution history](#) [Action statuses](#)

Last one hour ▾

| Trigger time | State | Comment |
|---------------------------|----------|---------|
| 2020-05-06T23:52:35+02:00 | ✓ OK | |
| 2020-05-06T23:47:35+02:00 | ▶ Firing | |
| 2020-05-06T23:42:35+02:00 | ✓ OK | |
| 2020-05-06T23:37:35+02:00 | ✓ OK | |
| 2020-05-06T23:32:35+02:00 | ✓ OK | |
| 2020-05-06T23:27:35+02:00 | ✓ OK | |
| 2020-05-06T23:22:35+02:00 | ✓ OK | |
| 2020-05-06T23:17:35+02:00 | ✓ OK | |
| 2020-05-06T23:12:35+02:00 | ✓ OK | |
| 2020-05-06T23:07:35+02:00 | ✓ OK | |

Rows per page: 10 ▾

Ilustración 70: Historial de ejecución Watcher Fuerza Bruta

6.0 Hardening de Elastic

En este punto de la memoria se va a explicar el proceso de Hardening de la plataforma Elastic Stack (Elastic, Elastic Guide 7.1 | Hardening, 2020). En una instalación por defecto, la seguridad brilla por su ausencia, ya que la plataforma no implementa ninguna medida de seguridad que proporcione confidencialidad, integridad de nuestros datos y la disponibilidad de la plataforma, ya que cualquier usuario con acceso a la red podría acceder al control de Kibana. Es por eso que es preciso realizar unas ciertas mejoras para evitar brechas de seguridad.

Los problemas de seguridad más comunes en instalaciones por defecto son las siguientes:

- Ausencia de mecanismos de autenticación o autorización en el acceso a Elasticsearch y a la API.
- Tráfico sin cifrar mediante HTTP.
- Denegación de servicio sobre nodos del clúster.
- Fugas de información de sensible.
- Operaciones de administración de la instancia de Elasticsearch sin protecciones.

El proceso de hardening de la plataforma Elastic se basa principalmente en tres aspectos:

- Uso de protocolo SSL/TLS para comunicaciones encriptadas.
- Disposición de un Realm nativo y de archivos para crear y administrar usuarios.
- Control de acceso basado en roles para controlar el acceso de usuarios a los índices y las API de cluster, que también permite multitenancy para Kibana con seguridad para los Kibana Spaces.

Con estos tres puntos vamos a cubrir una gran parte de los problemas de seguridad que tiene una instalación por defecto.

Es el primer paso para proteger los datos que fluyen a través de Elasticsearch, Kibana, Beats y Logstash de usuarios no autorizados y modificaciones no intencionales es disponer de un sistema de autenticación ([Creación de usuarios y Gestión de Roles en Elastic](#)). Elastic implementa diversos tipos de integración con los sistemas de gestión de identidad más habituales, como Active Directory, LDAP o el realm nativo de Elasticsearch. También se pueden implementar opciones de inicio de sesión único (SSO) como certificados, Kerberos y SAML, o desarrollar un realm personalizado.

Como segundo paso, es importante la implementación del modelo RBAC (Gestión de Acceso Basado en Roles). Esto permite tener un control de quién puede hacer qué dentro del Elastic Stack mediante la asignación de roles y privilegios de usuario.

De esta forma, se controla qué usuarios y roles pueden acceder a cada espacio, incluidas las características específicas de Kibana y las apps dentro de ellas.

Finalmente, el tercer punto es el cifrado de las comunicaciones ([Configurar SSL/TLS](#)). Con cifrado SSL/TLS, se implementa la seguridad de nodo a nodo, empleando HTTPS como protocolo para trasportar el tráfico del cliente a lo largo de Elastic Stack.

Junto a esto, es importante destacar la característica de filtrado mediante listas blancas. El filtrado IP previene que usuarios no autorizados se unan a un clúster o se comuniquen con él, consiguiendo tráfico confiable de “nodo a nodo”.

Algunas de las características de seguridad requieren de una licencia de pago para poder ser utilizadas. Cada tipo de licencia incluye más o menos características de seguridad, las cuales se pueden consultar en la siguiente ilustración:

| | OPEN SOURCE | BASIC | ORO | PLATINO | ENTERPRISE |
|---|-------------|-------|-----|---------|------------|
| OPERACIONES Y GESTIÓN DEL ELASTIC STACK | | | | | |
| Seguridad del Elastic Stack | | | | | |
| Configuración segura | ✓ | ✓ | ✓ | ✓ | ✓ |
| Comunicaciones encriptadas | — | ✓ | ✓ | ✓ | ✓ |
| Control de acceso basado en roles | — | ✓ | ✓ | ✓ | ✓ |
| Autenticación nativa y de archivos | — | ✓ | ✓ | ✓ | ✓ |
| Kibana Spaces | — | ✓ | ✓ | ✓ | ✓ |
| Controles de características de Kibana | — | ✓ | ✓ | ✓ | ✓ |
| Gestión de claves de API | — | ✓ | ✓ | ✓ | ✓ |
| Logs de auditoría | — | — | ✓ | ✓ | ✓ |
| Filtrado IP | — | — | ✓ | ✓ | ✓ |
| Autenticación de LDAP, PKI*, Active Directory | — | — | ✓ | ✓ | ✓ |
| Servicio de token de Elasticsearch | — | — | ✓ | ✓ | ✓ |
| Inicio de sesión único (SAML, OpenID Connect, Kerberos) | — | — | — | ✓ | ✓ |
| Control de acceso basado en atributos | — | — | — | ✓ | ✓ |
| Seguridad a nivel de campo y documento | — | — | — | ✓ | ✓ |
| Realms de autenticación y autorización personalizados | — | — | — | ✓ | ✓ |
| Soporte de encryption at rest | — | — | — | ✓ | ✓ |
| Modo FIPS 140-2 | — | — | — | ✓ | ✓ |

Ilustración 71: Características de Seguridad por tipo de licencia

Además de los puntos comentados anteriormente, es preciso seguir un diseño de la plataforma utilizando un modelo de seguridad en capas, lo que permite proteger toda la infraestructura. En este caso, es preciso gestionar que usuarios pueden acceder a qué, y para ello es necesario formularse las siguientes preguntas, con el objetivo de conocer el propósito esperado para los usuarios:

- ¿Quién puede comunicarse con un clúster?
- ¿Quién puede agregar o eliminar documentos en un índice?
- ¿Quién puede acceder a documentos confidenciales?
- ¿Quién puede acceder a ciertos campos?

Como datos opcionales, aunque totalmente recomendables, comentar la posibilidad de disponer de logs de auditoría y cumplir los estándares de seguridad y buenas prácticas de la industria.

Mantener el sistema seguro requiere vigilancia. Las características de logs de auditoría permiten llevar un registro completo de toda la actividad del sistema y de los usuarios. Es posible filtrar la actividad para que registre solo lo que se necesite o para que registre toda la actividad que tiene lugar en el Elastic Stack. Es decir, se puede ver fácilmente quién está accediendo al clúster y qué está haciendo. Se puede configurar el nivel de auditoría, que tiene en cuenta el tipo de eventos que se registran. Estos eventos incluyen intentos de autenticación fallidos, accesos de usuario denegados, conexiones a nodos denegadas, etc. Al analizar los patrones de acceso y los intentos fallidos de acceso al clúster, se puede obtener información sobre los intentos de ataque y las brechas de datos. Mantener un registro auditable de la actividad en el clúster es de utilidad puesto que puede ayudar a diagnosticar problemas operativos.

El cumplimiento de los estándares de seguridad forma parte de Elastic (FIPS 140-2 y Sección 508), aunque se pueden emplear estándares de seguridad adicionales como HIPAA, PCI DSS, FISMA, ISO o GDPR que ayudan a cumplir y mantener el cumplimiento.

7.0 Monitorización con Elastic

Elastic es una herramienta muy utilizada para la monitorización de la infraestructura. Proporciona una serie de herramientas destinadas a la creación de Dashboards para la visualización de datos en tiempo real, a través de Kibana.

En este apartado, se mostrarán determinados casos de uso enfocados a la monitorización de la infraestructura y el análisis de la seguridad. Existen multitud de posibilidades, aunque se mostrarán algunos de los más comunes en las organizaciones.

7.1 Estado de salud de los servidores mediante métricas

Mediante el Dashboard de Overview ECS se puede ver un resumen en tiempo real del estado de los servidores monitorizados en cuanto a las métricas del sistema. Este Dashboard resulta de utilidad, ya que nos permite ver en tiempo real como están las métricas de salud de nuestros hosts, o en un momento dado. Los paneles de visualizaciones que podemos ver son los siguientes:

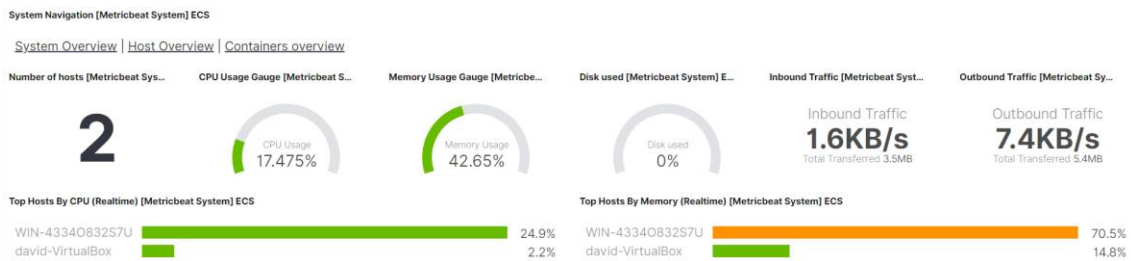


Ilustración 72: System Overview Dashboard

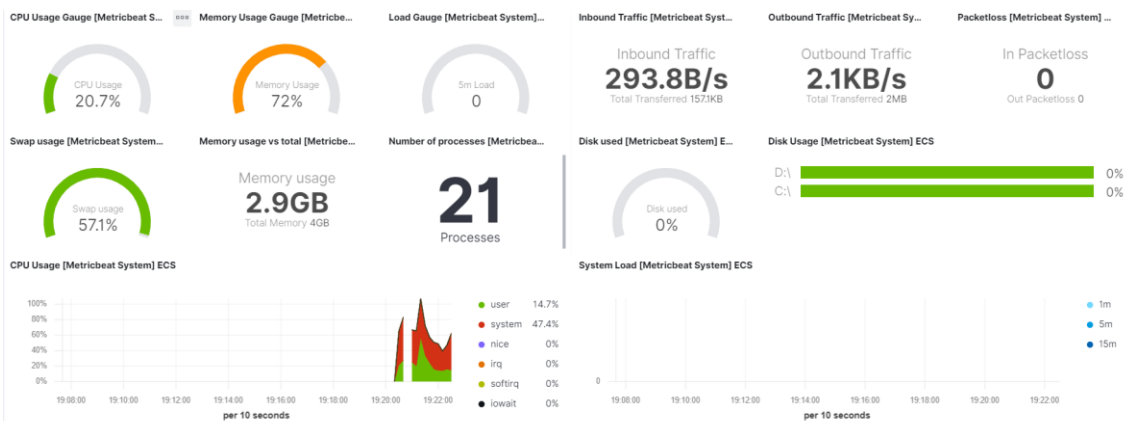


Ilustración 73: Host Overview Dashboard 1



Ilustración 74: Host Overview Dashboard 2

Estas métricas nos dan información sobre el funcionamiento de los componentes de nuestros sistemas (infraestructura, dispositivos, redes), por ejemplo el uso del CPU, la cantidad de memoria consumida, la capacidad disponible en disco, el número de procesos y dispositivos activos, la cantidad de fallos en el sistema, la cantidad de redes disponibles, el estado de la comunicación entre los dispositivos, etc. Estos valores nos permiten tener una visión clara sobre el uso de los recursos. Su simple lectura nos orienta en la gestión de la capacidad de la infraestructura para obtener mejores resultados en coste o rendimiento en un determinado momento. Aunque estas métricas siguen dando información de bajo nivel, ayudan a entender el comportamiento del sistema y su interacción.

