Monitorización de métricas y logs en tiempo real utilizando **ELK stack**, **Metricbeat** y **Filebeat**

Que comience nuestro aprendizaje...

Tabla de **contenidos**

01 Definición de **herramientas**

Cómo funcionan cada una de las herramientas

ELK STACK

¿Qué es el stack ELK?

- Elastic Stack es un conjunto de herramientas de software libre y de código abierto que forman un potente conjunto de herramientas para la gestión de registros, análisis de datos en tiempo real y visualización de información.

Estas herramientas que componen la pila ELK son “Elasticsearch”, “Logstash” y “Kibana”

Hablemos de estas herramientas...

Elasticsearch

Elasticsearch es un motor de búsqueda y análisis de datos diseñado para almacenar, buscar y analizar grandes volúmenes de datos en tiempo real. Algunas de sus características y funcionalidades incluyen:

➔ Búsqueda de texto completo: Elasticsearch es capaz de indexar y buscar grandes cantidades de texto en tiempo real, lo que lo hace ideal para aplicaciones de búsqueda de texto completo.

➔ Escalabilidad: Elasticsearch es altamente escalable y puede manejar grandes cantidades de datos y consultas de búsqueda.

➔ Análisis de datos: Elasticsearch permite realizar análisis en tiempo real de los datos indexados, incluyendo agregaciones, estadísticas y más.

➔ Integración con otros componentes del Elastic Stack: Elasticsearch se integra con otros componentes del Elastic Stack para proporcionar una solución completa de búsqueda y análisis de datos.

Elasticsearch

API Elasticsearch

La API de Elasticsearch es una interfaz de programación de aplicaciones que proporciona un conjunto de endpoints para interactuar con el clúster de Elasticsearch. Esta API permite realizar una variedad de operaciones, como indexar, buscar, actualizar y eliminar documentos, administrar índices y configuraciones, realizar agregaciones y mucho más. La API de Elasticsearch es accesible a través de solicitudes HTTP/RESTful y responde en formato JSON.

Elasticsearch

Datos a destacar:

➔ Puerto sobre el que trabaja: 9200

➔ Fichero de configuración principal (en este fichero se especifican diferentes ajustes y opciones de configuración para el funcionamiento de un nodo de Elasticsearch): elasticsearch.yml

➔ Muy recomendable la creación y uso de un usuario de monitorización

➔ Almacena los datos en índices y cada dato almacenado se denomina documento

Elasticsearch

Rol para usuario de monitorización:

```
curl -H 'Content-Type: application/json' -u elastic:aPvUyEMp8WXbaS7NW9wH -XPUT
"http://localhost:9200/_security/role/usuario_monitoring" -d '{
  "indices": [
  {
    "names": ["*"],
    "privileges": ["create_index", "create_doc", "manage", "auto_configure", "all"]
   }
 ],
  "cluster": ["monitor"]
}'
```

Elasticsearch

Usuario de monitorización:

```
curl -H 'Content-Type: application/json' -u elastic:aPvUyEMp8WXbaS7NW9wH -XPOST
"http://localhost:9200/_security/user/angelsuarez" -d '{
  "password" : "Thieth5ta:u?faig",
  "roles" : ["kibana_admin", "monitoring_user", "usuario_monitoring"]
}'
```

Logstash

Logstash es una herramienta de procesamiento de datos que permite la ingestión, transformación y envío de datos a Elasticsearch o a otros destinos. Algunas de sus características y funcionalidades incluyen:

➔ Ingestión de datos: Logstash puede recibir datos de diferentes fuentes, incluyendo archivos de registro, bases de datos, servicios web y más.
➔ Transformación de datos: Logstash permite transformar los datos antes de enviarlos a Elasticsearch o a otros destinos, lo que puede incluir filtrar, enriquecer y modificar los datos.
➔ Conexión con diferentes sistemas: Logstash se conecta a diferentes sistemas y servicios, como bases de datos SQL, MongoDB, Kafka y más.
➔ Escalabilidad: Logstash es altamente escalable y puede manejar grandes cantidades de datos y procesamiento de datos.

Logstash

Datos a destacar:

➔ Puertos sobre los que trabaja: 5044 (entrada de Beats), 5000 (entrada de logs)

➔ Fichero de configuración principal (Contiene configuraciones generales que se aplican a todo el entorno de Logstash.): logstash.yml

➔ Ruta de ficheros de pipelines (en estos ficheros configuramos el input y output de los datos además del procesamiento denominado filter): /etc/logstash/conf.d/

Kibana

Kibana es una interfaz de usuario web para Elasticsearch que permite visualizar y analizar los datos indexados. Algunas de sus características y funcionalidades incluyen:

➔ Visualización de datos: Kibana permite crear visualizaciones de datos en tiempo real, incluyendo gráficos, tablas, mapas y más.
➔ Análisis de datos: Kibana permite realizar análisis de datos en tiempo real, incluyendo agregaciones, estadísticas y más.
➔ Búsqueda de texto completo: Kibana permite realizar búsquedas de texto completo en los datos indexados.
➔ Integración con Elasticsearch: Kibana se integra con Elasticsearch para proporcionar una solución completa de búsqueda y análisis de datos.

