ELK STACK + DOCUMENTACIÓN INTEGRADA (OBSERVABILITY)

Presenta: Axel Bautista y Emanuel López

HERRAMIENTAS ASIGNADAS

ELK Stack: Elasticsearch, Logstash/Fluentd, Kibana

Documentación Integrada: Manual de seguridad del cluster completo

OBJETIVOS DEL PROYECTO

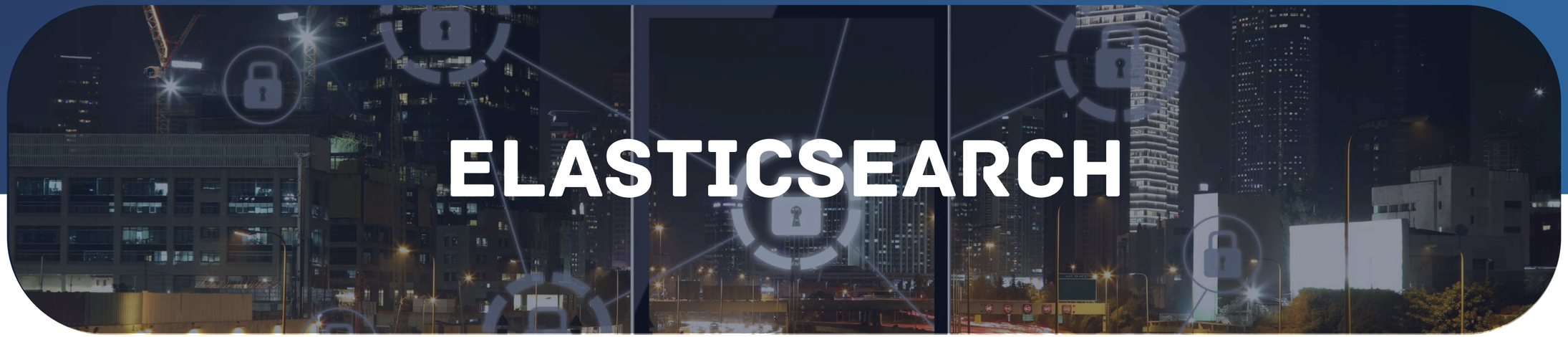
1. Desplegar ELK Stack en Kubernetes.
2. Centralizar logs de todos los equipos.
3. Crear 5+ dashboards de seguridad en Kibana.
4. Documentar todos los procesos de seguridad del cluster.
5. Crear manual de operaciones de seguridad.

1. ELK STACK

ELK Stack es el **acrónimo** de tres herramientas integradas: **Elasticsearch**, **Logstash** (o Fluentd como alternativa) y **Kibana**.

Con ellas se construye una **plataforma** de **observabilidad**, capaz de **recolectar**, **procesar**, **almacenar**, **analizar** y **visualizar** grandes volúmenes de datos principalmente **logs**, **métricas** y **eventos de seguridad**.

El **propósito** fundamental del ELK Stack es **centralizar** la **información** de **todos** los **sistemas**, permitiendo detectar anomalías, correlacionar eventos, crear alertas y generar reportes visuales en tiempo real.



ELASTICSEARCH

1. ¿Qué es?
2. ¿Para qué sirve?
3. ¿Cómo funciona?

ELASTICSEARCH

1.

Elasticsearch es un **motor de búsqueda y análisis de datos** distribuido, desarrollado originalmente por Elastic.co. Está basado en **Apache Lucene**, una biblioteca de indexación de texto de alto rendimiento. En el contexto de seguridad, actúa como el **repositorio central** donde se almacenan los **logs** y **eventos procesados**.

ELASTICSEARCH

2.

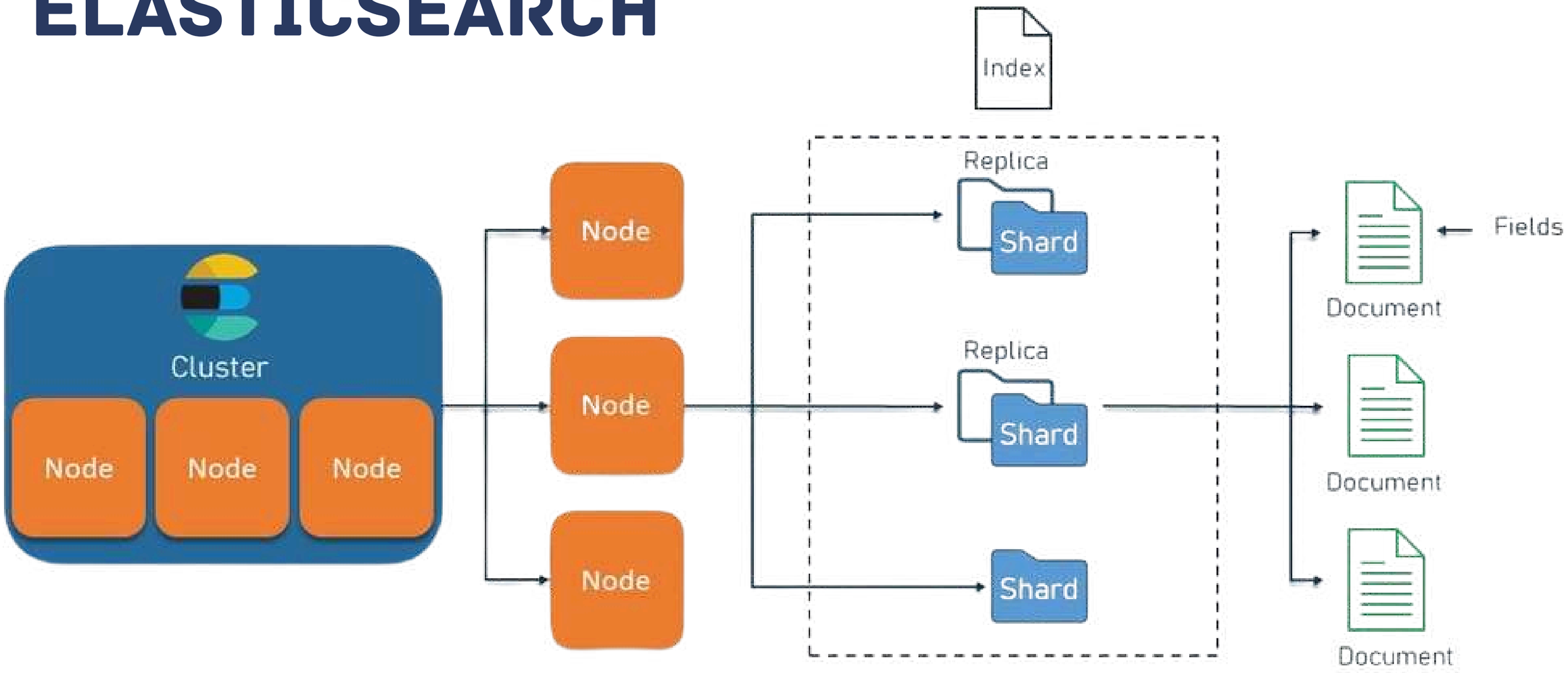
- Almacenar **logs estructurados y no estructurados** provenientes de múltiples fuentes.
- **Indexar** la información para hacerla **consultable** en **milisegundos**.
- Permitir **búsquedas complejas** (por campo, texto libre, rangos de tiempo, filtros, etc.).
- Soportar **agregaciones** y **análisis estadísticos** para dashboards de seguridad.
- Integrar políticas de **retención** y **rotación** de **datos** (ILM), garantizando eficiencia y cumplimiento normativo.