7.2 Controlar la disponibilidad de servicios

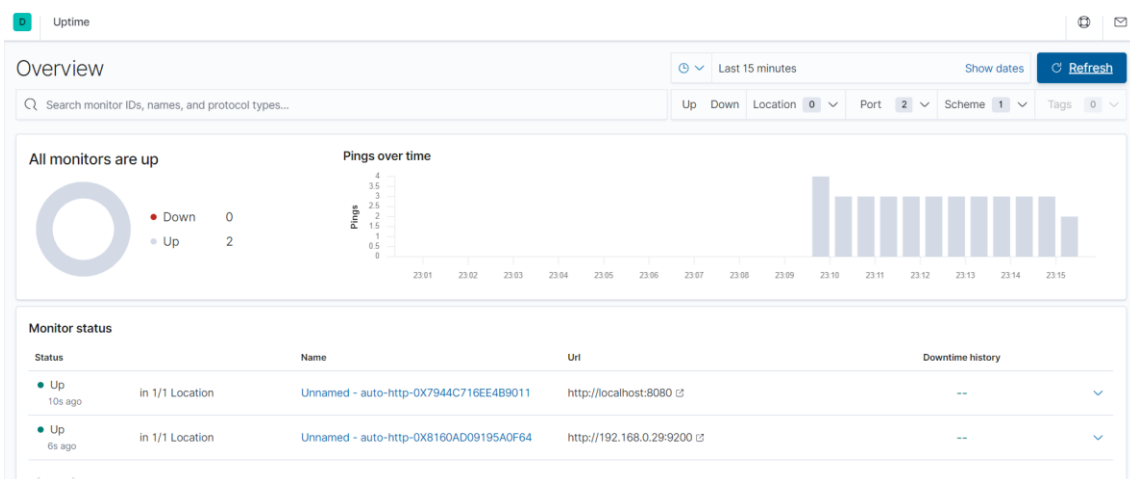


Ilustración 75: upTime Monitor

Mediante upTime Monitor se puede verificar la disponibilidad de un servicio o sistema. upTime Monitor realiza las siguientes comprobaciones:

- Comprueba si aún funciona con un recordatorio ICMP básico.
- Inspección de puertos.
- Control de HTTP/HTTPS verificando que los puntos finales específicos estén activos y devuelvan los códigos de texto y estado correctos.
- Monitoreo de API a través de validaciones y pruebas avanzadas.

En este caso, la imagen muestra la monitorización de un servidor web localizado en el Servidor Windows Server.

Además, está comprobando que el servicio de Elasticsearch esté levantado.

De esta manera, se puede mantener un control de los servicios/hosts "Alive" y "Not Alive" de la infraestructura.

7.3 Análisis de logs de Linux

Mediante el uso del Dashboard Syslog ECS se puede analizar una visión holística de los sistemas Linux y visualizar la actividad por host.

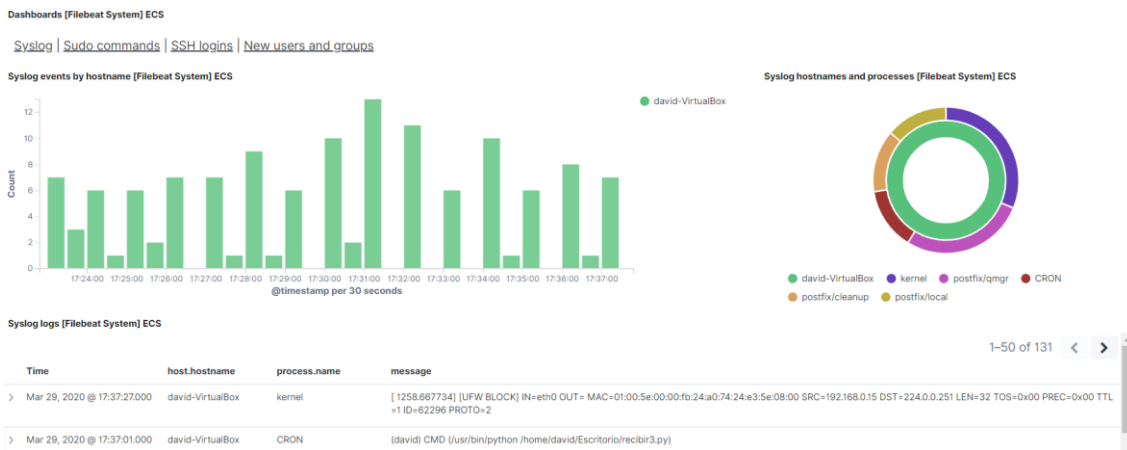


Ilustración 76: Syslog Events Dashboard Linux

Este Dashboard nos proporciona una visión general de la información de los eventos del sistema, más propiamente los log de Syslog. El registro Syslog contiene los mensajes de servicio que provienen de aplicaciones y del kernel. Con Syslog se pueden determinar problemas de seguridad, como por ejemplo:

- Un intento de acceso con contraseña equivocada
- Un acceso correcto al sistema
- Anomalías: variaciones en el funcionamiento normal del sistema
- Alertas cuando ocurre alguna condición especial
- Información sobre las actividades del sistema operativo
- Errores del hardware o el software

Con este Dashboard, se puede tener una visión general de la actividad de un host Linux en tiempo real.

7.4 Análisis de logs de Windows

Mediante el uso del Winlogbeat Dashboard ECS se puede analizar una visión holística de los sistemas Windows y visualizar la actividad por host.



Ilustración 77: WinlogBeat Dashboard Windows

De los registros de Windows se puede extraer mucha información. Por ejemplo, analizando los eventos de seguridad, se pueden ver los éxitos de inicio de sesión (4624) y los fracasos (4625), cuando se conecta un dispositivo de almacenamiento (4663) o cuando se instala un nuevo servicio (4798) en el sistema.

De la misma manera que en Linux, este Dashboard nos permite tener una visión general de la actividad en tiempo real de un host Windows.

7.5 Control de logins

Mediante el Dashboard de Links de Auditbeat System ECS se puede disponer de una visualización generalizada de los eventos de Login de los sistemas monitorizados. Con este tipo de Dashboard, se puede detectar de una forma visual un intento de sesión no autorizado.

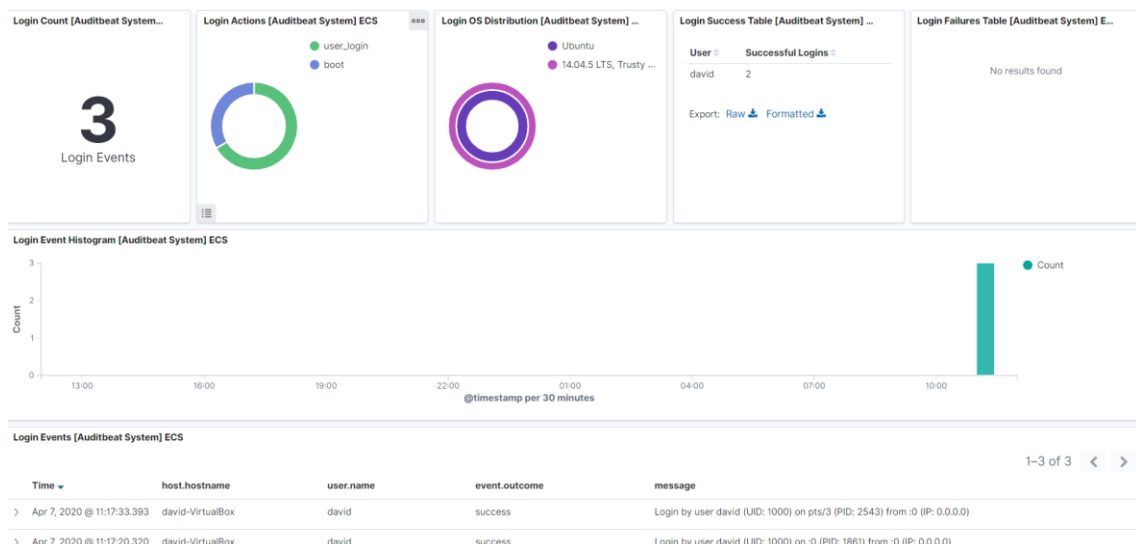


Ilustración 78: Dashboard control de logins

7.6 Control de usuarios

Mediante el Dashboard de Links de Auditbeat System, también se dispone de un apartado específico para el control de usuarios. En este Dashboard se muestra de una forma gráfica e intuitiva lo que está sucediendo con los usuarios de los sistemas monitorizados.

Con esto, se puede detectar si se crean usuarios sospechosos o si hay cambios en los permisos de los usuarios, síntomas de que los dispositivos pueden estar siendo atacados.

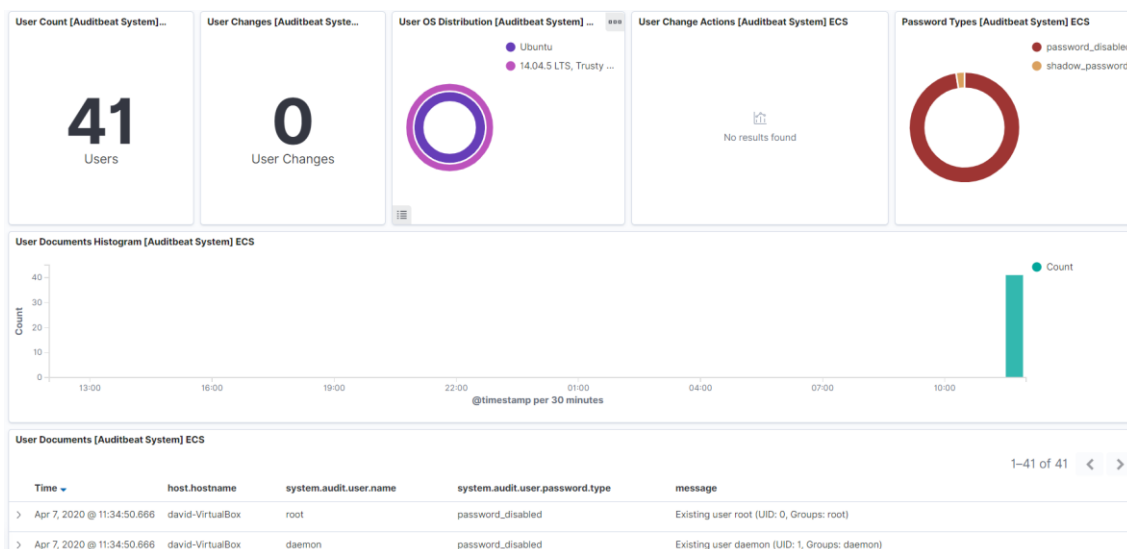


Ilustración 79: Dashboard control de users

7.7 Control de integridad de ficheros

Mediante el Dashboard de File Integrity de Auditbeat se dispone de una visión generalizada acerca de la integridad de los ficheros de los sistemas monitorizados. La monitorización de la Integridad de Ficheros (también conocido como FIM) funciona analizando los cambios de los ficheros, con lo que se puede determinar un indicio de un ataque o brecha de seguridad. La idea detrás del monitoreo de integridad de archivos es la de examinar archivos para ver qué cambió, cuándo, dónde y quién lo modificó, y así disponer de la información suficiente para recuperarse rápidamente ante un incidente o cambio no autorizado.