Kibana

Discover

Es uno de los apartados que considero más importantes de Kibana, donde podemos crear una nueva visualización de datos apuntando a nuestro índice de Elasticsearch y tener una visualización de nuestros datos recopilados, podemos especificar el rango de tiempo sobre el que queremos ver los datos recopilados y otras muchas opciones como aplicar filtros...

Dashboards

Los dashboards de Kibana son interfaces visuales que permiten mostrar y explorar datos de manera intuitiva y personalizada.

En estos dashboards podemos crear paneles, filtrar, explorar y analizar datos recopilados de manera más intuitiva.

Para crear un dashboard tenemos que tener clara la función de cada métrica recopilada para saber que datos nos ofrece y que panel podemos crear utilizando “x” métrica de los campos que hayamos exportado

Kibana

Datos a destacar:

➔ Puerto sobre el que trabaja: 5601

➔ Fichero de configuración principal (se utiliza para configurar opciones como la conexión con Elasticsearch, el puerto de escucha, la configuración de seguridad, las opciones de visualización y muchas más.): kibana.yml

Metricbeat

¿Qué es Metricbeat?

Metricbeat es una herramienta desarrollada por Elastic como parte de la pila ELK para recopilar y enviar datos de métricas del sistema y servicios a Elasticsearch o a otros destinos. A continuación, se presentan algunas de las principales características y funcionalidades de Metricbeat:

➔ Recopilación de métricas: Metricbeat recopila una amplia variedad de métricas de diferentes sistemas y servicios, como CPU, memoria, red, disco, bases de datos, servidores web y contenedores.

➔ Modularidad: Metricbeat está diseñado con un enfoque modular que permite agregar y configurar módulos específicos para recopilar métricas de sistemas y servicios específicos. Los módulos preconstruidos incluyen, por ejemplo, Apache, Docker, Elasticsearch, MySQL, MongoDB, Redis, entre otros.

➔ Monitoreo en tiempo real: Metricbeat envía los datos de métricas recopilados a Elasticsearch o a otros destinos en tiempo real, lo que permite el monitoreo y análisis en tiempo real de la infraestructura.

Metricbeat

Datos a destacar:

➔ Fichero de configuración principal (Aquí es donde se definen los ajustes globales y se habilitan los módulos específicos que deseas utilizar): metricbeat.yml

Módulos:

Los módulos de Metricbeat son componentes preconfigurados que facilitan la recolección de métricas específicas de diversos servicios y tecnologías.

Cada módulo tiene su propio conjunto de métricas y configuraciones específicas.

Los módulos de Metricbeat se almacenan en la ruta "/etc/metricbeat/modules.d", por defecto todos los módulos están deshabilitados, excepto el módulo de sistema, que por defecto recopila datos de cpu, carga, memoria, red...

Filebeat

¿Qué es Filebeat?

Filebeat es un componente del conjunto de herramientas de Elastic Stack, desarrollado por Elastic. Se utiliza para la recopilación, el envío y el procesamiento ligero de logs y otros datos de diferentes fuentes. Algunas de las principales características y funcionalidades de Filebeat son las siguientes:

➔ Recopilación de logs: Filebeat puede leer y recopilar logs de archivos de texto, logs de sistema y eventos de diferentes fuentes, como servidores, aplicaciones, servicios y más. Puede seguir de forma continua los cambios en los archivos de logs o enviar logs de forma periódica.

➔ Envío de logs a destinos: Filebeat puede enviar logs y otros datos a distintos destinos, como Elasticsearch, Logstash, Kafka y otros sistemas de almacenamiento o procesamiento.

➔ Módulos preconfigurados: Filebeat proporciona una serie de módulos preconfigurados para la recopilación de logs de fuentes populares, como Apache, Nginx, MySQL, Docker, sistema operativo, entre otros. Estos módulos simplifican la configuración y permiten una integración rápida con diferentes fuentes de logs.

Filebeat

Datos a destacar:

➔ Fichero de configuración principal (en este archivo definimos como Filebeat debe recolectar, procesar y enviar los archivos de logs a diferentes destinos): filebeat.yml

Redis

¿Qué es Redis?

Redis es una base de datos de tipo clave-valor, extremadamente rápida y versátil. Se destaca por su rendimiento excepcional y su capacidad para manejar diversas estructuras de datos.

Algunas características clave de Redis incluyen:

Almacenamiento en memoria: Redis almacena todos los datos en la memoria principal, lo que le permite ofrecer tiempos de respuesta extremadamente rápidos.

Estructuras de datos versátiles: Redis admite una amplia gama de estructuras de datos, incluyendo cadenas, listas, conjuntos, mapas hash, conjuntos ordenados y más.

Persistencia de datos: Aunque Redis almacena datos en memoria, también ofrece opciones para persistir los datos en disco.

