Desplegar la herramienta "Zeek IDS" y su posterior explotación para el análisis de actividades sospechosas en la red

Ignacio Galván Vitas
Máster Universitario en Seguridad de las Tecnologías de la Información y de las Comunicaciones
Análisis de datos

Borja Guaita Pérez
Helena Rifà Pous

Junio 2020

Agenda de la presentación

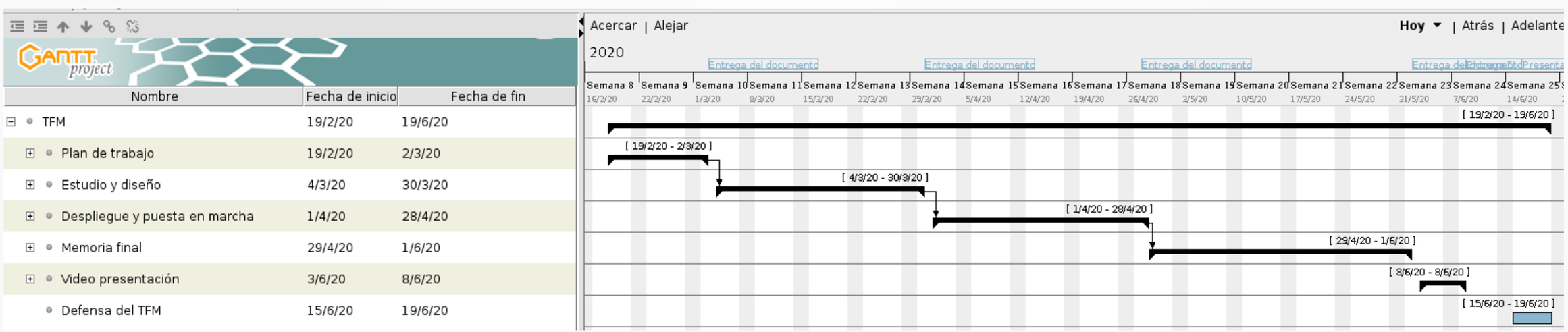
- Introducción.
- Plan de trabajo.
- Estudio y diseño.
- Despliegue.
- Resultados obtenidos.
- Demostración en vivo.
- Conclusiones.

Introducción – Objetivos

- Objetivo principal ⇒ Evaluar Zeek IDS combinado con ELK Stack para el análisis de actividades sospechosas en la red.
- Objetivos secundarios:
  - Despliegue automatizado ⇒ Ansible/Docker.
  - Software de fuentes abiertas.
  - Aprendizaje máquina integrado ELK.
  - Integrar listas de reputación de software malicioso.
  - Cuadros de mandos información relevante.
  - Conjunto de recomendaciones y mejoras.

Plan de trabajo – Fases

- Seis fases:
  - Definición del Plan de Trabajo.
  - Estudio y diseño de la solución.
  - Despliegue y puesta en marcha.
  - Análisis y evaluación de los datos obtenidos.
  - Video presentación.
  - Defensa del Trabajo de Fin de Máster.
- Total ⇒ 244 horas de trabajo.



Estudio y diseño – Hardware

- Un servidor físico:
  - Intel i3-2130.
  - 8 GB DDR3 1333 Mhz.
  - 2 TB almacenamiento SATA.
- Contratado con Kimsufi tipo KS-7.

Estudio y diseño – Software base

- Sistema operativo ⇒ CentOS 7.
- Heralding:
  - Honeypot baja interacción.
  - «Cebo» para aumentar interacción rastreo y conexiones.
- Zeek IDS ⇒ Analizador de red.
- HA Proxy:
  - Proporcionar autenticación a Kibana.
  - Cifrado comunicaciones mediante SSL.