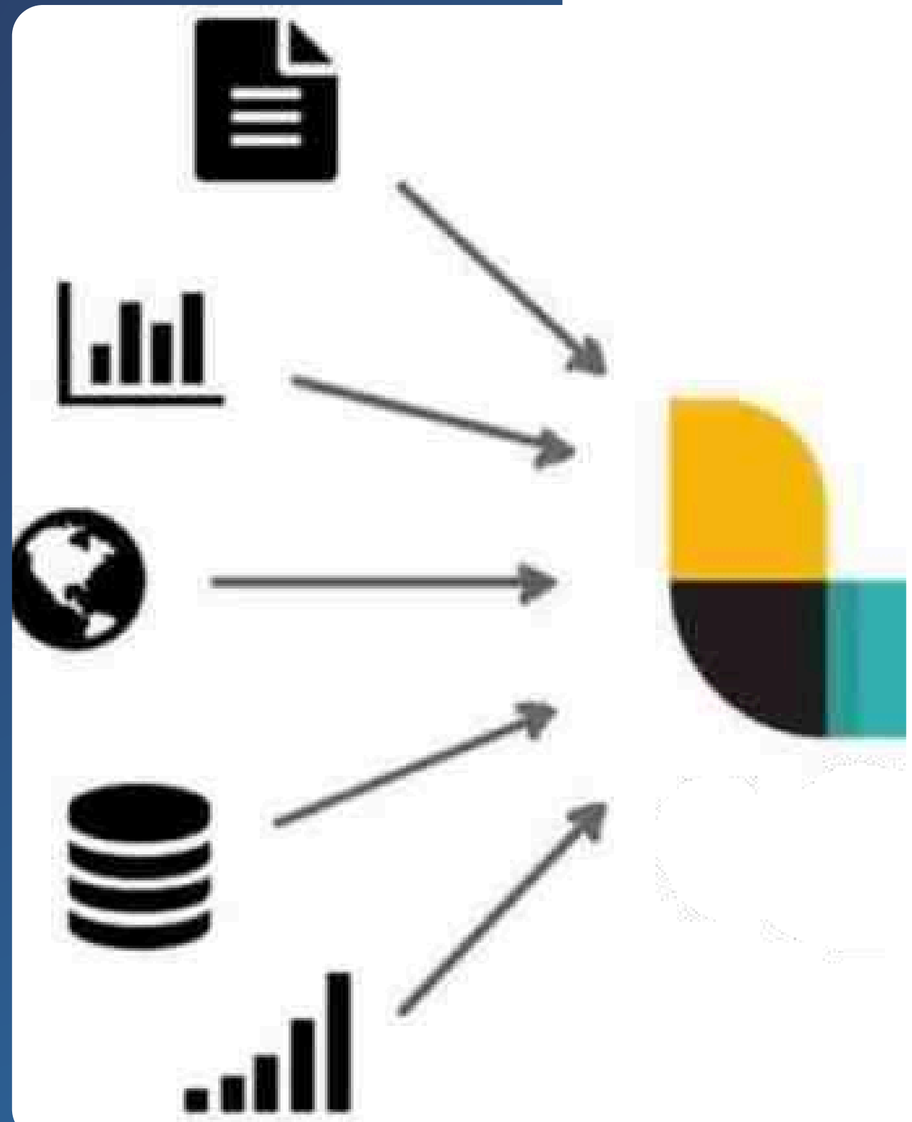
ELASTICSEARCH

3.

1. **Ingesta**: recibe datos de Logstash o Fluentd a través de una API REST o plugin nativo.
2. **Indexación**: organiza los documentos en índices, donde cada índice se divide en shards (fragmentos) distribuidos entre nodos.
3. **Búsqueda y agregación**: cuando un usuario o Kibana consulta, Elasticsearch utiliza índices invertidos (estructuras de Lucene) para localizar rápidamente coincidencias.
4. **Alta disponibilidad**: en modo cluster, replica los shards entre nodos para garantizar tolerancia a fallos.

ELASTICSEARCH





LOGSTASH / FLUENTD

Estas dos herramientas cumplen una función similar en el ecosistema ELK: **ingestar**, **transformar** y **enrutar** los **logs** desde su origen hasta **Elasticsearch**. Pueden usarse indistintamente, aunque cada una tiene su propio enfoque.

LOGSTASH

1. ¿Qué es?
2. ¿Para qué sirve?
3. ¿Cómo funciona?

LOGSTASH

1.

Logstash es una herramienta de **procesamiento de datos** desarrollada también por Elastic. Su **objetivo** es **recibir datos de múltiples fuentes**, **transformarlos** mediante filtros y enviarlos a destinos como Elasticsearch, bases de datos o archivos.

LOGSTASH

2.

- **Centralizar** logs de servidores, contenedores y aplicaciones.
- **Estandarizar** el **formato** de los mensajes (por ejemplo, convertir texto plano a JSON).
- **Filtrar** información sensible antes de almacenarla.
- **Aplicar transformaciones**: añadir campos, normalizar IPs, parsear timestamps, etc.
- **Redirigir** los datos a diferentes **outputs** (Elasticsearch, Kafka, S3...).