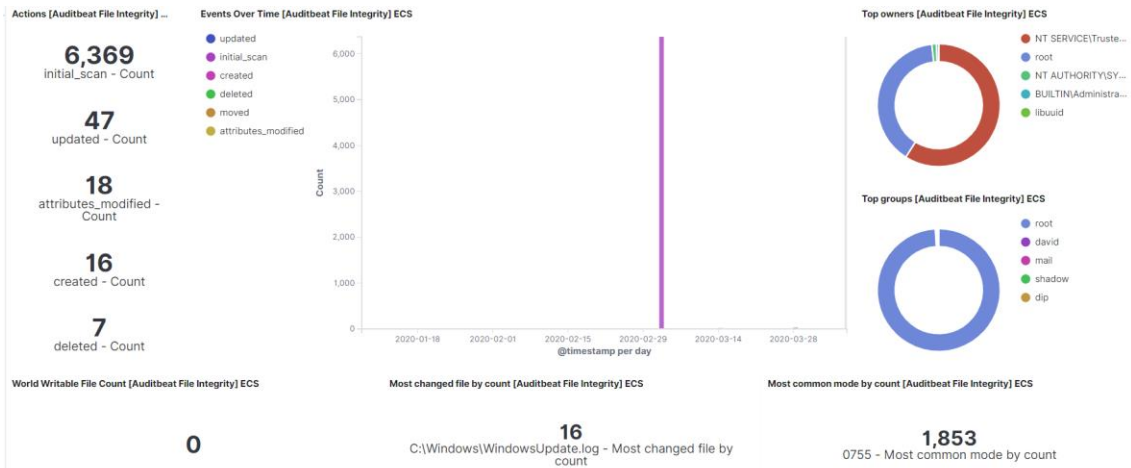


Ilustración 80: Dashboard control de integridad – 1

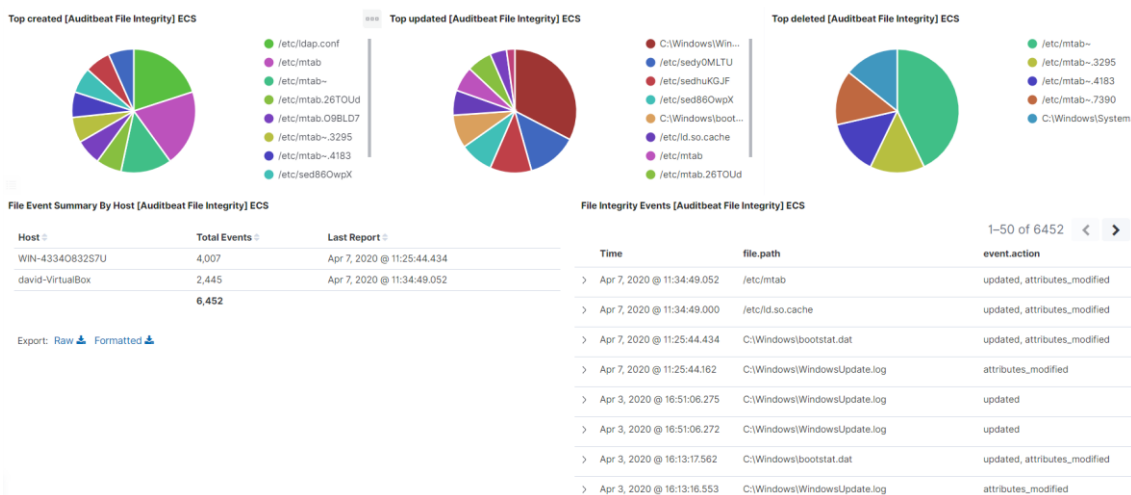


Ilustración 81: Dashboard control de integridad - 2

7.8 Control de conexiones SSH

El Dashboard de Filebeat SSH Login Attempts nos muestra aquellos intentos de conexión, ya sean satisfactorios como fallidos, al servicio SSH de los dispositivos monitorizados. Además, el Dashboard muestra información de qué usuarios han tenido intentos de conexión, la dirección IP de origen y su geolocalización.

El enfoque de seguridad de este tipo de Dashboard es la de conocer si se está recibiendo un ataque al servicio SSH. Los ataques a servicios SSH pueden realizarse mediante fuerza bruta, para intentar adivinar las combinaciones de usuarios y contraseñas. También es útil para detectar la posibilidad de accesos no autorizados a los sistemas desde orígenes sospechosos.

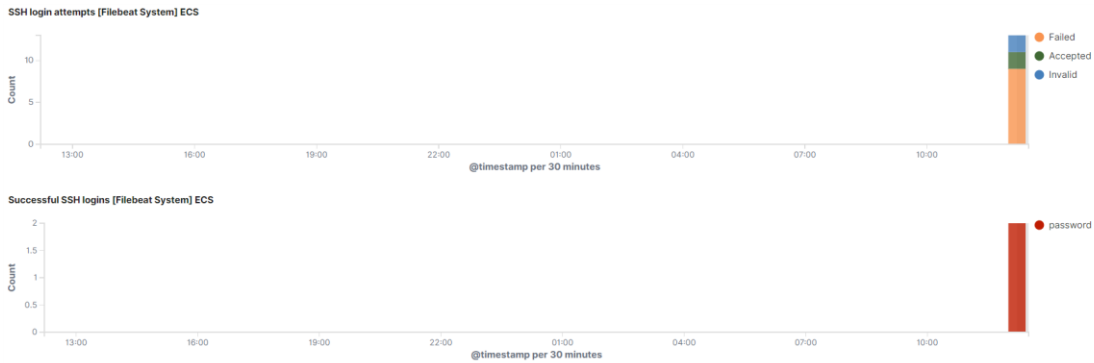


Ilustración 82: Dashboard control SSH - 1

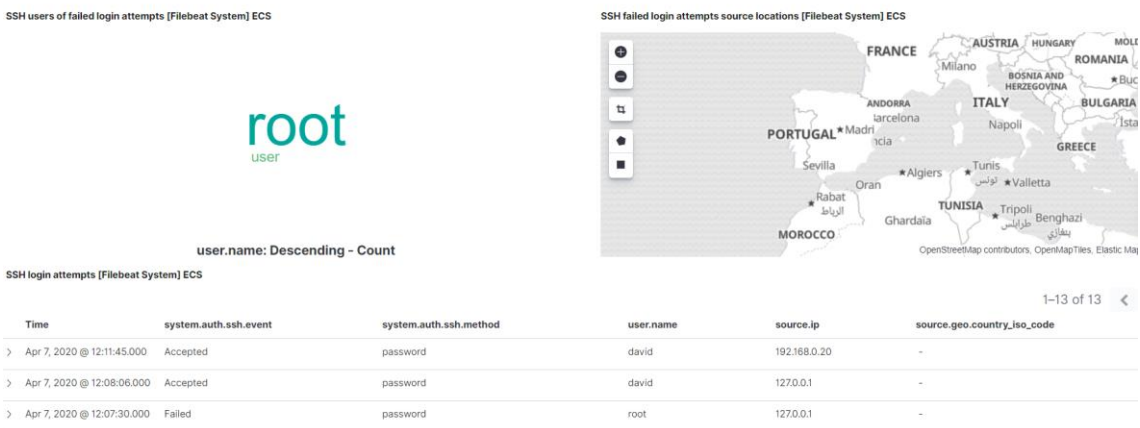


Ilustración 83: Dashboard control SSH - 2

Cabe destacar que existe una regla en el SIEM que también muestra este tipo de señales, detectando en tiempo real si se está recibiendo un ataque desde el exterior de la red.

7.9 Control de procesos

A través del Dashboard Links de Auditbeat System se puede visualizar en tiempo real los procesos en ejecución de los sistemas monitorizados. Aparte de la utilidad de saber la salud de los servicios en ejecución para detectar servicios caídos, este Dashboard permite controlar la aparición de procesos sospechosos que puedan indicar la presencia de Malware.

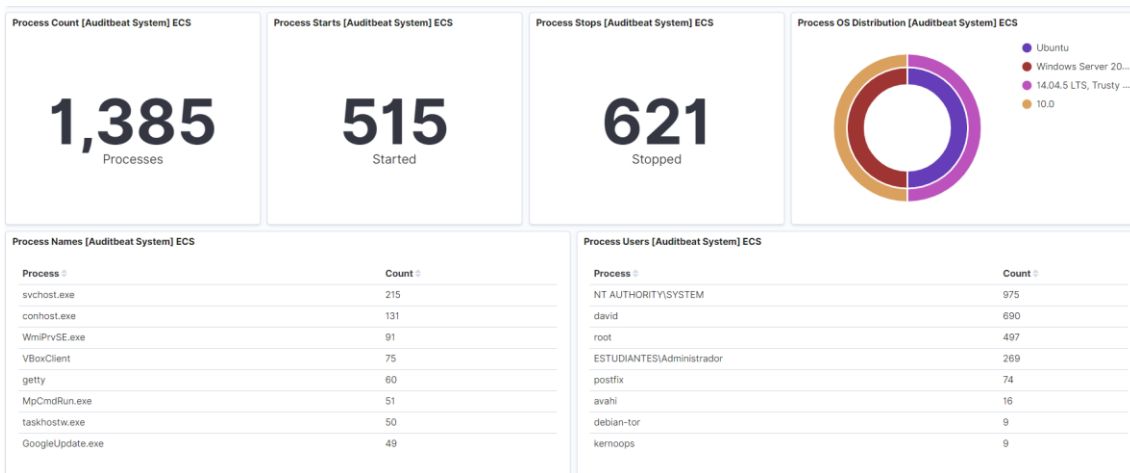


Ilustración 84: Dashboard control de procesos - 1

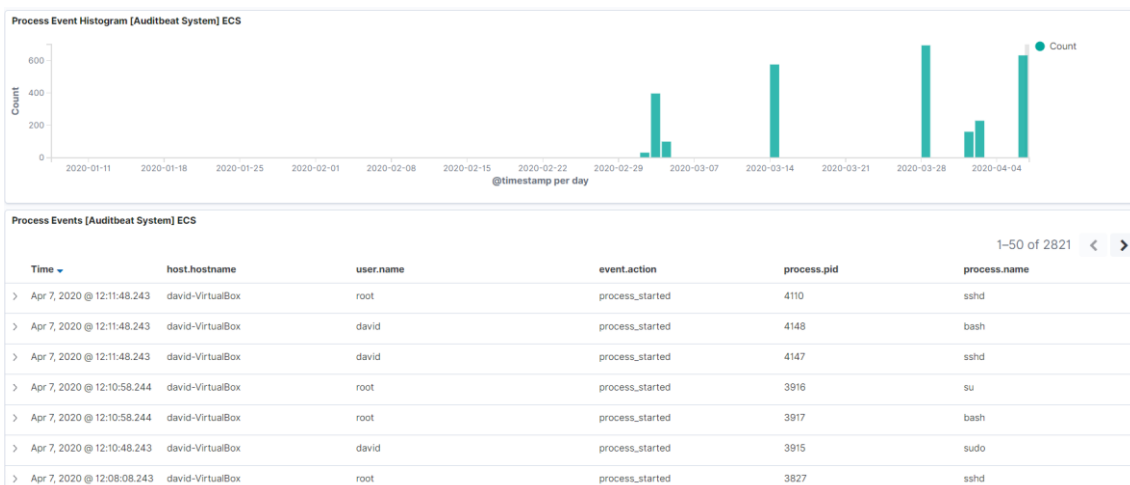


Ilustración 85: Dashboard control de procesos – 2

7.10 Flow de paquetes

A través del Dashboard Navigation de packbeats se puede visualizar en tiempo real los flujos de red de la infraestructura.

Packetbeat es capaz de recopilar e informar de las estadísticas de los flujos de la red. Un flujo es un grupo de paquetes enviados durante el mismo período de tiempo que comparten propiedades comunes, como la misma dirección y protocolo de origen y destino. Se puede utilizar esta función para analizar el tráfico de red sobre protocolos específicos en la red, como por ejemplo HTTP, con objetivo de monitorizarlos.