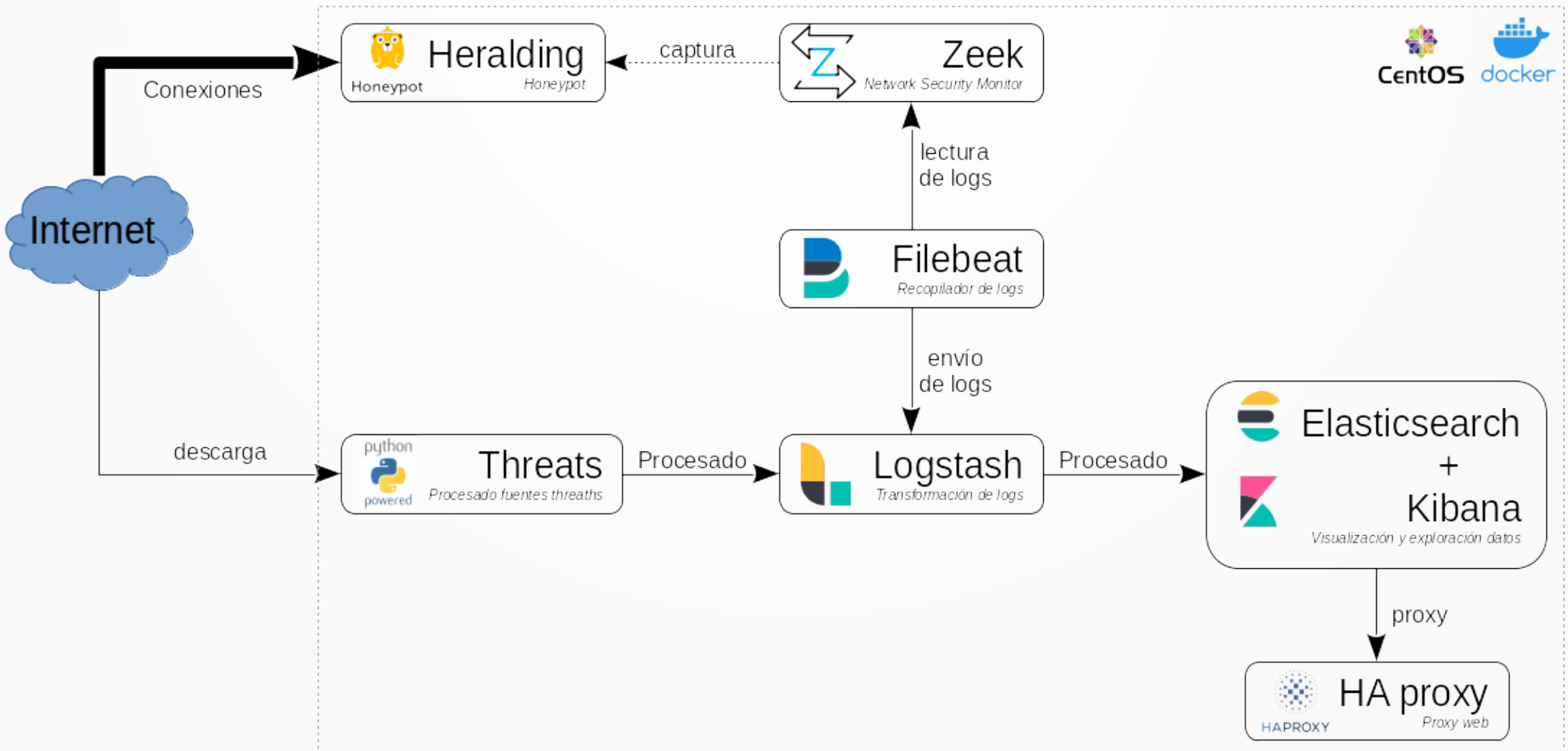
Estudio y diseño – Elastic Stack

- Filebeat ⇒ Reenvío de registros desde Zeek a Logstash.
- Logstash:
  - Centralización de registros.
  - Transformación de datos.
  - Envío a Elasticsearch.
- Elasticsearch:
  - Almacén información.
  - Búsqueda y análisis en tiempo real.
- Kibana ⇒ Interfaz web para usuarios.

Estudio y diseño – Fuentes información amenazas

- Origen abierto basadas en listas negras IP.
- Fuentes elegidas:
  - Blocklist.de ⇒ Especialista fraude. Servidores propios.
  - BadIPs.com ⇒ Servicio comunitario. Recopilación IP.
  - Turris.cz ⇒ Enrutadores Turris.
  - Dan.Me.uk ⇒ Proyecto individual. Nodos salida red Tor.
  - Malware Bazaar ⇒Abuse.ch. Lista «hashes» ficheros malware.
- Procesado con scripts en Python a CSV.
- Integración con Telegram ⇒ Avisos ejecución.

