José Juan López Abellán
Oficina de Seguridad
Nuria Prieto Pinedo
CSIRT-SOC
Rafael Calzada Pradas.
Resp. de Seguridad de la Información

cert@uc3m.es

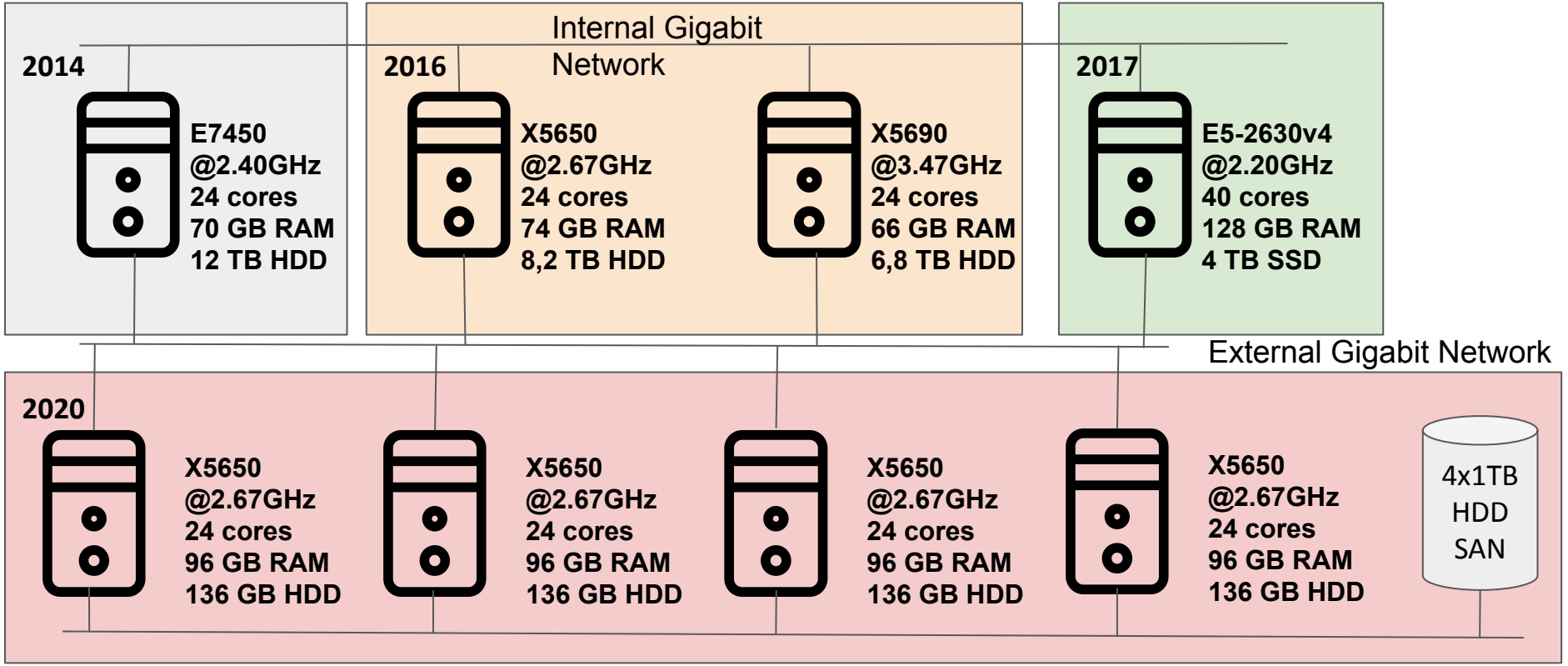
Migración del SIEM: De Elasticsearch en nodos físicos a OpenSearch sobre contenedores Docker en modo swarm

Agenda

- Situación inicial
  - Descripción
  - Problemas detectados
  - Parches
- Migración
  - Opensearch
  - Docker
- Conclusiones
  - Ventajas e inconvenientes

Situación inicial: Descripción

TLP:GREEN



Descripción: Problemas detectados

TLP:GREEN

- Nodos no homogéneos:
  - Capacidad de proceso
  - Memoria disponible
  - Capacidad y tipo de almacenamiento
  - **Consecuencias**:
    - Nodos saturados en CPU y/o disco
    - Nodos ociosos
    - No se pueden establecer roles de gestión en exclusiva
- Instalación en nodos físicos
  - Actualizaciones complejas:
    - Se trata de un sistema de streaming
    - Dependencias de plugins no incluidos en el paquete estándar
    - Dificultad para volver al estado anterior
  - **Consecuencias**:
    - No se actualiza de forma regular