Para cada flujo, Packetbeat informa el número de paquetes y el número total de bytes enviados desde el origen hasta el destino. Cada evento de flujo también contiene información sobre los hosts de origen y destino, como su dirección IP.

Es interesante destacar que conocer los flujos de red permite detectar intentos de conexión sospechosos, los cuales pueden provenir de algún Malware instalado en los equipos, ya que es común que los Malware traten de conectarse a la red.

Además, también se incluye una vista de tráfico DNS que, al igual que Cisco Umbrella, se puede utilizar para detectar resoluciones de nombres maliciosas o cuyo destino sea no autorizado.



Ilustración 86: Flujo packetbeat – 1

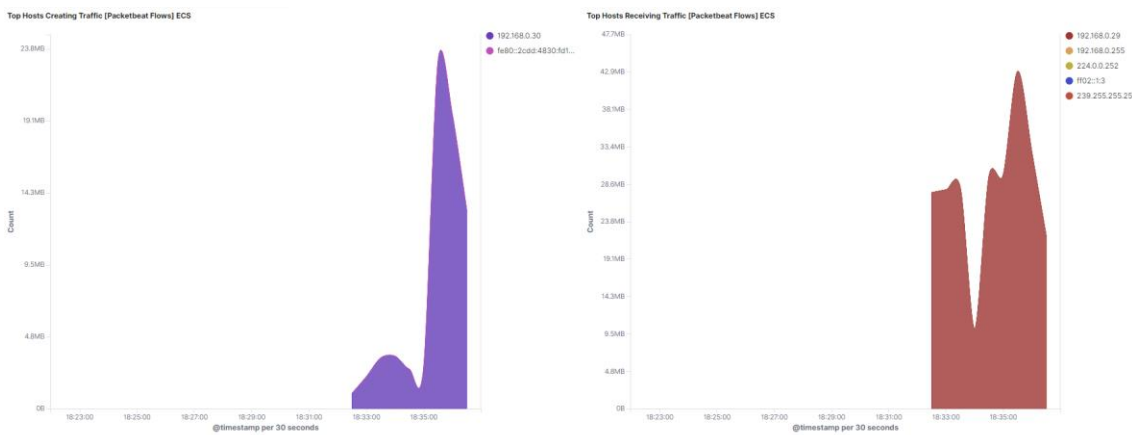


Ilustración 87: Flujo packetbeat – 2

Network Traffic Between Hosts (Packetbeat Flows) ECS

| Source IP | Destination IP | Source Bytes | Destination Bytes |
|---------------------------|-----------------|--------------|-------------------|
| 192.168.0.30 | 192.168.0.29 | 71.1MB | 250.7MB |
| 192.168.0.30 | 239.255.255.250 | 6.8KB | 0B |
| 192.168.0.30 | 192.168.0.255 | 3.5KB | 0B |
| 192.168.0.30 | 224.0.0.252 | 2.8KB | 0B |
| fe80::2cdd:4830:f111:3c5d | m02:1:3 | 3.6KB | 0B |

Export: Raw | Formatted

Ilustración 88: Flujo packetbeat – 3

8.0 Conclusiones

Durante la realización del proyecto, se ha investigado principalmente sobre las funcionalidades de Elastic enfocadas a la analítica de datos de seguridad. Con esta investigación, se ha obtenido conocimiento acerca del funcionamiento de la analítica de seguridad y su implementación en un entorno corporativo.

El análisis de seguridad es el proceso de usar herramientas de recopilación, agregación y análisis de datos para la monitorización de la seguridad de la infraestructura y detección de amenazas.

En función de los tipos de herramientas utilizadas en la organización, se pueden incorporar conjuntos de datos de diversas fuentes. En este caso, se han utilizado fuentes de datos provenientes de los productos de seguridad más utilizados en el mercado, como Fortinet, Cisco, Checkpoint, etc. Esto permite disponer de fuentes de información que aportan datos que se utilizarán para la analítica de seguridad, tales como:

- Tráfico de red.
- Los datos de comportamiento del usuario.
- Aplicaciones de negocios.
- Datos contextuales no informáticos.
- Datos de gestión de la identidad y el acceso.
- Fuentes de inteligencia de amenazas externas.
- Errores en los sistemas.

Gracias a los recientes avances en la analítica de seguridad, Elastic incluye sistemas de aprendizaje automático adaptativo, que afinan los modelos de detección basados en la experiencia y aprendizajes, así como la lógica de detección de anomalías. Este tipo de avances son muy positivos, ya que permiten el análisis de datos en tiempo real, necesario para la respuesta inmediata a amenazas (Incident Response) así como para investigaciones (Threat Hunting).

El análisis de seguridad tiene una variedad de casos de uso, entre los que destacan:

- Analizar el comportamiento del usuario para detectar patrones potencialmente sospechosos.
- Analizar el tráfico de la red para determinar las tendencias que indican posibles ataques.
- Identificar el uso indebido de las cuentas de usuario.
- Detección de Malware.
- Detección de la exfiltración de datos por los atacantes.
- Detección de amenazas internas.
- Identificar las cuentas comprometidas.
- Investigar los incidentes de seguridad.
- Threat Hunting.

- Demostración del cumplimiento durante las auditorías.

Es necesario destacar que, el objetivo principal de la analítica de seguridad, es convertir los datos sin procesar de fuentes dispares en información procesable, con el propósito de identificar eventos que requieren de una respuesta inmediata. Esto se consigue a través de la correlación de eventos y alertas.

Elastic es en una herramienta que agrega un filtro crítico a los volúmenes de datos obtenidos, y con esto se consigue obtener de forma eficaz una respuesta satisfactoria ante las brechas de seguridad de una organización.

Se puede concluir que Elastic es una solución muy completa y recomendable para todo tipo de organizaciones, especialmente para aquellas pequeñas o medianas empresas que no pueden realizar grandes inversiones en materia de ciberseguridad. Las características Open-Source de Elastic son un arma completa para establecer algunas de las medidas necesarias para la detección de amenazas y para la monitorización de la infraestructura y servicios, aunque para poder disponer de una solución más completa y avanzada es necesario obtener una licencia de pago, ya que permite disfrutar de todas las características que ofrece la herramienta.

También existe la posibilidad de utilizar estudiar la implementación de otras herramientas del mercado, o de la propia integración de productos de otras compañías con Elastic, en función de las necesidades de cada organización.

Es posible que existan herramientas en el mercado más completas que Elastic y con mayores posibilidades, pero por otro lado el desembolso económico es probable que sea bastante mayor.

En razón de lo expuesto anteriormente, los objetivos planteados inicialmente para el proyecto se han cumplido. Por un lado, los objetivos generales:

- ✓ Estudiar las capacidades de Elastic SIEM.
- ✓ Estudiar el funcionamiento del Machine Learning en la detección de amenazas de seguridad.
- ✓ Estudiar las capacidades de Elastic relacionadas con la analítica de datos y la utilidad que esto ofrece para la ciberseguridad.

Por otro lado, los objetivos específicos:

- ✓ Una herramienta que permita mantener el control de la seguridad de los activos en un entorno empresarial.
- ✓ Una herramienta que provee de un sistema de monitorización en tiempo real.
- ✓ Una herramienta que provee de un sistema de detección de amenazas en tiempo real.
- ✓ Una herramienta que provee de un sistema de prevención de amenazas de seguridad.
- ✓ Una herramienta que permita gestionar los principales riesgos de seguridad del entorno empresarial asociados a fallas de seguridad.
- ✓ Estudiar la viabilidad, en cuanto a ciberseguridad, de esta solución empresarial.

De todos modos, es necesario también comentar aquellas tareas que no se han podido realizar, así como los problemas surgidos durante la realización de ciertas tareas que también afectan a las cualidades de Elastic. A continuación, las reseñas más destacables:

- ❖ El proceso de normalización de eventos del SIEM requiere de intervención humana. El *out of the box* de Elastic es se queda un poco escaso en cuanto a la compatibilidad de los módulos con diferentes fuentes de información, por lo que no es compatible con todas las herramientas del mercado, como es lógico. Aunque es cierto que están trabajando en ello, la lista de compatibles todavía es un poco escasa. Por ello, es preciso realizar una tarea de normalización de eventos al estándar ECS, lo cual implica invertir bastante tiempo. Si una empresa dispone de múltiples fuentes de información de las cuales Elastic no es compatible, esto podría ser inviable.
- ❖ No se ha podido crear un laboratorio de Malware para realizar pruebas de seguridad. Aparte del riesgo que supone, no se disponía del material suficiente para la creación de una *sandbox* física. Aunque se ha intentado con una máquina virtual, la inmensa mayoría del Malware no llegaba a ejecutarse dado que detectaban que estaban siendo ejecutados en máquinas virtuales. De todos modos, durante la realización de las pruebas se pudieron visualizar en Elastic algunos de los procesos que se ejecutaban, los cuales eran obviamente sospechosos, así como algún intento de conexión al exterior sospechoso.
- ❖ Las reglas de correlación de eventos pre-construidas son insuficientes. Esto conlleva a que los administradores necesiten incorporar múltiples reglas de correlación que permitan mantener la infraestructura segura, lo cual requiere de invertir mucho tiempo. Esto es un punto importante a tener en cuenta, aunque es necesario tener en cuenta que esta solución es gratuita.
- ❖ Las reglas de detección de amenazas se han encontrado poco personalizables. En este momento, no se permite crear reglas en base a umbrales o cuentas durante un período de tiempo.
- ❖ Las características de Machine Learning requieren de unas capacidades de Hardware elevadas. Este tipo de analítica requiere de potencia computacional elevada, ya que trabajan sobre grandes volúmenes de datos. Es por eso que no se pudo emplear el Machine Learning para la detección de anomalías de seguridad relacionadas con el Hardware.

Con todo lo comentado anteriormente, la planificación de las tareas se ha visto afectada. Aunque se ha intentado seguir al pie de la letra las tareas expuestas al comienzo de esta memoria, algunas se han tenido que ver modificadas:

- Dado que los logs de Cisco Umbrella contenían más información, se ha prescindido de los logs del servicio bind de DNS.
- Dado que los logs de Checkpoint, Cisco Umbrella y Fortinet no eran compatibles con Elastic, se ha trabajado en la normalización de eventos de los logs de Cisco Umbrella y de Fortinet para optimizar el tiempo. Por tanto, las tareas relacionadas con la analítica de datos solo se han llevado con los comentados.
- Aunque se ha trabajado en ello y se ha intentado, las tareas relacionadas con la instalación y detección de Malware no se han introducido en la memoria puesto que no se han alcanzado los objetivos deseados.

En cuanto a la planificación temporal, la estimación ha resultado satisfactoria, pudiendo haber una disparidad de +/- 2 días de media en algunas tareas, aunque se han cumplido los objetivos en todos los hitos establecidos.

8.1 Líneas de trabajo futuro

En cuanto a líneas de trabajo futuro, se podrían llevar a cabo las siguientes determinaciones:

- Incrementar las fuentes de información, mediante la ingesta de más logs, como logs de servicios, aplicaciones web, bases de datos, correo electrónico, Directorio Activo, etc; y mediante el uso de más herramientas de seguridad, como sistemas IPS/IDS, Elastic Endpoint, Antivirus, Sandbox, Proxies, Biometría, etc.
- Cruzar las fuentes de información con fuentes de inteligencia, tales como OSINT, HUMINT, etc para mejorar la detección de anomalías y amenazas de seguridad.
- Utilizar Elastic para la detección de Malware en tiempo real mediante el uso de técnicas de Machine Learning para la detección de anomalías de seguridad.
- Llevar a cabo un estudio de la viabilidad de Elastic para tareas de Threat Hunting.
- Estudiar la herramienta Elastic para llevar a cabo tareas de cumplimiento (Compliance).
- Utilizar Elastic para la monitorización y analítica de seguridad en plataformas Cloud.
- Integrar Elastic con otras herramientas SIEM como IBM Qradar, ArchSight, AliemVault OSSIM, etc, para estudiar su viabilidad y características en cuanto a la detección de anomalías de seguridad.
- Integrar Elastic con los escáneres de vulnerabilidades más utilizados del mercado, como Nessus, Qualys, OpenVAS, OWASP Zap, Burp Suite, Acunetix, etc, que permitan el uso de Elastic como herramienta para la gestión de las vulnerabilidades de una organización.