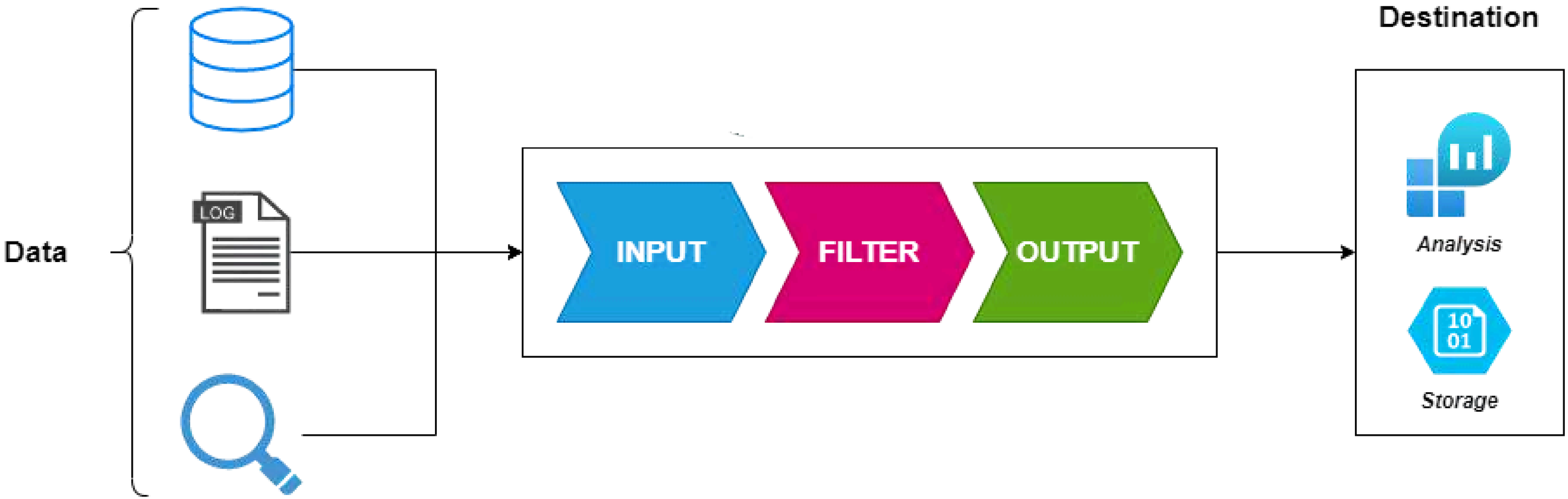
LOGSTASH

3.

Logstash se basa en un **pipeline de tres etapas**:

1. **Input**: define de dónde **provienen** los datos (archivos, syslog, beats, APIs).
2. **Filter**: aplica **transformaciones** (grok, mutate, date, geoip, etc.).
3. **Output**: envía los **resultados** a Elasticsearch u otro destino.

LOGSTASH


Data
LOG
INPUT
FILTER
OUTPUT
Destination
Analysis
10
01
Storage

FLUENTD

1. ¿Qué es?
2. ¿Para qué sirve?
3. ¿Cómo funciona?

FLUENTD

1.

Fluentd es un **colector de datos de código abierto** diseñado por Treasure Data (ahora parte de CNCF — Cloud Native Computing Foundation). Cumple el mismo rol que Logstash, pero está más **optimizado** para **entornos Kubernetes** y **contenedores**.

FLUENTD

2.

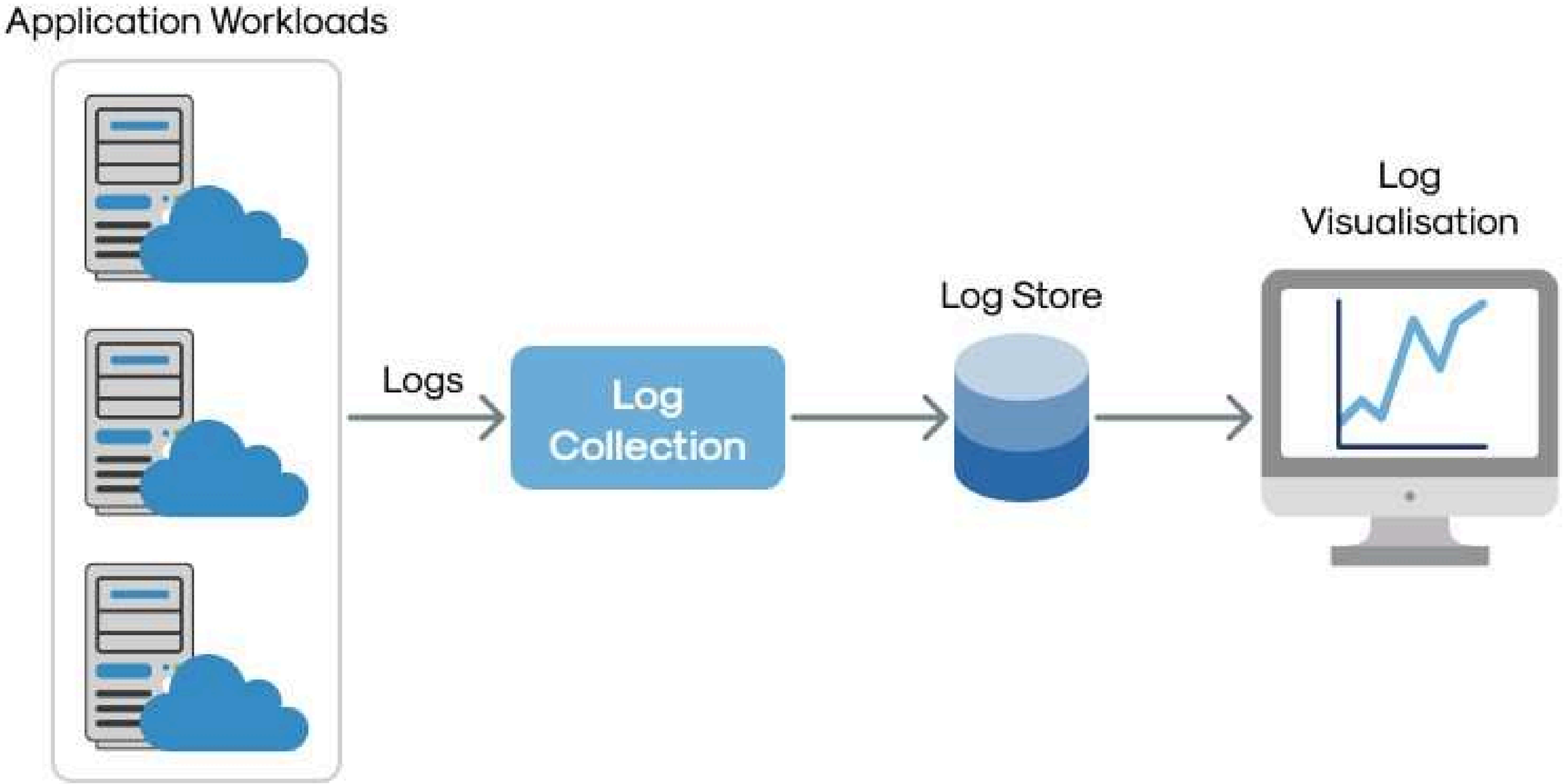
- **Recolectar logs** de cada **pod** o **nodo** del **cluster**.
- **Unificar** logs del sistema, del kernel, de Falco, OPA, API Server, etc.
- **Enviar** datos a **Elasticsearch** en formato **JSON**.
- Actuar como **DaemonSet** (un pod por nodo) para **capturar automáticamente** todos los **logs** del sistema.