Descripción: Parches

- Distribución de carga de trabajo:
  - Limitación de espacio de almacenamiento basado en porcentaje de ocupación en disco
  - División de la ingesta de eventos en varios procesos de logstash
  - Gestión de indices para eliminar aquellos que ya no sean relevantes
- Actualizaciones
  - Ventana de trabajo en los meses de verano

Migración: Opensearch: Un poco de historia

- Elastic Stack (Elastic, Logstash y Kibana)
  - Basada en Apache Lucene
  - Desarrollada y soportada por la empresa Elastic
  - Diferentes suscripciones: Ciertas funcionalidades incluidas solo en las suscripciones de pago:
    - Autenticación SAML/LDAP
    - Seguridad a nivel de campo y documento
    - Enmascaramiento de datos/pseudonimización
  - Licenciamiento basado en número de nodos
    - Muy caro para pocos nodos
    - Asumible para clústers de varias decenas de nodos
    - Empiezan a moverse a pago por uso, en la nube prinicipalmente
- Más información en:
  - https://www.elastic.co/es/subscriptions

Migración: Historia, OpenDistro y OpenSearch

- OpenDistro:
  - Fusión de varios repositorios de Elasticsearch y Searchguard
  - Problemas de autoría:
    - 2019: https://www.elastic.co/es/blog/dear-search-guard-users
    - Resueltos en 2022: https://www.elastic.co/es/blog/elastic-and-amazon-reach-agreement-on-trademark-infringement-lawsuit
- Opensearch:
  - Fork de Elasticsearch 7.10.2 y Kibana 7.10.2
  - Excluye X-Pack
  - Problemas con algunos Beats, que requierían X-Pack
  - Se puede migrar de logstash a Fluentd
  - Kibana pasa a ser Opensearch-Dashboards
  - Más información en https://www.opensearch.org/ y https://aws.amazon.com/es/what-is/opensearch

Suscripciones de Elastic: Principales carencias de Free

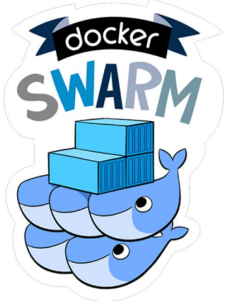
	Free	Platino	Enterprise
Gestión centralizada de pipelines logstash Logs de auditoria, filtrado IP, autenticación LDAP, PKI Alertas geofencing/agregaciones geométricas Informes en PDF/PNG	🚫	👍	👍
Replicación entre clústers Autenticación SAML, Seguridad a nivel de campo y documento Machine Learning Conector JDBC, ODBC y Tableau Exploración de grafos Detección de anomalías/Pronóstico series temporales	🚫	👍	👍

Más información de suscripciones: https://www.elastic.co/es/subscriptions

Elasticsearch Free vs Opensearch

	Free	Opensearch
Gestión centralizada de pipelines logstash	🚫	🚫
Logs de auditoria, filtrado IP, autenticación LDAP, PKI	🚫	👍
Alertas geofencing/agregaciones geométricas	🚫	?
Informes en PDF/PNG	🚫	👍
Replicación entre clústers	🚫	👍
Autenticación SAML, Seguridad a nivel de campo y documento	🚫	👍
Machine Learning	🚫	👍 (RCF y K-NN)
Conector JDBC, ODBC y Tableau	🚫	👍 (JDBC/ODBC)
Exploración de grafos	🚫	🚫
Detección de anomalías/Pronóstico series temporales	🚫	?

Contenedores: Docker, Docker Swarm mode y Kubernetes



¿Cómo migrar teniendo en cuenta las incompatibilidades y sin dejar de dar servicio?

- Contenedores Docker
- Elegimos Docker Swarm mode
  - Es más sencillo de administrar que kubernetes
  - Permite gestionar un cluster de nodos docker
  - No es tan flexible como kubernetes