9.0 Glosario

Amenaza: Circunstancia desfavorable que puede ocurrir y que cuando sucede tiene consecuencias negativas sobre los activos provocando su indisponibilidad, funcionamiento incorrecto o pérdida de valor. Una amenaza puede tener causas naturales, ser accidental o intencionada. Si esta circunstancia desfavorable acontece a la vez que existe una vulnerabilidad o debilidad de los sistemas o aprovechando su existencia, puede derivar en un incidente de seguridad.

Ciberseguridad: Protección de activos de información, a través del tratamiento de amenazas que ponen en riesgo la información que es procesada, almacenada y transportada por los sistemas de información que se encuentran interconectados

Cortafuegos/Firewall: Sistema de seguridad compuesto o bien de programas (software) o de dispositivos hardware situados en los puntos limítrofes de una red que tienen el objetivo de permitir y limitar, el flujo de tráfico entre los diferentes ámbitos que protege sobre la base de un conjunto de normas y otros criterios. La funcionalidad básica de un cortafuego es asegurar que todas las comunicaciones entre la red e Internet se realicen conforme a las políticas de seguridad de la organización o corporación.

Fuerza Bruta: Un ataque de fuerza bruta es un procedimiento para averiguar una contraseña que consiste en probar todas las combinaciones posibles hasta encontrar la combinación correcta.

IDS: Un sistema de detección de intrusos (o IDS de sus siglas en inglés Intrusion Detection System) es una aplicación usada para detectar accesos no autorizados a un ordenador o a una red. Estos accesos pueden ser ataques realizados por usuarios malintencionados con conocimientos de seguridad o usando herramientas automáticas. A diferencia de los IPS, estos sistemas sólo detectan intentos de acceso y no tratan de prevenir su ocurrencia.

Incidente de Seguridad: Cualquier suceso que afecte a la confidencialidad, integridad o disponibilidad de los activos de información de la empresa, por ejemplo: acceso o intento de acceso a los sistemas, uso, divulgación, modificación o destrucción no autorizada de información.

Log: Un log o registro es un archivo en el que se almacenan cronológicamente los acontecimientos que han ido afectando a un sistema informático.

Malware: Es un tipo de software que tiene como objetivo dañar o infiltrarse sin el consentimiento de su propietario en un sistema de información. Palabra que nace de la unión de los términos en inglés de software malintencionado: malicious software. Dentro de esta definición tiene cabida un amplio elenco de programas maliciosos: virus, gusanos, troyanos, backdoors, spyware, etc. La nota común a todos estos programas es su carácter dañino o lesivo.

Metodología ágil: Son metodologías de desarrollo de proyectos basadas en el desarrollo iterativo e incremental. Fomentan respuestas rápidas y flexibles al cambio mediante una planificación adaptada, la identificación de requisitos y la racionalización entre el equipo de trabajo.

Parsear: Proceso de analizar una secuencia de símbolos a fin de determinar su estructura gramatical con respecto a una gramática formal dada.

Phishing: Se define como el conjunto de técnicas que persiguen el engaño a una víctima ganándose su confianza, haciéndose pasar por una persona, empresa o servicio de confianza, para manipularla y conseguir que realice acciones a la voluntad del atacante.

SIEM: SIEM o Gestión de Eventos e Información de Seguridad (Security Information and Event Management) es una categoría de software que tiene como objetivo otorgar a las organizaciones información útil sobre potenciales amenazas de seguridad de sus redes críticas de negocio, a través de la estandarización de datos y priorización de amenazas.

URL: Las siglas URL (Uniform Resource Locator) hacen referencia a la dirección que identifica un contenido colgado en Internet.

10.0 Bibliografía

- Kanban. (2019). En J. Edge, *La guía definitiva de la metodología Kanban para el desarrollo de software ágil*. James Edge.
- ECS Guide 1.5 | Service Fields. (30 de 03 de 2020). Obtenido de <https://www.elastic.co/guide/en/ecs/master/ecs-service.html>
- Elasti. (19 de 04 de 2020). *Packetbeat | Elastic Guide 7.6.0*. Obtenido de <https://www.elastic.co/guide/en/beats/packetbeat/current/packetbeat-installation.html>
- Elastic. (09 de 03 de 2020). *Beats*. Obtenido de <https://www.elastic.co/es/beats>
- Elastic. (9 de 03 de 2020). *Características de Elastic Stack*. Obtenido de <https://www.elastic.co/es/elastic-stack/features>
- Elastic. (01 de 04 de 2020). *Configuración de SSL, TLS y HTTPS para asegurar Elasticsearch, Kibana, Beats y Logstash*. Obtenido de <https://www.elastic.co/es/blog/configuring-ssl-tls-and-https-to-secure-elasticsearch-kibana-beats-and-logstash#enable-tls-kibana>
- Elastic. (16 de 03 de 2020). *Dashboard | Kibana Guide 7.6.0*. Obtenido de <https://www.elastic.co/guide/en/kibana/current/dashboard.html>
- Elastic. (23 de 03 de 2020). *ECS DNS | ECS Guide 1.5*. Obtenido de <https://www.elastic.co/guide/en/ecs/current/ecs-dns.html>
- Elastic. (26 de 03 de 2020). *ECS Guide 1.5 | DNS*. Obtenido de <https://www.elastic.co/guide/en/ecs/current/ecs-dns.html>
- Elastic. (26 de 03 de 2020). *ECS guide 1.5 | Event*. Obtenido de <https://www.elastic.co/guide/en/ecs/current/ecs-event.html>
- Elastic. (26 de 03 de 2020). *ECS Guide 1.5 | File*. Obtenido de <https://www.elastic.co/guide/en/ecs/current/ecs-file.html>
- Elastic. (30 de 03 de 2020). *ECS Guide 1.5 | Vulnerability Fields*. Obtenido de <https://www.elastic.co/guide/en/ecs/current/ecs-vulnerability.html>
- Elastic. (25 de 02 de 2020). *Elastic*. Obtenido de <https://www.elastic.co/es/>
- Elastic. (22 de 03 de 2020). *Elastic Common Schema (ECS) | ECS Guide 1.5*. Obtenido de <https://www.elastic.co/guide/en/ecs/current/index.html>
- Elastic. (10 de 03 de 2020). *Elastic Discuss*. Obtenido de <https://discuss.elastic.co>
- Elastic. (20 de 05 de 2020). *Elastic Guide 7.1 | Hardening*. Obtenido de <https://www.elastic.co/guide/en/elasticsearch/reference/current/secure-cluster.html>
- Elastic. (01 de 04 de 2020). *Elastic Guide 7.6 | Detection Engine Overview*. Obtenido de <https://www.elastic.co/guide/en/siem/guide/current/detection-engine-overview.html>
- Elastic. (02 de 04 de 2020). *Elastic Guide 7.6 | Rules Create*. Obtenido de <https://www.elastic.co/guide/en/siem/guide/7.6/rules-ui-create.html>
- Elastic. (01 de 03 de 2020). *Elastic Stack*. Obtenido de <https://www.elastic.co/es/elastic-stack>
- Elastic. (01 de 03 de 2020). *Elasticserach*. Obtenido de <https://www.elastic.co/es/what-is/elasticsearch>
- Elastic. (14 de 03 de 2020). *Extracting Fields and Wrangling Data*. Obtenido de <https://www.elastic.co/guide/en/logstash/current/field-extraction.html>

- Elastic. (14 de 03 de 2020). *File input plugin | Logstash Reference 7.6.0*. Obtenido de <https://www.elastic.co/guide/en/logstash/current/plugins-inputs-file.html>
- Elastic. (03 de 03 de 2020). *Guía Elastic 7.6.0*. Obtenido de <https://www.elastic.co/guide/index.html>
- Elastic. (01 de 03 de 2020). *Historia de Elasticsearch*. Obtenido de <https://www.elastic.co/es/about/history-of-elasticsearch>
- Elastic. (15 de 03 de 2020). *Index management | Kibana Guide 7.6.0*. Obtenido de <https://www.elastic.co/guide/en/kibana/current/managing-indices.html>
- Elastic. (10 de 03 de 2020). *Input plugins | Logstash Reference 7.6.0*. Obtenido de <https://www.elastic.co/guide/en/logstash/current/input-plugins.html>
- Elastic. (28 de 02 de 2020). *Logstash*. Obtenido de <https://www.elastic.co/es/logstash>
- Elastic. (04 de 05 de 2020). *Machine Learning | Kibana Guide 7.6.0*. Obtenido de <https://www.elastic.co/guide/en/machine-learning/current/ml-getting-started.html>
- Elastic. (20 de 04 de 2020). *Machine Learning trabajos preconstruidos | Elastic Guide 7.6.0*. Obtenido de <https://www.elastic.co/guide/en/siem/guide/current/prebuilt-ml-jobs.html>
- Elastic. (14 de 03 de 2020). *Output plugins | Logstash Reference 7.6.0*. Obtenido de <https://www.elastic.co/guide/en/logstash/current/output-plugins.html>
- Elastic. (26 de 02 de 2020). *Qué es Elasticsearch*. Obtenido de <https://www.elastic.co/es/what-is/elasticsearch>
- Elastic. (8 de 03 de 2020). *Qué es Kibana*. Obtenido de <https://www.elastic.co/es/what-is/kibana>
- Elastic. (05 de 10 de 2020). *Watcher | Kibana Guide 7.6.0*. Obtenido de <https://www.elastic.co/guide/en/kibana/current/watcher-ui.html>
- Elastic. (15 de 05 de 2020). *Watcher Avanzado | Elastic Guide 7.6.0*. Obtenido de <https://www.elastic.co/guide/en/elasticsearch/reference/7.6/how-watcher-works.html#watch-definition>
- INCIBE. (12 de 03 de 2020). *Glosario Ciberseguridad*. Obtenido de https://www.incibe.es/sites/default/files/contenidos/guias/doc/guia_glosario_ciberseguridad_metad.pdf
- Kumar, P. S. (2019). *Elastic Stack 7.0 Second Edition*. En P. S. Kumar, *Learning Elastic Stack 7.0 Second*. Packt>.
- Martin, M. (10 de 07 de 2019). *Migración a Elastic Common Schema (ECS) en ambientes de Beats*. Obtenido de <https://www.elastic.co/es/blog/migrating-to-elastic-common-schema-in-beats-environments>
- Trello. (19 de 02 de 2020). *Trello*. Obtenido de <https://trello.com/>

11.0 Anexos

A continuación, se muestran como anexos los pasos seguidos para la instalación y configuración del Software empleado en el desarrollo del proyecto.