Redis

Datos a destacar:

➔ Puerto sobre el que trabaja: 6379

➔ Fichero de configuración principal (Este fichero contiene una serie de opciones y parámetros que permiten personalizar y ajustar el comportamiento de Redis según las necesidades del entorno de implementación.): redis.conf

02

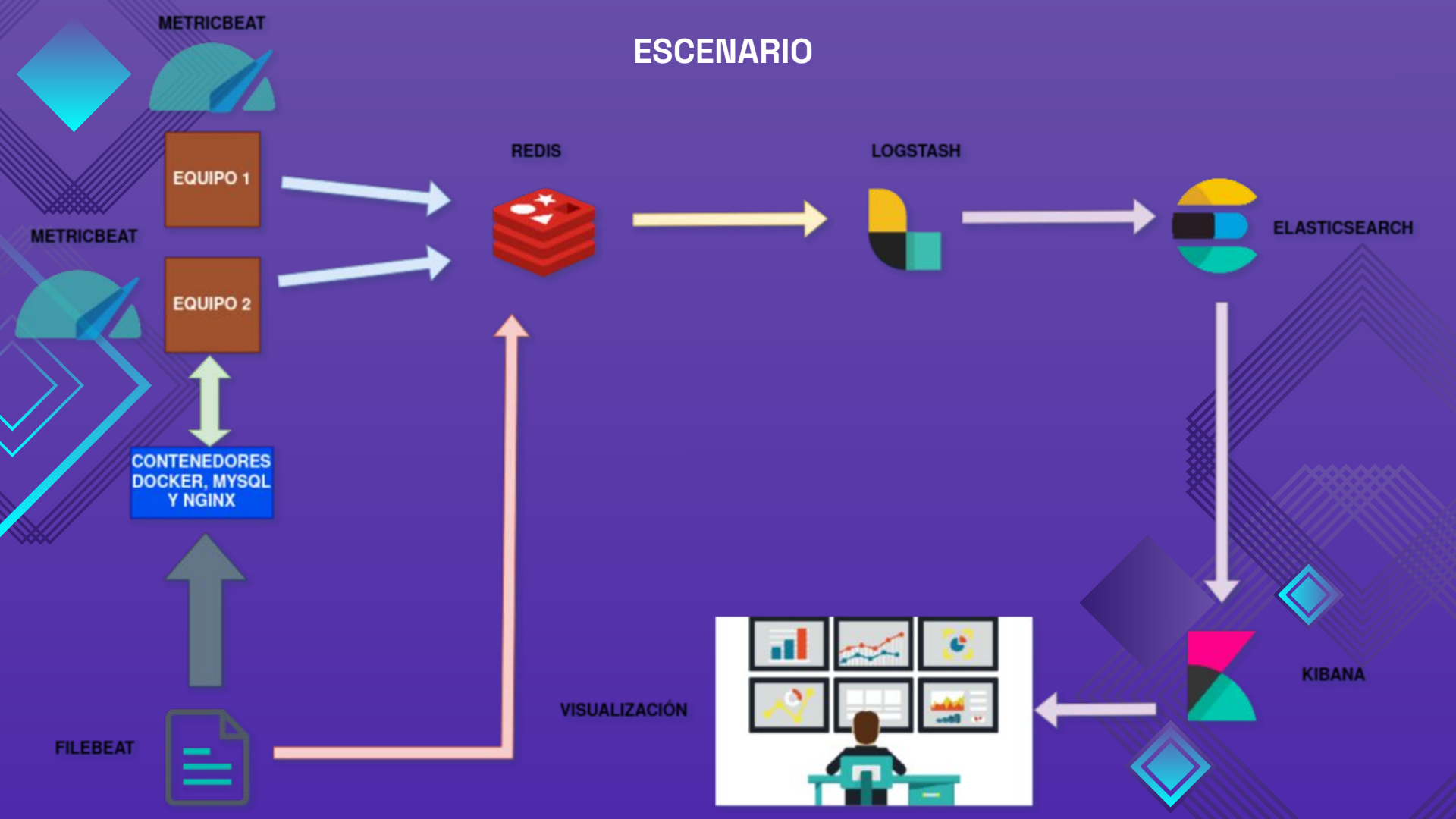
Escenario sobre el que he trabajado

Presentación del escenario de la demo

Cómo funciona el escenario

1. Metricbeat y Filebeat recopilan métricas y logs de los hosts configurados ➔ 2. Envían estos datos a una “key” de Redis, donde serán almacenados esperando ser procesados por Logstash ➔ 3. Logstash lee la cola de Redis y procesa esos datos para posteriormente enviarlos ya procesados a Elasticsearch ➔ 4. Elasticsearch los almacena en índices y documentos listos para que Kibana acceda a estos datos ➔ 5. Kibana accede a los índices configurados en Elasticsearch para configurar dashboards para la visualización sencilla e intuitiva de los datos recopilados

ESCENARIO


METRICBEAT
EQUIPO 1
METRICBEAT
EQUIPO 2
CONTENEDORES DOCKER, MYSQL Y NGINX
FILEBEAT
REDIS
LOGSTASH
ELASTICSEARCH
KIBANA
VISUALIZACIÓN