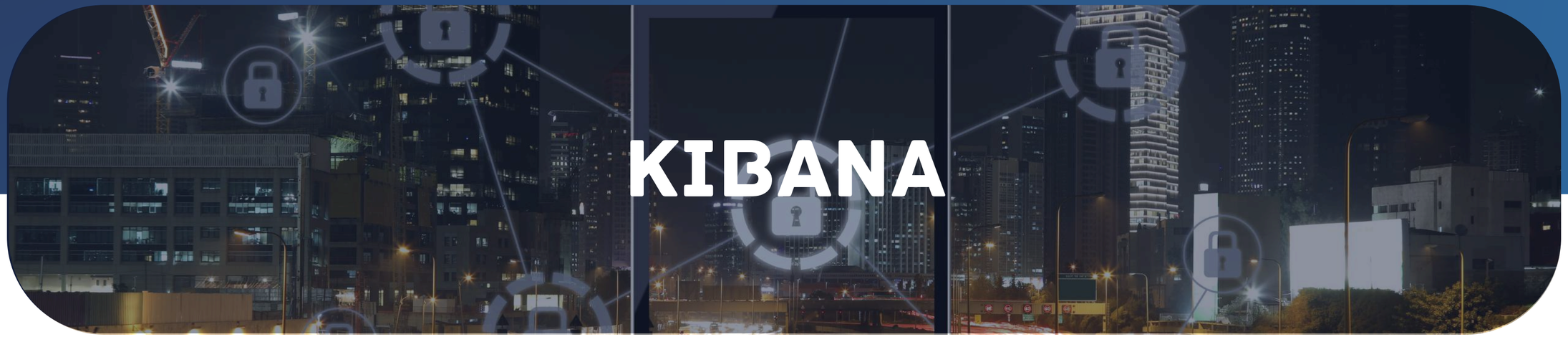
FLUENTD

3.

1. **Input plugins:** definen qué fuentes se leen (archivos, systemd, TCP/UDP, etc.).
2. **Filter plugins:** manipulan los datos (añaden etiquetas, cambian formato).
3. **Output plugins:** especifican el destino final (por ejemplo, out_elasticsearch).
4. **Buffer**: Fluentd puede almacenar temporalmente los logs si hay cortes de conexión con el servidor.

FLUENTD


Application Workloads
Logs
Log Collection
Log Store
Log Visualisation



KIBANA

1. ¿Qué es?
2. ¿Para qué sirve?
3. ¿Cómo funciona?

KIBANA

1.

Kibana es la **interfaz gráfica** y de **análisis** del ELK Stack. Es una **aplicación web** que permite **visualizar**, **explorar** y **gestionar** los datos almacenados en **Elasticsearch**.

KIBANA

2.

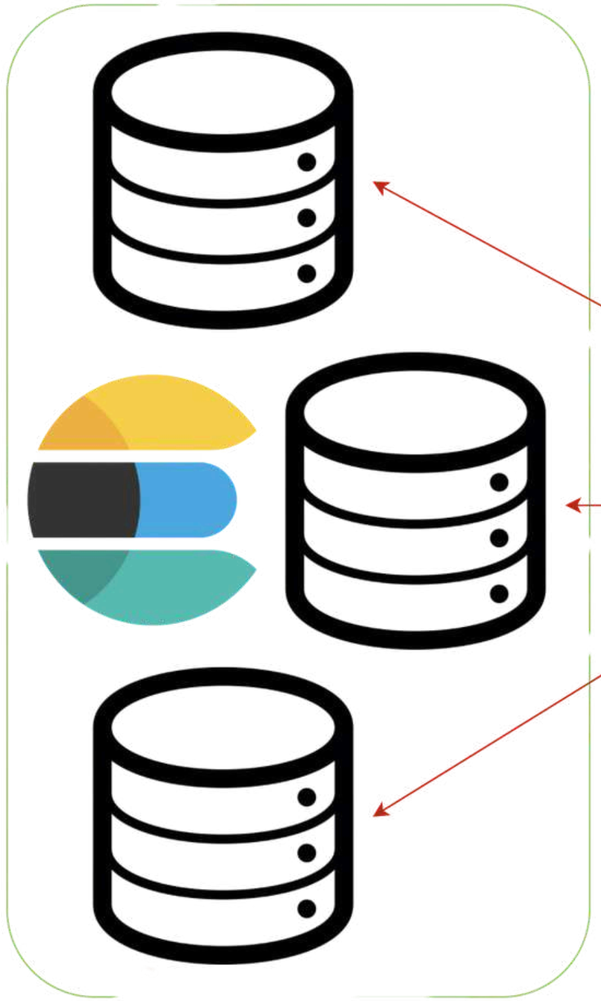
- Crear **dashboards interactivos** de **seguridad**, **rendimiento** y **auditoría**.
- **Buscar** y **filtrar** logs **específicos** con lenguaje KQL (Kibana Query Language).
- **Configurar alertas** y **notificaciones** basadas en condiciones (por ejemplo, “más de 10 fallos de autenticación en 5 minutos”).
- **Explorar anomalías** y **correlaciones** entre distintos índices.
- Exportar **reportes** o **visualizaciones** para **auditorías** o presentaciones.

KIBANA

3.

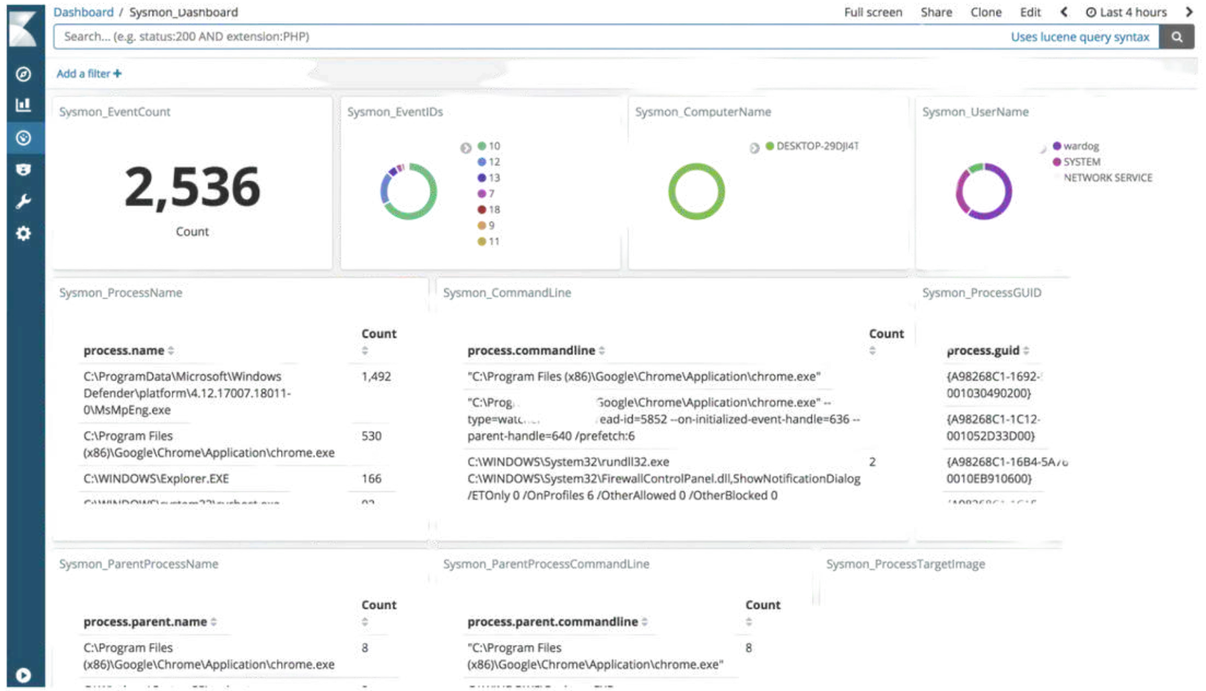
1. Se **conecta** a Elasticsearch mediante su **API REST.**
2. **Consulta** índices, **ejecuta** agregaciones y **muestra** resultados **visualmente** (gráficas, tablas, mapas, cronogramas).
3. Permite **crear** Saved Searches, Visualizations y Dashboards.
4. También **gestiona usuarios**, **roles** y **políticas** si la seguridad de Elastic está habilitada (xPack).