Estudio y diseño - Diagrama de la solución



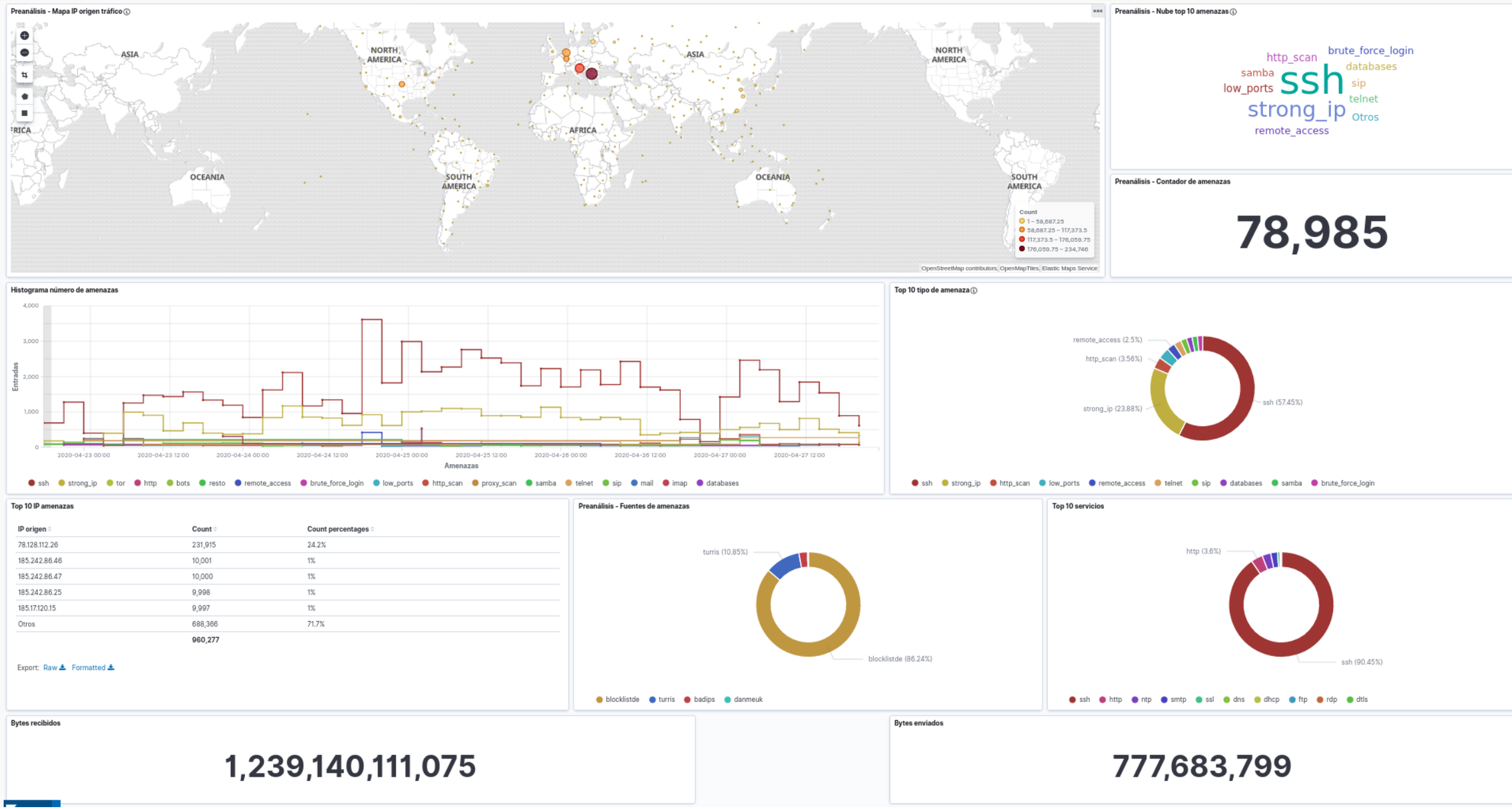
Despliegue – Docker y Ansible

- Objetivos:
  - Independizar software del entorno de ejecución.
  - Facilitar el despliegue y puesta en marcha.
  - Reproducible.
  - Rapidez ⇒ En quince minutos.
- Detalle de la implementación ⇒ Memoria TFM.