11.1 Instalación y configuración de ElasticSearch, Kibana y Logstash

El primer paso es descargar el Software desde la página oficial de Elastic. Para ello, se descargan los archivos RPM de las siguientes URLs:

- <https://www.elastic.co/es/downloads/elasticsearch>
- <https://www.elastic.co/es/downloads/kibana>
- <https://www.elastic.co/es/downloads/logstash>

11.1.1 Instalación de Elasticsearch

Una vez descargados los paquetes RPM, se utiliza el siguiente comando para instalar Elasticsearch:

```
yum -y install elasticsearch-7.6.0-x86_64.rpm
```

```
Tamaño total: 475 M
Tamaño instalado: 475 M
Downloading packages:
Running transaction check
Running transaction test
Transaction test succeeded
Running transaction
Creating elasticsearch group... OK
Creating elasticsearch user... OK
  Instalando      : elasticsearch-7.6.0-1.x86_64                      1/1
### NOT starting on installation, please execute the following statements to con
figure elasticsearch service to start automatically using systemd
  sudo systemctl daemon-reload
  sudo systemctl enable elasticsearch.service
### You can start elasticsearch service by executing
  sudo systemctl start elasticsearch.service
Created elasticsearch keystore in /etc/elasticsearch
  Comprobando    : elasticsearch-7.6.0-1.x86_64                      1/1

Instalado:
  elasticsearch.x86_64 0:7.6.0-1

¡Listo!
[root@localhost Descargas]# █
```

Ilustración 89: Instalación Elasticsearch

Para que el Software se ejecute en el inicio tras encender el servidor, es necesario realizar el siguiente comando:


```
systemctl enable elasticsearch
```

Para arrancar el servicio Elasticsearch se utiliza el siguiente comando:

```
systemctl start elasticsearch
```

Para verificar que Elasticsearch se ha instalado correctamente, se accede a la siguiente URL:

<http://localhost:9200>

El resultado debería ser el siguiente:

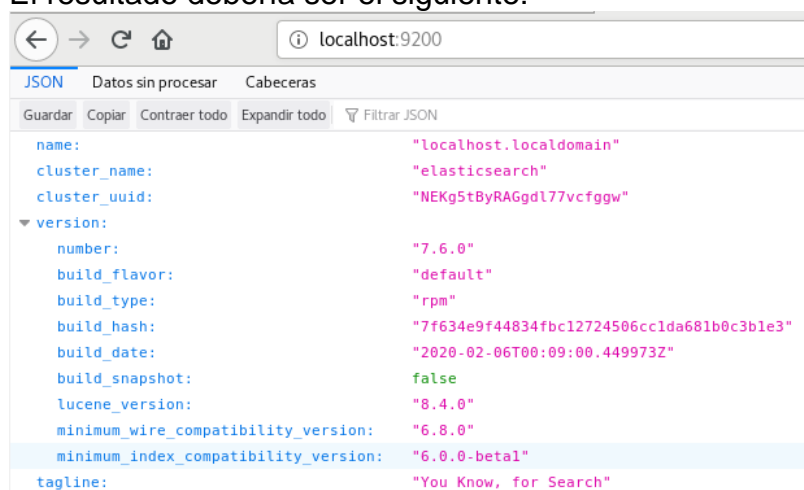


Ilustración 90: Verificación de instalación de Elasticsearch

11.1.2 Configuración de Elasticsearch

El fichero de configuración de Elasticsearch se encuentra en la ruta */etc/elasticsearch*

Ahí, hay un fichero que incluye las opciones de configuración. El fichero que se debe modificar es *elasticsearch.yml*

Para este laboratorio, se dejará por defecto a excepción de las siguientes variables:

```
# ----- Network -----  
#  
# Set the bind address to a specific IP (IPv4 or IPv6):  
#  
network.host: 0.0.0.0  
#  
# Set a custom port for HTTP:  
#  
#http.port: 9200  
#
```

Ilustración 91: Elasticsearch.yml – 1

```
# ----- Discovery -----
#
# Pass an initial list of hosts to perform discovery when this node is started:
# The default list of hosts is ["127.0.0.1", "[::1]"]
#
#discovery.seed_hosts: ["host1", "host2"]
#
# Bootstrap the cluster using an initial set of master-eligible nodes:
#
cluster.initial_master_nodes: ["node-1", "node-2"]
#
# For more information, consult the discovery and cluster formation module documentation.
#
```

Ilustración 92: Elasticsearch.yml – 2

11.1.3 Instalación de Kibana

Una vez descargados los paquetes RPM, se utiliza el siguiente comando para instalar Kibana:

```
yum -y install kibana-7.6.0-x86_64.rpm
```

```
Instalando:
  kibana      x86_64      7.6.0-1      /kibana-7.6.0-x86_64      676 M

Resumen de la transacción
=====
Instalar 1 Paquete

Tamaño total: 676 M
Tamaño instalado: 676 M
Downloading packages:
Running transaction check
Running transaction test
Transaction test succeeded
Running transaction
  Instalando      : kibana-7.6.0-1.x86_64      1/1
  Comprobando    : kibana-7.6.0-1.x86_64      1/1

Instalado:
  kibana.x86_64 0:7.6.0-1

¡Listo!
[root@localhost Descargas]#
```

Ilustración 93: Instalación de Kibana

Para que el Software se ejecute en el inicio tras encender el servidor, es necesario realizar el siguiente comando:

```
systemctl enable kibana
```

Para arrancar el servicio Kibana se utiliza el siguiente comando:

```
systemctl start kibana
```

Para verificar que Kibana se ha instalado correctamente, se accede a la siguiente URL:
<http://localhost:5601>

El resultado debería ser el siguiente:

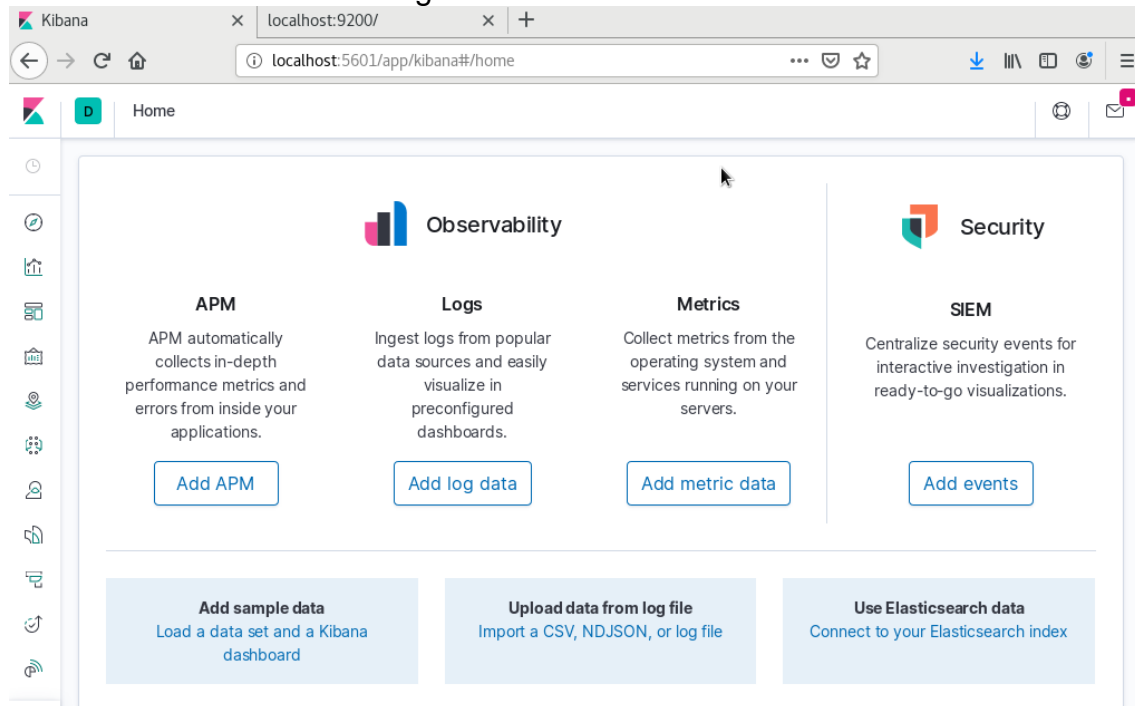


Ilustración 94: Verificación de instalación de Kibana

11.1.4 Configuración de Kibana

El fichero de configuración de Kibana se encuentra en la ruta `/etc/kibana`. Ahí, hay un fichero que incluye las opciones de configuración. El fichero que se debe modificar es `kibana.yml`

Para este laboratorio, se dejará la configuración por defecto.

11.1.5 Instalación de Logstash

Una vez descargados los paquetes RPM, se utiliza el siguiente comando para instalar Kibana:

```
yum -y install logstash-7.6.0-x86_64.rpm
```

```

Resumen de la transacción
=====
Instalar 1 Paquete

Tamaño total: 283 M
Tamaño instalado: 283 M
Downloading packages:
Running transaction check
Running transaction test
Transaction test succeeded
Running transaction
  Instalando   : 1:logstash-7.6.0-1.noarch                               1/1
Using provided startup.options file: /etc/logstash/startup.options
/usr/share/logstash/vendor/bundle/jruby/2.5.0/gems/pleaserun-0.0.30/lib/pleaserun/platform/b
ase.rb:112: warning: constant ::Fixnum is deprecated
Successfully created system startup script for Logstash
  Comprobando  : 1:logstash-7.6.0-1.noarch                               1/1

Instalado:
  logstash.noarch 1:7.6.0-1

¡Listo!
[root@localhost Descargas]# █

```

Ilustración 95: Instalación de Logstash

Para que el Software se ejecute en el inicio tras encender el servidor, es necesario realizar el siguiente comando:

```
systemctl enable logstash
```

Para arrancar el servicio logstash se utiliza el siguiente comando:

```
systemctl start logstash
```

11.1.6 Configuración de Logstash

El fichero de configuración de Logstash se encuentra en la ruta */etc/logstash*. Ahí, hay un fichero que incluye las opciones de configuración. El fichero que se debe modificar es *logstash.yml*

Para este laboratorio, se dejará por defecto.

11.2 Instalación y configuración de MetricBeat

Para instalar MetricBeat hay que descargarlo de la página oficial de Elastic: <https://www.elastic.co/es/downloads/beats/metricbeat>

Una vez descargado se procede a su instalación.

Caso Ubuntu

Se instala el paquete DEB mediante el siguiente comando:

```
sudo dpkg -i metricbeat-7.6.0-amd64.deb
```

Una vez instalado, se accede a la ruta `/etc/metricbeat` para su configuración. Se ha de modificar el fichero `metricbeat.yml` con la siguiente información:

```
#===== Kibana =====  
  
# Starting with Beats version 6.0.0, the dashboards are loaded via the Kibana AS  
# This requires a Kibana endpoint configuration.  
setup.kibana:  
  
  # Kibana Host  
  # Scheme and port can be left out and will be set to the default (http and 56$  
  # In case you specify and additional path, the scheme is required: http://loc$  
  # IPv6 addresses should always be defined as: https://[2001:db8::1]:5601  
  host: "192.168.0.29:5601"  
  
  # Kibana Space ID
```

Ilustración 96: Configuración metricbeat Ubuntu – 1

```
#----- Elasticsearch output -----  
output.elasticsearch:  
  # Array of hosts to connect to.  
  hosts: ["192.168.0.29:9200"]  
  
  # Protocol - either `http` (default) or `https`.  
  #protocol: "https"  
  
  # Authentication credentials - either API key or username/password.  
  #api_key: "id:api_key"  
  #username: "elastic"  
  #password: "changeme"
```

Ilustración 97: Configuración metricbeat Ubuntu - 2

Se ejecutan los siguientes comandos para cargar los Dashboards y arrancar el servicio:

```
sudo metricbeat setup  
sudo service metricbeat start
```

Comprobamos desde Kibana que efectivamente ser reciben los datos:

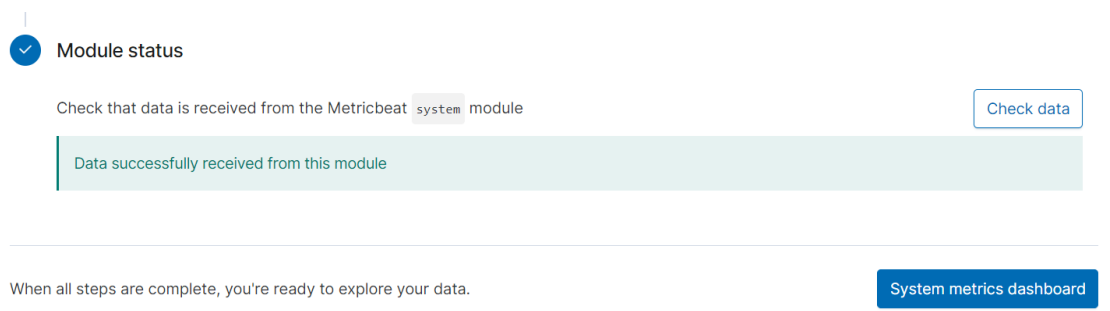


Ilustración 98: Metricbeat ubuntu verificación

Caso Windows

Los pasos a seguir son los siguientes:

- Se descomprime el fichero ZIP en *C:\Program Files*
- Se renombra la carpeta y se le llama *metricbeat*
- Se abre una PowerShell como Administrador y se ejecuta el siguiente comando:

```
.\install-service-  
metricbeat.ps1
```

- Se edita la configuración del fichero *metricbeat.yml* de la misma forma que en Caso Ubuntu.
- Se habilitan los módulos de *system* mediante el siguiente comando:

```
.\metricbeat.exe modules enable system
```

- Se cargan los Dashboards en Kibana y se arranca el servicio:

```
.\metricbeat.exe setup  
Start-Service metricbeat
```

```
PS C:\Program Files\metricbeat\metricbeat-7.6.0-windows-x86_64> .\metricbeat.exe modules enable system  
Module system is already enabled  
PS C:\Program Files\metricbeat\metricbeat-7.6.0-windows-x86_64> .\metricbeat.exe setup  
Overwriting ILM policy is disabled. Set `setup.ilm.overwrite:true` for enabling.  
Index setup finished.  
Loading dashboards (Kibana must be running and reachable)  
Loaded dashboards  
PS C:\Program Files\metricbeat\metricbeat-7.6.0-windows-x86_64> Start-Service metricbeat  
PS C:\Program Files\metricbeat\metricbeat-7.6.0-windows-x86_64>
```

Ilustración 99: Servicio metricbeat

Para corroborar que están llegando datos se verifica en Kibana:

The screenshot shows the Kibana 'Module status' page. At the top, there is a dropdown menu with 'Module status' selected. Below it, a message reads 'Check that data is received from the Metricbeat system module' with a 'Check data' button. A green notification bar below that states 'Data successfully received from this module'. At the bottom of the page, there is a message 'When all steps are complete, you're ready to explore your data.' and a blue button labeled 'System metrics dashboard'.

Ilustración 100: metricbeat Windows verificación

11.3 Añadir logs del sistema en Linux a Elasticsearch

Para recoger los logs del sistema en sistemas operativos Linux se utiliza filebeat. Para utilizarlo, hay que descargarlo de la página oficial de Elastic:

<https://www.elastic.co/es/downloads/beats/filebeat>

Una vez descargado se procede a su instalación:

Se instala el paquete DEB mediante el siguiente comando:

```
sudo dpkg -i filebeat-7.6.0-amd64.deb
```

Una vez instalado, se accede a la ruta `/etc/filebeat` para su configuración.

Se ha de modificar el fichero `filebeat.yml` con la siguiente información:

```
#===== Kibana =====
# Starting with Beats version 6.0.0, the dashboards are loaded via the Kibana AS
# This requires a Kibana endpoint configuration.
setup.kibana:
  # Kibana Host
  # Scheme and port can be left out and will be set to the default (http and 56$
  # In case you specify an additional path, the scheme is required: http://loc$
  # IPv6 addresses should always be defined as: https://[2001:db8::1]:5601
  host: "192.168.0.29:5601"
  # Kibana Space ID
  # ID of the Kibana Space into which the dashboards should be loaded. By defau$
  # the Default Space will be used.
  #space.id:
```

Ilustración 101: Configuración filebeat – 1

```
#----- Elasticsearch output -----
output.elasticsearch:
  # Array of hosts to connect to.
  hosts: ["192.168.0.29:9200"]
  # Protocol - either `http` (default) or `https`.
  #protocol: "https"
  # Authentication credentials - either API key or username/password.
  #api_key: "id:api_key"
  #username: "elastic"
  #password: "changeme"
```

Ilustración 102: Configuración filebeat - 2

Finalmente, se ejecutan los siguientes comandos para habilitar el módulo de sistema, cargar los Dashboards y arrancar el servicio:

```
sudo filebeat modules enable system
sudo filebeat setup
sudo service filebeat start
```

Se comprueba si se reciben datos:

Module status

Check that data is received from the Filebeat `system` module Check data

Data successfully received from this module

When all steps are complete, you're ready to explore your data. System logs dashboard

Ilustración 103: Verificación de instalación de filebeat

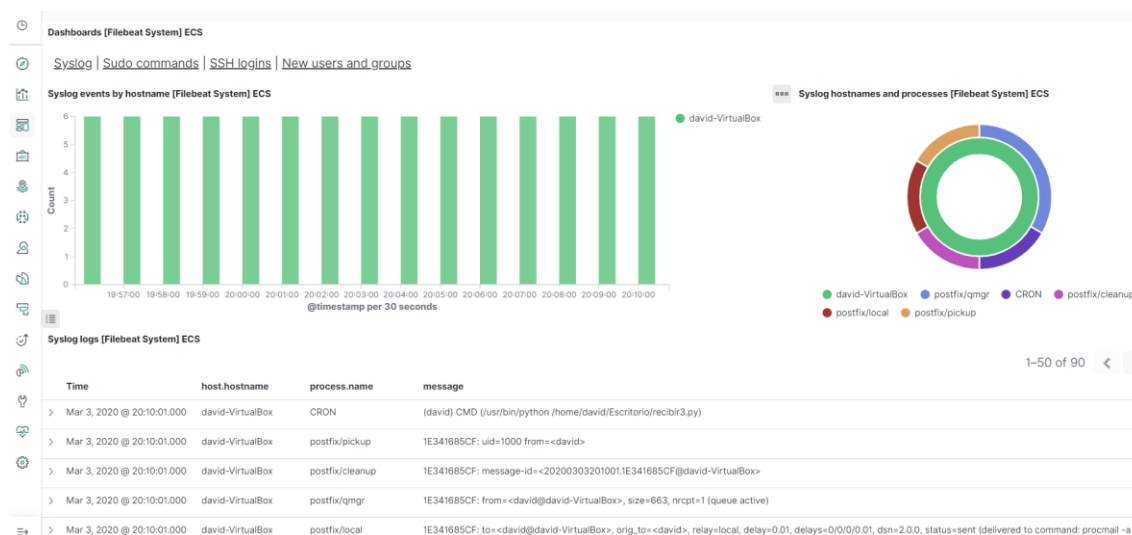


Ilustración 104: Análisis de logs de Linux con filebeats

11.4 Añadir Eventos de Windows a ElasticSearch

Para recoger los eventos del sistema en Windows e integrarlos con el SIEM se utiliza Winlogbeat. Para utilizarlo, hay que descargarlo de la página oficial de Elastic:

<https://www.elastic.co/es/downloads/beats/winlogbeat>

Los pasos a seguir para la instalación son los siguientes:

- Se descomprime el fichero ZIP en `C:\Program Files`
- Se renombra la carpeta y se le llama `winlogbeat`
- Se abre una PowerShell como Administrador y se ejecuta el siguiente comando:

```
.\install-service-winlogbeat.ps1
```

- Se edita la configuración del fichero `winlogbeat.yml` de la siguiente forma:


```

setup.kibana:
# Kibana Host
# Scheme and port can be left out and will be set to the default (http and 5601)
# In case you specify and additional path, the scheme is required: http://localhost:5601/path
# IPv6 addresses should always be defined as: https://[2001:db8::1]:5601
host: "192.168.0.29:5601"

# Kibana Space ID
# ID of the Kibana Space into which the dashboards should be loaded. By default,
# the Default Space will be used.
#space.id:

#----- Elastic Cloud -----
# These settings simplify using Winlogbeat with the Elastic Cloud (https://cloud.elastic.co/).

# The cloud.id setting overwrites the `output.elasticsearch.hosts` and
# `setup.kibana.host` options.
# You can find the `cloud.id` in the Elastic Cloud web UI.
#cloud.id:

# The cloud.auth setting overwrites the `output.elasticsearch.username` and
# `output.elasticsearch.password` settings. The format is `<user>:<pass>`.
#cloud.auth:

#----- Outputs -----
# Configure what output to use when sending the data collected by the beat.

#----- Elasticsearch output -----
output.elasticsearch:
# Array of hosts to connect to.
hosts: ["192.168.0.29:9200"]

# Protocol - either `http` (default) or `https`.
#protocol: "https"

# Authentication credentials - either API key or username/password.
#api_key: "id:api_key"
#username: "elastic"
#password: "changeme"

```

Ilustración 105: Configuración winlogbeat

- Se cargan los Dashboards en Kibana y se arranca el servicio:

```

.\winlogbeat.exe setup
Start-Service winlogbeat

```

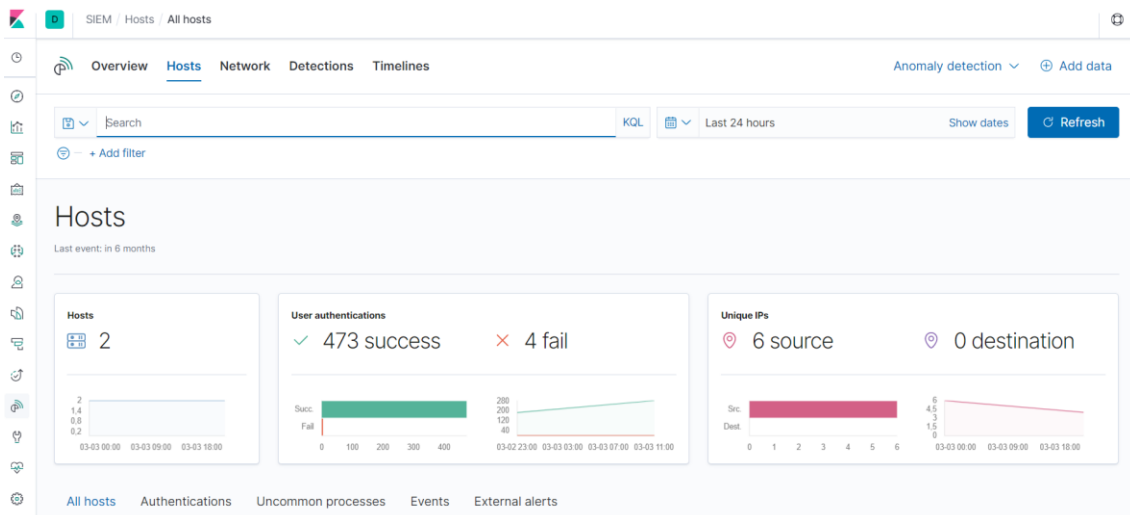


Ilustración 106: Verificación winlogbeat

11.5 Instalación y configuración de Heartbeat

Para verificar si la disponibilidad de una web se utiliza *heartbeat*. Para utilizarlo, hay que descargarlo de la página oficial de Elastic:

<https://www.elastic.co/es/downloads/beats/heartbeat>

Caso Ubuntu

Una vez descargado se procede a su instalación:
Se instala el paquete DEB mediante el siguiente comando:

```
sudo dpkg -i heartbeat-7.6.0-amd64.deb
```

Una vez instalado, se accede a la ruta `/etc/heartbeat` para su configuración.
Se ha de modificar el fichero `heartbeat.yml` con la siguiente información:

```
# Configure monitors inline
heartbeat.monitors:
- type: http

  # List or urls to query
  urls: ["http://192.168.0.29:9200"]

  # Configure task schedule
  schedule: '@every 10s'

  # Total test connection and data exchange timeout
  #timeout: 16s
```

Ilustración 107: Configuración heartbeat – 1

```
#===== Kibana =====
# Starting with Beats version 6.0.0, the dashboards are loaded via the Kibana AS
# This requires a Kibana endpoint configuration.
setup.kibana:

  # Kibana Host
  # Scheme and port can be left out and will be set to the default (http and 56$
  # In case you specify and additional path, the scheme is required: http://loc$
  # IPv6 addresses should always be defined as: https://[2001:db8::1]:5601
  host: "192.168.0.29:5601"

  # Kibana Space ID
  # ID of the Kibana Space into which the dashboards should be loaded. By defau$
  # the Default Space will be used.
  #space.id:
```

Ilustración 108: Configuración heartbeat – 2

```
#----- Elasticsearch output -----
output.elasticsearch:
  # Array of hosts to connect to.
  hosts: ["192.168.0.29:9200"]

  # Protocol - either `http` (default) or `https`.
  #protocol: "https"

  # Authentication credentials - either API key or username/password.
  #api_key: "id:api_key"
  #username: "elastic"
  #password: "changeme"
```

Ilustración 109: Configuración heartbeat - 3

En este caso, se verificará la disponibilidad de Elasticsearch.