KIBANA



Dashboards

- Global Dashboard
- Network Dashboard
- Sysmon Dashboard



IMPORTANCIA DE ELK STACK PARA LA SEGURIDAD

1. **Elasticsearch**

En seguridad, Elasticsearch permite consultar incidentes, correlacionar eventos y alimentar sistemas de detección (SIEM).

2. **Logstash / Fluentd**

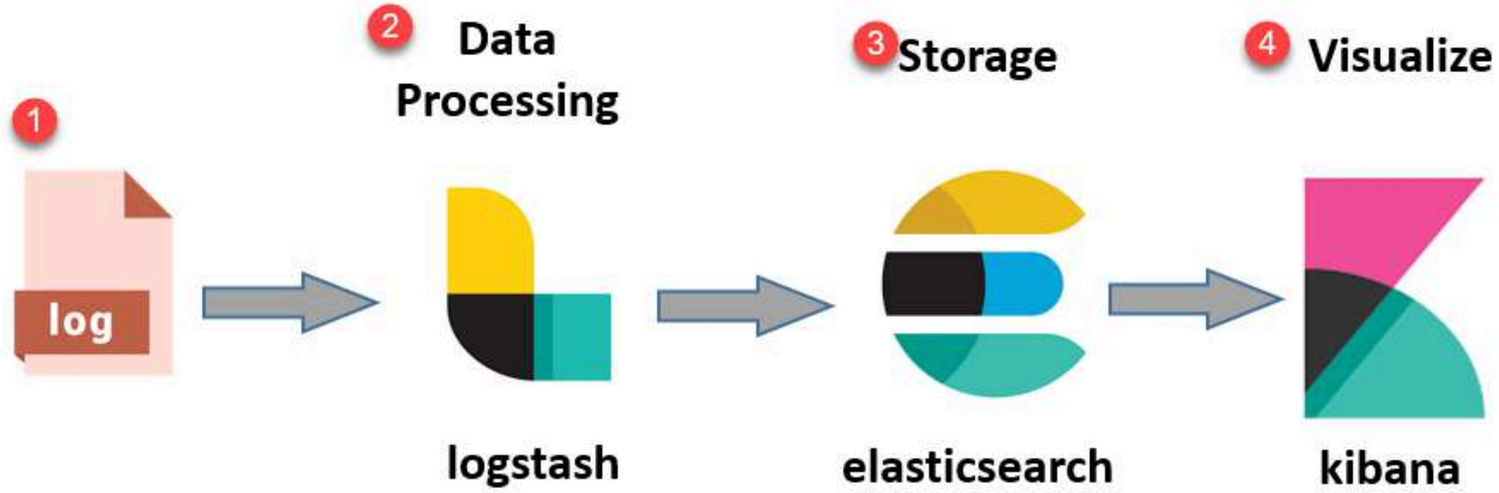
En seguridad, Logstash ayuda a limpiar y estructurar los logs, eliminando ruido y resaltando eventos críticos (por ejemplo, intentos de acceso no autorizado o cambios en privilegios).

En seguridad, Fluentd facilita el monitoreo de todo el cluster Kubernetes, ya que capta logs de nodos, pods y contenedores sin necesidad de configuración manual.