Despliegue – Pruebas y captura preliminar de datos

- Pruebas de integración ⇒ Verificación del funcionamiento de los componentes.
- Captura y análisis preliminar:
  - Período de siete días.
  - Poco tiempo ⇒ Información probablemente sesgada.
  - Cuadro de mandos especifico.

Despliegue – Cuadro de mandos preliminar



Análisis y evaluación datos – Cuadros de mandos

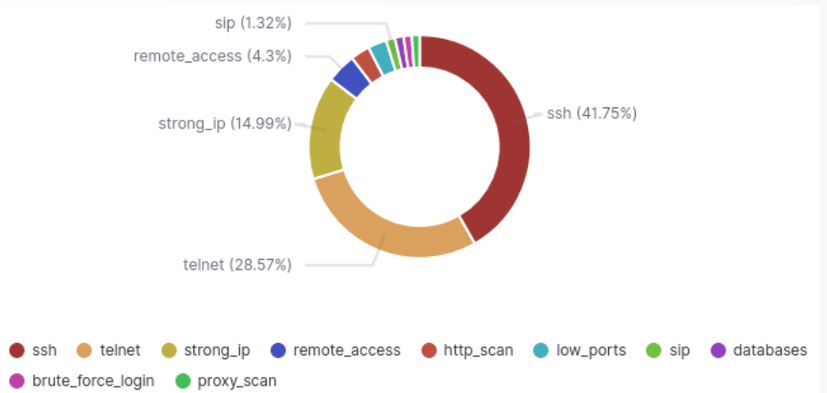
- Diez días adicionales de captura de datos.
- Cuadros de mandos:
  - 1 cuadro general.
  - 17 cuadros de mandos específicos:

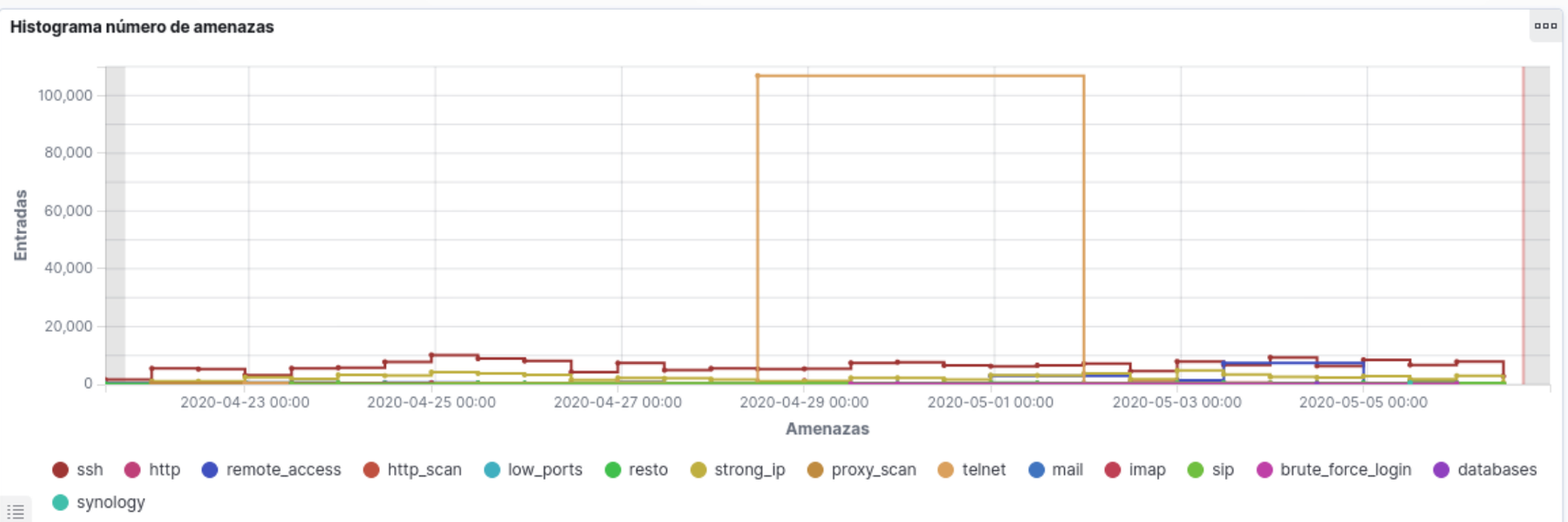
DHCP	DNS	DPD	FILES	HTTP	Kerberos
NTP	RDP	SIP	SMTP	SSH	SSL
Syslog	Túneles	Tráfico extraño		Certificados X.509	