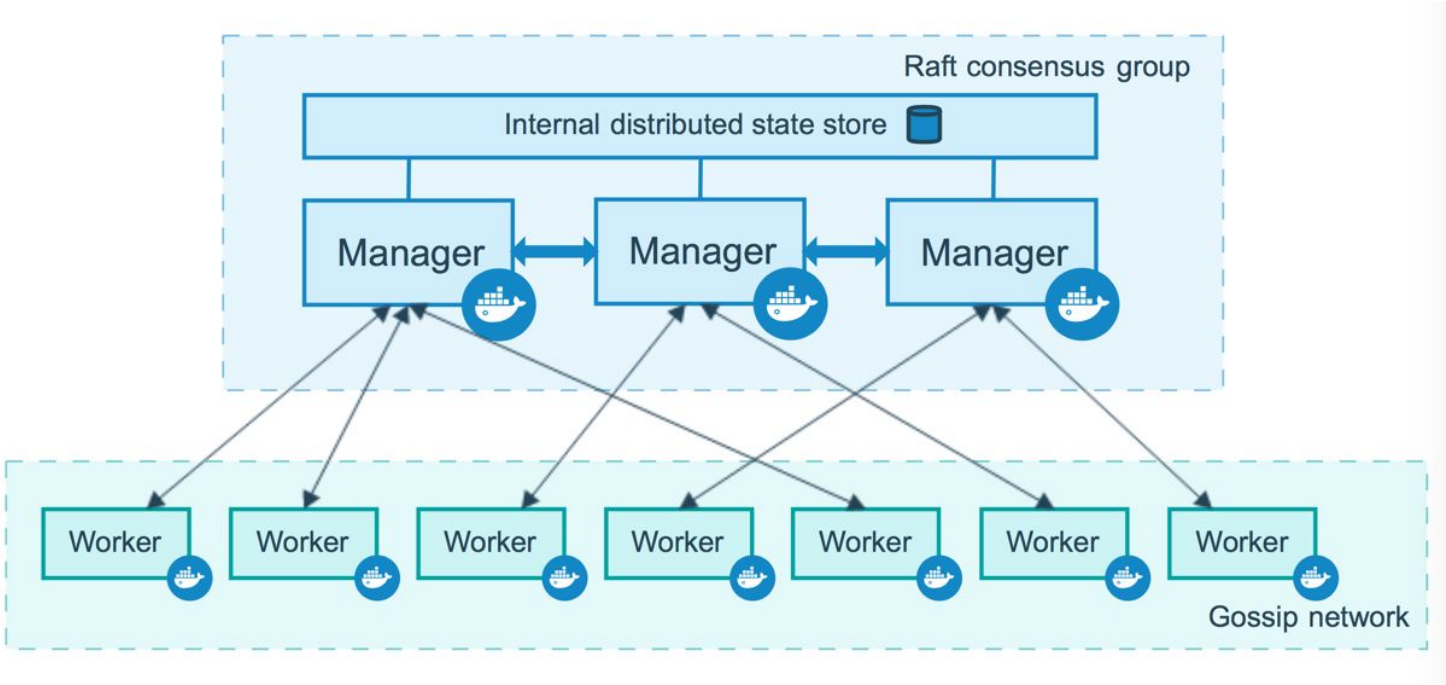
https://www.ibm.com/blog/docker-swarm-vs-kubernetes-a-comparison/

Docker swarm: Conceptos básicos

- Roles existentes:
  - Manager: Gestión del clúster
  - Worker: Ejecutan tasks
- Task: Contenedor en ejecución en un nodo del clúster
- Service: Definición de task a ejecutar, incluyendo el número de instancias.
  - Globales: Se ejecuta una task en cada nodo
  - Replicados: Se ejecutan tantas tasks como se indique en la configuración
  - Se emplea un servicio de DNS interno para balancear la carga
- Stack: Conjunto de servicios relacionados
- Exposed Ports:
  - Puntos de acceso a los servicios
  - Disponibles en cualquier nodo worker del clúster
- Network overlay:
  - Similar a Docker stand-alone.
  - Se pueden definir redes y asociar servicios a ellas.
  - Uso intensivo de NAT