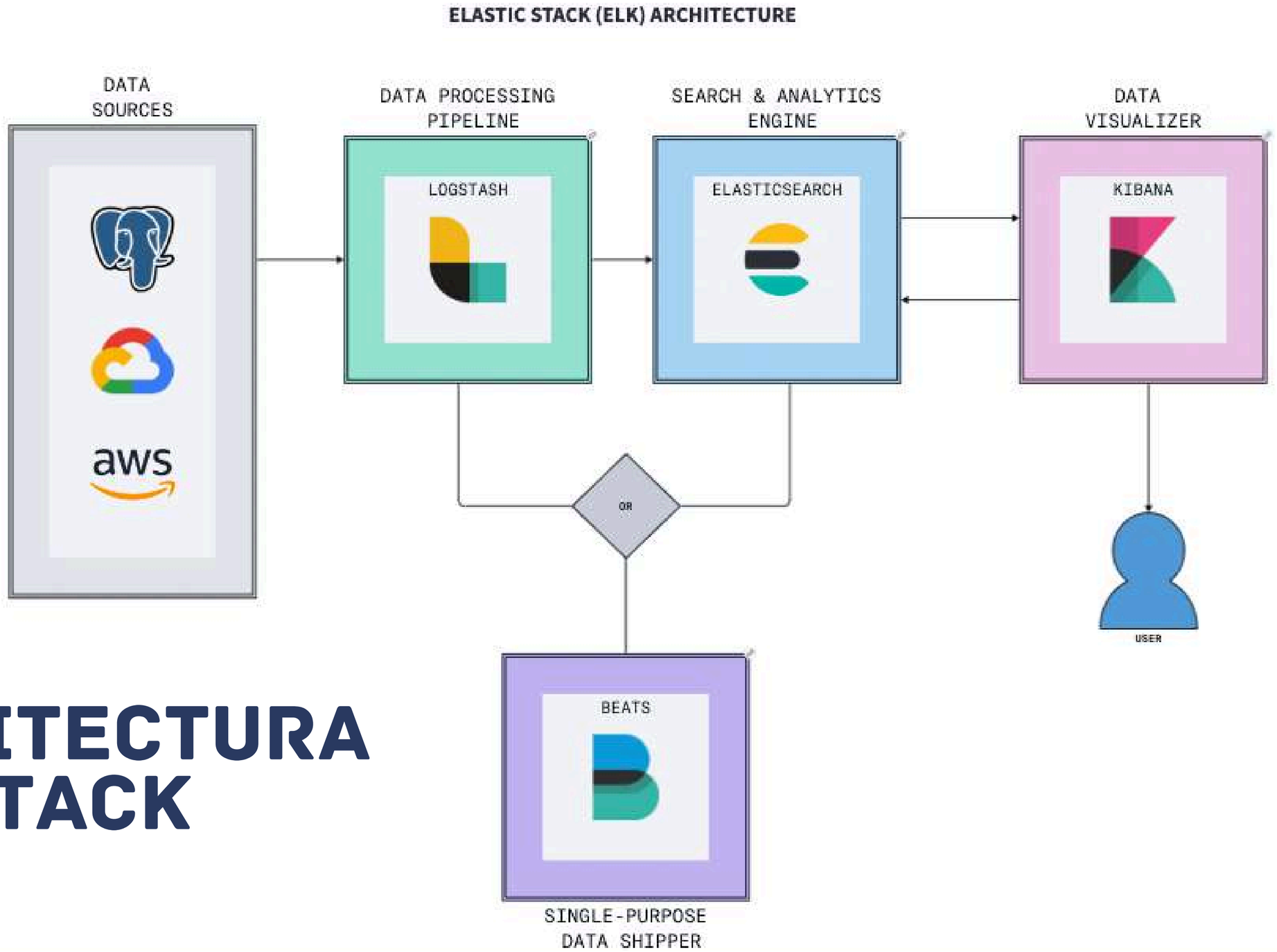
IMPORTANCIA DE ELK STACK PARA LA SEGURIDAD

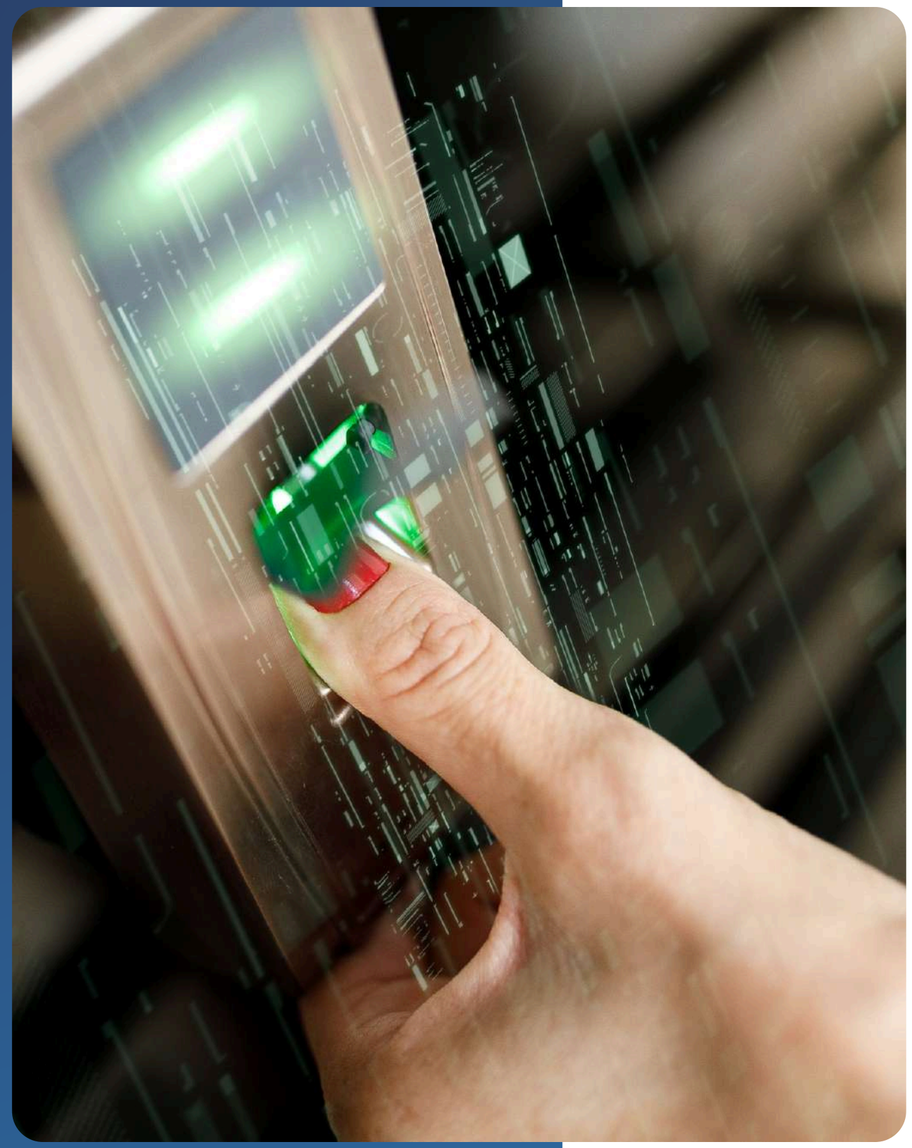
3. **Kibana**

En seguridad, Kibana es el centro de monitoreo donde los analistas visualizan alertas de Falco, violaciones de OPA, eventos del API Server, y cualquier otra fuente integrada.



ARQUITECTURA ELK STACK





DOCUMENTACIÓN INTEGRADA (OBSERVABILITY)

Cuando se habla de Documentación Integrada (Observability) usando ELK Stack, nos referimos a una estrategia avanzada de monitoreo y análisis dentro de sistemas distribuidos (como Kubernetes, servidores Linux o aplicaciones en contenedores), basada en la centralización y visualización de toda la información operativa del sistema en tiempo real.

DOCUMENTACIÓN INTEGRADA (OBSERVABILITY)

Observability es la capacidad de un sistema para permitirnos entender lo que ocurre internamente mediante la recopilación y el análisis de:

- Logs (registros de eventos)
- Metrics (métricas numéricas del rendimiento)
- Traces (rastros de peticiones o transacciones)

¿QUÉ ES LA DOCUMENTACIÓN INTEGRADA?

Se refiere a un repositorio centralizado de información operativa donde todo lo que el sistema genera —registros, métricas, errores, alertas— se documenta de manera automática y visual.

- No tienes que escribir manualmente qué pasó.
- Los logs, métricas y dashboards se convierten en una documentación viva del comportamiento del sistema.
- Todo está interrelacionado y consultable desde una misma interfaz (por ejemplo, Kibana).

¿CÓMO LO IMPLEMENTA EL ELK STACK?