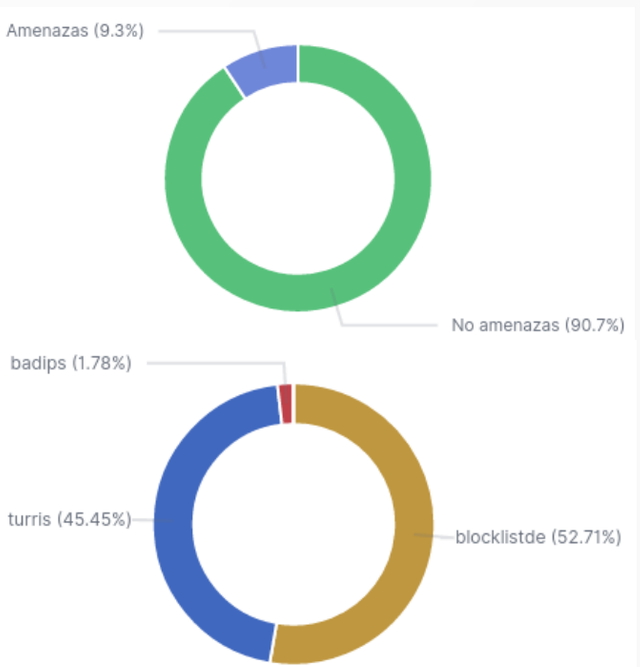
Análisis y evaluación datos – Resultados obtenidos

- Conexiones:
  - 3.307.045 conexiones, 336.565 etiquetadas como amenaza. Media: 22.437 amenazas diarias.
- Países:
  - China (33,56%), España (32,28%) Francia (7,86%), EE.UU (6,1%) y Rusia (1,83%). Resto: 18,37%.
- Tráfico:
  - 1,27 TB entrante y 2,76 GB saliente.
- Ratio de amenazas:
  - 9,3% etiquetadas del total analizado.
  - Por tipo de amenaza: SSH (43,49%), Telnet (25,81%), StrongIP (16,49%), RDP (4,02%). Resto: 10,19%.
- Ratio por fuente:
  - Blocklist.de (52,71%) , Turris (45,45%) , BadIPs (1,78%), DanMe (0,06%) y Malware Bazaar (0%).

Análisis y evaluación datos – Resultados obtenidos







Demostración en tiempo real

Conclusiones - Resumen

- Es viable disponer de un SIEM:
  - Basado en Zeek IDS y ELK Stack.
  - Basado en software de fuentes abiertas.
  - Fuentes de información sobre amenazas integradas.
- Dificultades:
  - La calidad de las fuentes de información abiertas.
  - Esfuerzo para mantenerlo actualizado y operativo.
- Sorpresa: Alto número de peticiones a lo largo del tiempo.

Conclusiones – Revisión objetivos

- Objetivo principal cumplido. ✓
- Objetivos secundarios:
  - Despliegue automatizado ⇒ Ansible/Docker. ✓
  - Software de fuentes abiertas. ✓
  - Aprendizaje máquina integrado ELK. ✗
  - Integrar listas de reputación de software malicioso. ✓
  - Cuadros de mandos información relevante. ✓
  - Conjunto de recomendaciones y mejoras. ✓

Conclusiones – Líneas futuras trabajo

- Evaluar nuevas fuentes de inteligencia de amenazas.
- Aprendizaje máquina compatible software fuentes abiertas.
- Sustitución «honeypot» por uno de alta interacción.
- Distribución de varios «honeypot» en diferentes ubicaciones en el mundo.
- Evaluar Graylog como sustituto de ELK Stack.

FIN