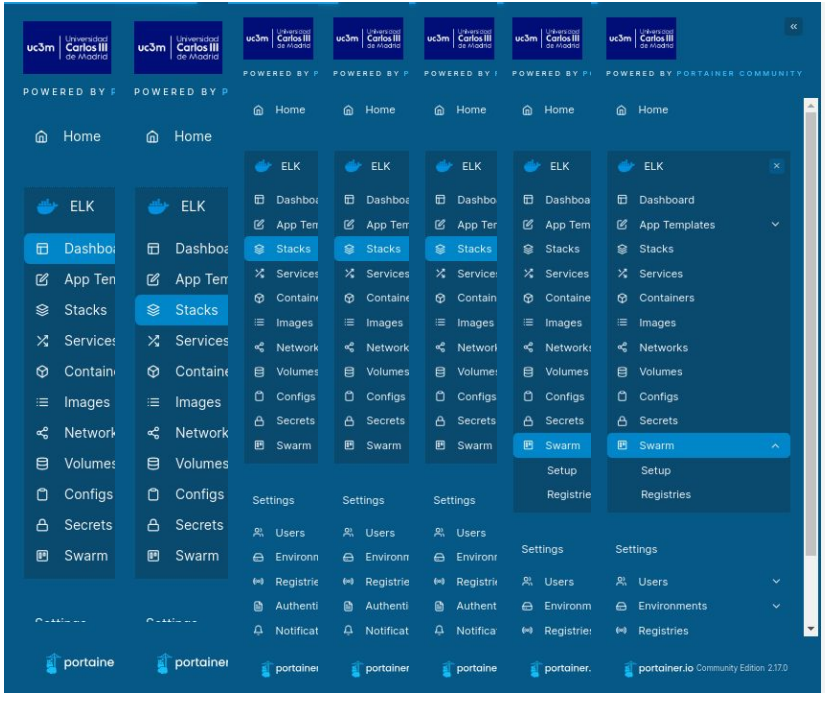
Docker swarm: arquitectura

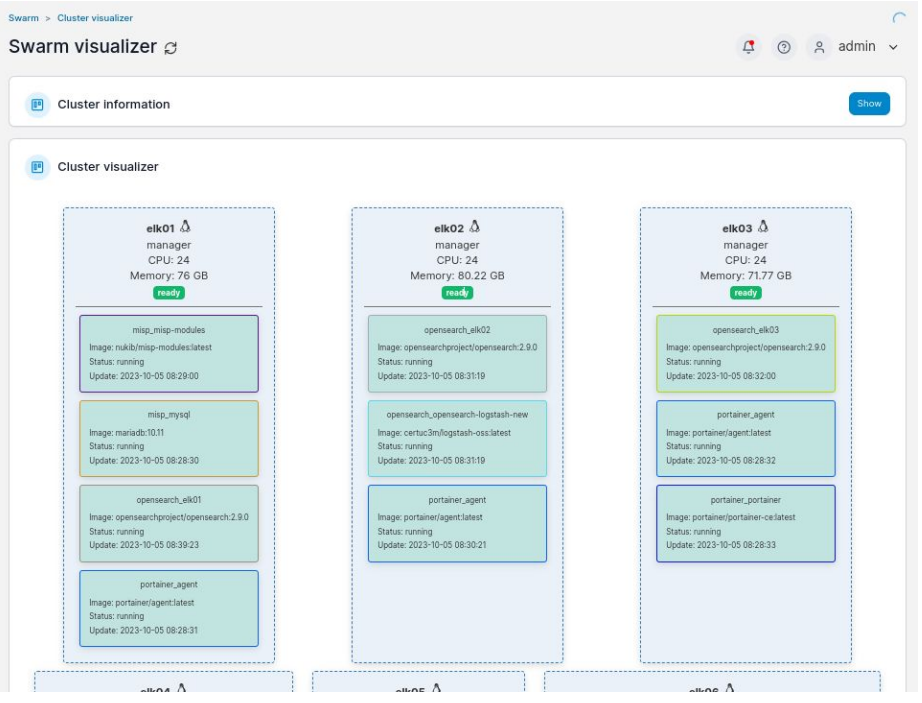
https://docs.docker.com/engine/swarm/how-swarm-mode-works/nodes/



Migración: Gestión de Contenedores

- Portainer-CE (https://www.portainer.io)