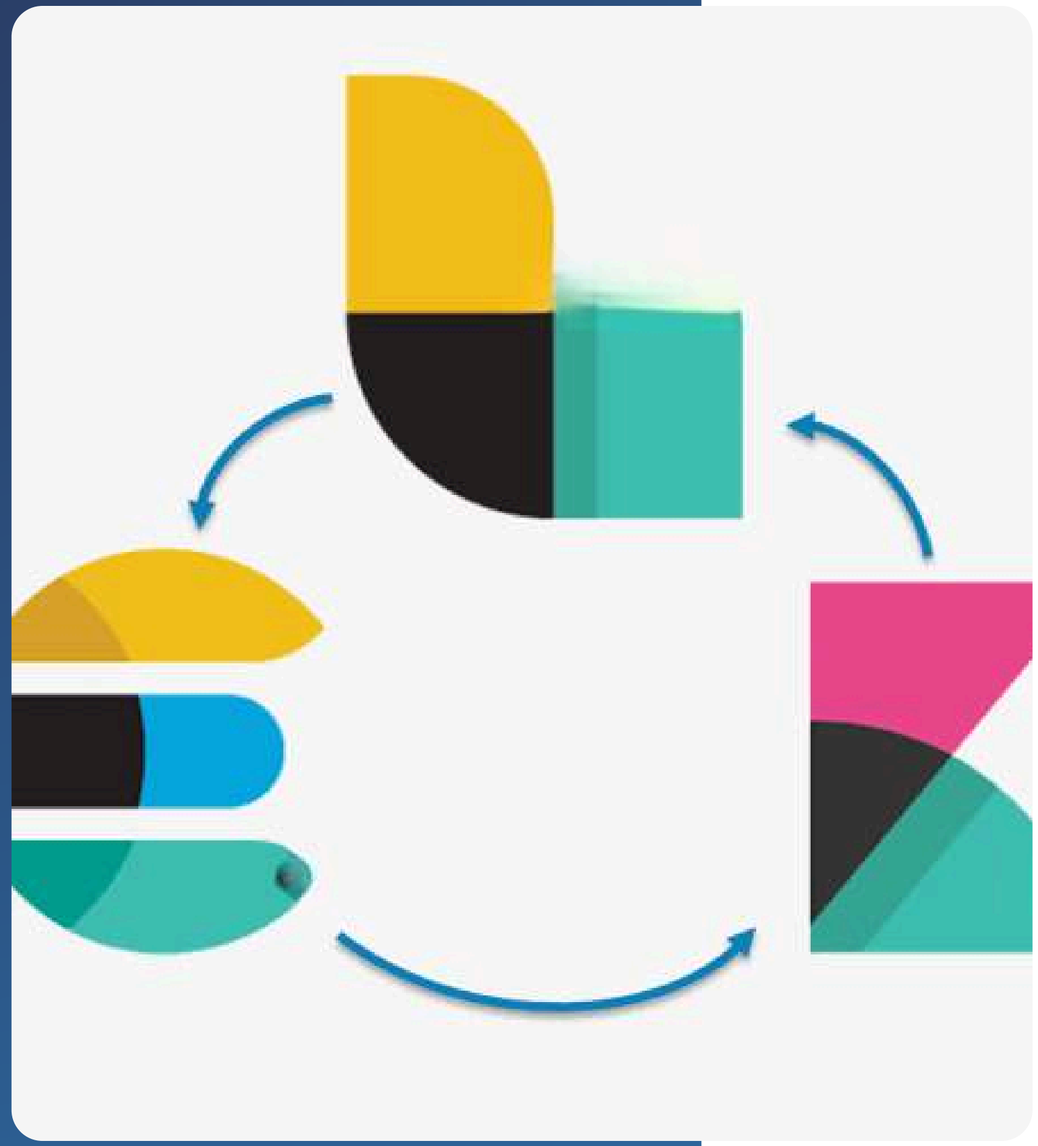
Componente	Función principal	Contribución a la observabilidad
Elasticsearch	Base de datos de búsqueda y análisis	Almacena logs, métricas y datos estructurados.
Logstash / Fluentd	Canal de ingestión y transformación	Recoge, limpia y transforma datos antes de indexarlos.
Kibana	Interfaz visual y de dashboards	Permite visualizar, consultar, filtrar y correlacionar datos.

EXTENSIÓN MODERNA DEL CONCEPTO (OBSERVABILITY STACK)

Hoy, el ELK Stack se suele complementar con herramientas como:

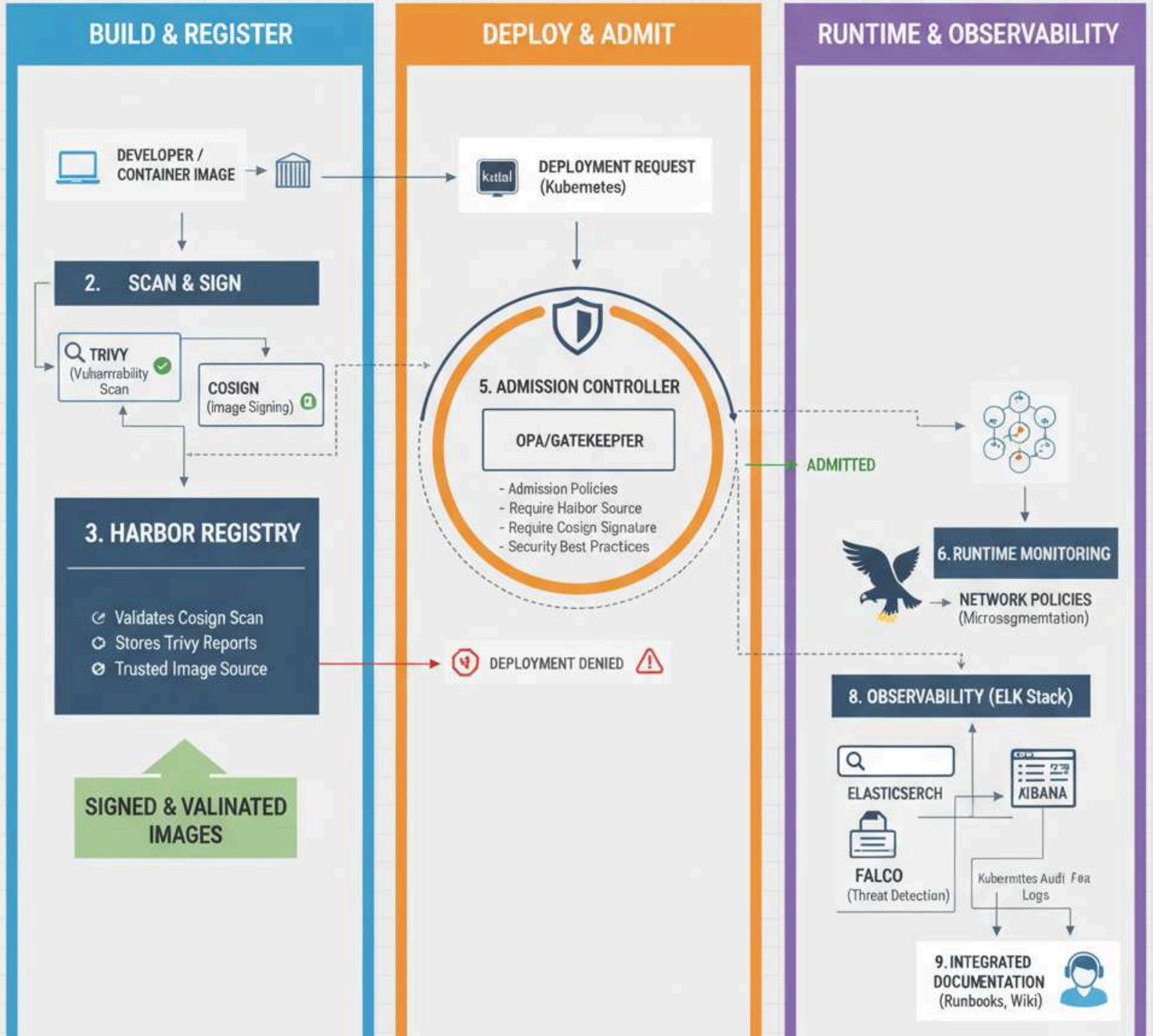
- Prometheus + Grafana (para métricas de rendimiento)
- Jaeger o OpenTelemetry (para trazas distribuidas)
- Alertmanager o ElastAlert (para alertas proactivas)

Así se logra una observabilidad completa de 360º: Logs + Métricas + Trazas = Sistema totalmente observable y documentado.