Finalmente, se ejecutan los siguientes comandos para indexar el índice en Kibana y arrancar el servicio:

```
sudo heartbeat setup
sudo service heartbeat-elastic start
```

Caso Windows

Los pasos a seguir son los siguientes:

- Se descomprime el fichero ZIP en *C:\Program Files*
- Se renombra la carpeta y se le llama *heartbeat*
- Se abre una PowerShell como Administrador y se ejecuta el siguiente comando:

```
.\install-service-heartbeat.ps1
```

- Se edita la configuración del fichero *heartbeat.yml* de la misma forma que en Caso Ubuntu, exceptuando la URL a monitorizar. Utilizaremos `http://localhost:8080`
- Se cargan los índices en Kibana y se arranca el servicio:

```
.\heartbeat.exe setup
Start-Service heartbeat
```

11.6 Instalar y configurar AuditBeat

Para recoger datos de auditoría se utiliza *auditbeat*. Para utilizarlo, hay que descargarlo de la página oficial de Elastic:

<https://www.elastic.co/es/downloads/beats/auditbeat>

Caso Ubuntu

Una vez descargado se procede a su instalación:
Se instala el paquete DEB mediante el siguiente comando:

```
sudo dpkg -i auditbeat-7.6.0-amd64.deb
```

Una vez instalado, se accede a la ruta `/etc/auditbeat` para su configuración.
Se ha de modificar el fichero `auditbeat.yml` con la siguiente información:

```
#===== Kibana =====  
  
# Starting with Beats version 6.0.0, the dashboards are loaded via the Kibana AS  
# This requires a Kibana endpoint configuration.  
setup.kibana:  
  
  # Kibana Host  
  # Scheme and port can be left out and will be set to the default (http and 56$  
  # In case you specify and additional path, the scheme is required: http://loc$  
  # IPv6 addresses should always be defined as: https://[2001:db8::1]:5601  
  host: "192.168.0.29:5601"  
  
  # Kibana Space ID  
  # ID of the Kibana Space into which the dashboards should be loaded. By defau$  
  # the Default Space will be used.  
  #space.id:
```

Ilustración 110: Configuración auditbeat - 1

```
#----- Elasticsearch output -----  
output.elasticsearch:  
  # Array of hosts to connect to.  
  hosts: ["192.168.0.29:9200"]  
  
  # Protocol - either `http` (default) or `https`.  
  #protocol: "https"  
  
  # Authentication credentials - either API key or username/password.  
  #api_key: "id:api_key"  
  #username: "elastic"  
  #password: "changeme"
```

Ilustración 111: Configuración auditbeat - 2

En este caso, se verificará la disponibilidad de la URL de Elasticsearch.

Finalmente, se ejecutan los siguientes comandos para cargar los Dashboards en Kibana y arrancar el servicio:

```
sudo auditbeat setup  
sudo service auditbeat start
```

Caso Windows

Los pasos a seguir son los siguientes:

- Se descomprime el fichero ZIP en `C:\Program Files`
- Se renombra la carpeta y se le llama `auditbeat`

- Se abre una PowerShell como Administrador y se ejecuta el siguiente comando:

```
.\install-service-auditbeat.ps1
```

- Se edita la configuración del fichero *auditbeat.yml* de la misma forma que en [Caso Ubuntu](#).
- Se cargan los índices en Kibana y se arranca el servicio:

```
.\auditbeat.exe setup  
Start-Service auditbeat
```

11.7 Establecer las comunicaciones de Elasticsearch y Kibana mediante SSL/TLS

Para establecer las comunicaciones mediante SSL/TLS es necesario crear certificados. Para ello, se realizan los siguientes pasos:

- Crear los certificados mediante la herramienta *certutil* de Elastic:

```
bin/elasticsearch-certutil cert ca -pem --out  
/etc/elasticsearch/certs/certs.zip
```

- Se descomprimen los certificados:

```
[root@localhost certs]# ls  
ca certs-zip instance  
[root@localhost certs]#
```

Ilustración 112: Certificados SSL/TLS

- Se añaden las siguientes líneas al fichero de configuración *elasticsearch.yml*:

```
xpack.security.enabled: true  
xpack.security.http.ssl.enabled: true  
xpack.security.transport.ssl.enabled: true  
xpack.security.http.ssl.key: certs/instance/instance.key  
xpack.security.http.ssl.certificate:  
certs/instance/instance.crt  
xpack.security.http.ssl.certificate_authorities:  
certs/ca/ca.crt
```

- Se deben configurar las contraseña de los usuarios integrados mediante la herramienta *setup-passwords* de Elastic:

```
bin/elasticsearch-setup-  
passwords interactive
```

El comando permite interactuar con la consola e introducir manualmente las contraseñas deseadas para cada usuario.

- Finalmente se reinicia el servicio de Elasticsearch.
- Ahora se configurará Kibana. Se copian los mismos certificados al directorio de Kibana y se descomprimen.
- Se añaden las siguientes líneas al fichero de configuración kibana.yml:

```
elasticsearch.hosts: ["https://192.168.0.29:9200"]  
elasticsearch.username: "kibana"  
elasticsearch.password: "password"  
server.ssl.enabled: true  
server.ssl.certificate: /etc/kibana/certs/instance/instance.crt  
server.ssl.key: /etc/kibana/certs/instance/instance.key  
elasticsearch.ssl.certificateAuthorities:["/etc/kibana/certs/ca/ca.c  
rt"]  
elasticsearch.ssl.verificationMode: none  
xpack.encryptedSavedObjects.encryptionKey:  
'fhjskloppd678ehkdfdlliverpoolfcr'
```

- Finalmente, se reinicia el servicio de Kibana.
- Es necesario destacar que para que el resto de servicios sigan funcionando hay que configurarlos para utilizar SSL/TLS. Para ello, hay que añadir las siguientes líneas en los ficheros de configuración pertinentes de cada servicio:

```
elasticsearch.output:  
  elasticsearch.hosts:  
    ["192.168.0.29:9200"]  
  protocol: "https"  
  elasticsearch.username: "elastic"  
  elasticsearch.password: "password"  
  ssl.verification_mode: "none"
```

11.8 Instalar y configurar packetbeat

Para recoger datos de la red se utiliza *packetbeat*. Para utilizarlo, hay que descargarlo de la página oficial de Elastic:

<https://www.elastic.co/es/downloads/beats/packetbeat>

En este caso, solamente se utilizará Packetbeat en Windows. Los pasos a seguir son los siguientes:

- Se instala la librería de sniffing de paquetes, *npcap*.
- Se descarga el fichero ZIP de *packetbeat*.
- Se descomprime el fichero ZIP en *C:\Program Files*
- Se renombra la carpeta y se le llama *packetbeat*
- Se abre una PowerShell como Administrador y se ejecuta el siguiente comando:

```
.\install-service-packetbeat.ps1
```

- Se edita la configuración del fichero *packetbeat.yml*:

```
elasticsearch.output:  
  elasticsearch.hosts:  
    ["192.168.0.29:9200"]  
  protocol: "https"  
  elasticsearch.username:  
    "elastic"  
  elasticsearch.password:  
    "password"  
  ssl.verification_mode: "none"
```

- Se cargan los índices en Kibana y se arranca el servicio:

```
Packetbeat.exe setup  
Start-Service packetbeat
```

11.9 Creación de usuarios y roles en Kibana

Previamente a la creación de usuarios desde Kibana, es necesario crear los usuarios por defecto de Elastic desde el terminal de comandos.

El comando a utilizar para establecer manualmente las contraseñas es el siguiente:

```
bin/elasticsearch-setup-passwords interactive
```

El *interactive* permite establecer para cada usuario la contraseña, de forma interactive. También es posible utilizar la palabra *auto* en lugar de *interactive* para realizarlo de forma automática.

Una vez establecidos los usuarios, se puede acceder a Kibana desde el navegador e iniciar sesión con el usuario elastic y la contraseña definida anteriormente.

Desde el área de ajustes de Kibana, dentro del área Security se podrá acceder al panel de gestión de usuarios y roles.

| <input type="checkbox"/> Full Name ↑ | User Name | Email Address | Roles | Reserved |
|--|-----------|-----------------------|---|----------|
| <input type="checkbox"/> David Vazquez | user | davidvazquezp@uoc.edu | apm_system, apm_user, beats_admin, beats_system, data_frame_transforms_admin, data_frame_transforms_user, enrich_user, ingest_admin, kibana_admin, kibana_dashboard_only_user, kibana_system, kibana_user, logstash_admin, logstash_system, machine_learning_admin, machine_learning_user, monitoring_user, remote_monitoring_agent, remote_monitoring_collector, reporting_user, rollup_admin, rollup_user, snapshot_user, superuser, transform_admin, transform_user, transport_client, watcher_admin, watcher_user, SIEM_admin | |
| <input type="checkbox"/> | elastic | | superuser | ✓ |
| <input type="checkbox"/> | kibana | | kibana_system | ✓ |

Ilustración 113: Gestión de usuarios

Roles Create role

Apply roles to groups of users and manage permissions across the stack.

Q Search...

| <input type="checkbox"/> Role ↑ | Reserved | Actions |
|---|----------|---------|
| <input type="checkbox"/> SIEM_admin | | |
| <input checked="" type="checkbox"/> apm_system | ✓ | |
| <input checked="" type="checkbox"/> apm_user | ✓ | |
| <input checked="" type="checkbox"/> beats_admin | ✓ | |
| <input checked="" type="checkbox"/> beats_system | ✓ | |
| <input checked="" type="checkbox"/> data_frame_transforms_admin | ✓ | |
| <input checked="" type="checkbox"/> data_frame_transforms_user | ✓ | |
| <input checked="" type="checkbox"/> enrich_user | ✓ | |
| <input checked="" type="checkbox"/> ingest_admin | ✓ | |
| <input checked="" type="checkbox"/> kibana_admin | ✓ | |
| <input checked="" type="checkbox"/> kibana_dashboard_only_user | ✓ | |
| <input checked="" type="checkbox"/> kibana_system | ✓ | |

Ilustración 114: Gestión de roles

Para crear nuevos usuarios o roles, tan solo se deberá acceder al botón de “Create user” o “Create role” y establecer la configuración deseada.