Migración: Imágenes de contenedores

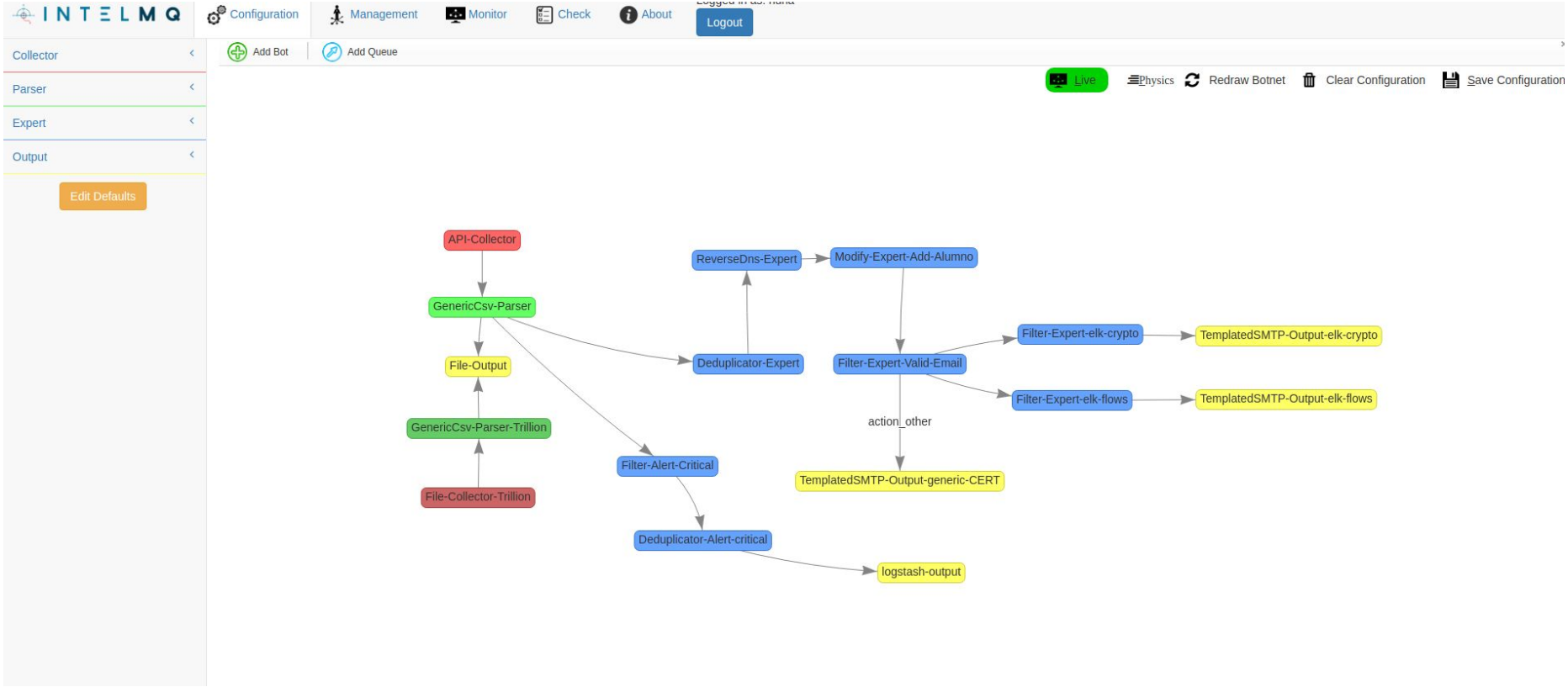
- opensearch: imagen oficial de opensearch
- logstash: imagen propia, para incluir plugins no estándar
  - filter-json-encode
  - filter-opensearch
  - output-google-bigquery
  - output-syslog
- opensearch-dashboards: imagen oficial de opensearch
- python: imagen propia, para ejecutar un script propio
- apache: imagen oficial, para servir ficheros estáticos
- intelmq: imagen propia
  - incluyendo algunas librerías python

IntelMQ

- Desarrollado por CERT Austria para CERTs/CSIRTs/SOCs.
- Respaldo de la Unión Europea, ENISA. IHAP Proyecto de automatización de gestión de incidentes.
- Recopila y procesa feeds mediante un protocolo de cola de mensajes.
- Métodos de instalación: paquetes nativos, Docker, paquete python y desde GitHub.
- Varios tipos de bots:
  - Collector Bots: MISP, Shodan, Twitter, ShadowServer, ESET, Microsoft Azure, fichero, Rsync, etc.
  - Parser Bots: MISP, Have I Been Pwned, CSV,ZonaH, etc.
  - Expert Bots: BlackHole,Sieve, Splunk, Tor Nodes, MaxMind, LookyLoo, etc.
  - Output Bots: SMTP, REST API, Postgre SQL, SQLite, MISP feed,file, etc.

IntelMQ

TLP:GREEN



Conclusiones: Ventajas e inconvenientes

- Ventajas:
  - Procedimiento de actualización más sencilla
    - Aproximadamente 1 hora en actualizar los 8 nodos sin dejar de prestar servicio.
  - Mejor uso de los nodos
    - Posibilidad de desplegar instancias con diferentes roles en un mismo nodo
    - Escalabilidad gestionada por Docker Swarm
    - Posibilidad de emplear los nodos para otros servicios *relacionados*, p.e. MISP
- Inconvenientes:
  - Curva de aprendizaje de Docker Swarm, por cambio de paradigma

uc3m | Universidad **Carlos III** de Madrid

Servicio de Informática y Comunicaciones

web: sdic.uc3m.es
twitter: @sdic_uc3m