ARQUITECTURA COMPLEMENTARIA DE ELK STACK Y HERRAMIENTAS DE SEGURIDAD

DEVSECOPS ARCHITECTURE: CONTAINER SECURITY WORKFLOW



REFERENCIAS

Documentación oficial y fuentes técnicas primarias

Elastic. (2024). *Elastic Observability: Documentation.* Elastic.co.
Recuperado de https://www.elastic.co/guide/en/observability/current/index.html

Elastic. (2024). *Logstash Reference Documentation.* Elastic.co.
Recuperado de https://www.elastic.co/guide/en/logstash/current/index.html

Elastic. (2024). *Kibana Guide.* Elastic.co.
Recuperado de https://www.elastic.co/guide/en/kibana/current/index.html

Fluentd. (2024). *Fluentd Documentation.* Treasure Data, Inc.
Recuperado de https://docs.fluentd.org/

Prometheus Authors. (2024). *Prometheus Documentation.* The Prometheus Authors.
Recuperado de https://prometheus.io/docs/

Grafana Labs. (2024). *Grafana Documentation*. Grafana Labs.
Recuperado de https://grafana.com/docs/

Cloud Native Computing Foundation (CNCF). (2024). Jaeger: Open Source, *End-to-End Distributed Tracing. CNCF.*
Recuperado de https://www.jaegertracing.io/docs/

OpenTelemetry Authors. (2024). *OpenTelemetry Documentation.* Cloud Native Computing Foundation.
Recuperado de https://opentelemetry.io/docs/

Prometheus Authors. (2024). *Alertmanager Guide*. The Prometheus Authors.
Recuperado de https://prometheus.io/docs/alerting/latest/alertmanager/

Yelp Engineering. (2023). *ElastAlert 2 Documentation.* Yelp Inc.
Recuperado de https://elastalert2.readthedocs.io/en/latest/

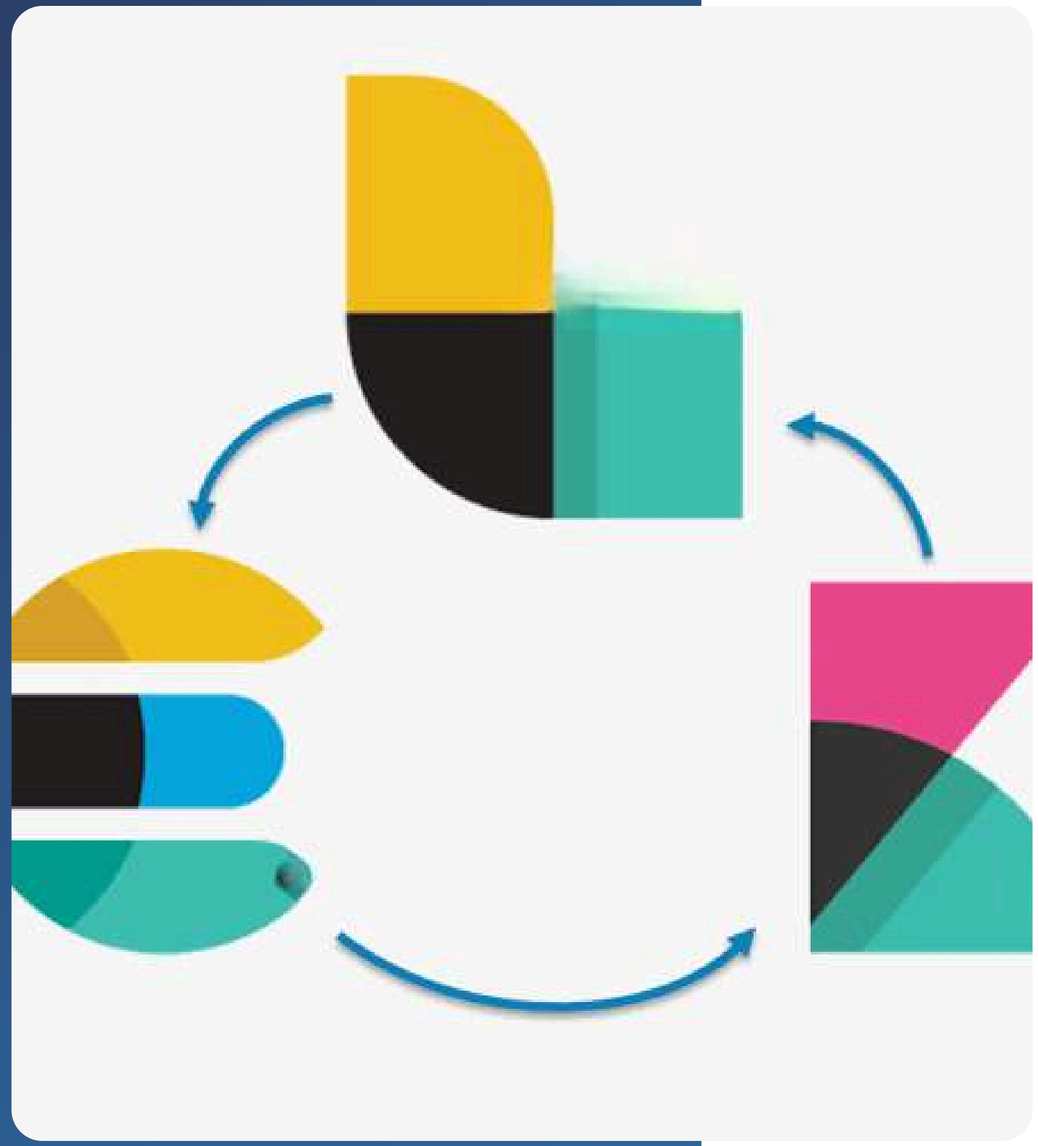
REFERENCIAS

Fuentes de referencia complementaria

Red Hat. (2024). *Understanding Observability in Kubernetes.* Red Hat, Inc. Recuperado de https://www.redhat.com/en/topics/observability

CNCF. (2024). *Cloud Native Observability Landscape Report.* Cloud Native Computing Foundation. Recuperado de https://landscape.cncf.io/

Docker Inc. (2024). *Docker Logging and Monitoring Overview.* Docker Docs. Recuperado de https://docs.docker.com/config/containers/logging/



DEMOSTRACIÓN DEL DESPLIEGUE DE ELK STACK EN KUBERNETES

¡POR SU ATENCIÓN GRACIAS!