VIRTUALISATION

TP7/8 Observabilité Docker - Logs

I. Objectifs :

L'objectif visé par ce TP est de comprendre et de déployer une première infrastructure de gestion des logs. On va se reposer sur les briques suivantes :

- Filebeat
- Logstash
- Elasticsearch
- Kibana

Cette stack est une bonne solution de départ. Elle sera complétée par d'autres briques en fonction de nos besoins de supervision et aussi en fonction des applications monitorées.

En entreprise, pour chacune de ces briques on va déployer plutôt un cluster pour des raisons de performance et de hautes disponibilités (HA) :

- Logstash
- Elasticsearch
- Kibana

Et sur chaque machine on va déployer une (ou plusieurs) instance BEATS ou dans les architectures récentes un "Elastic agent".

II. ACTIONS

Les étudiants doivent séquentiellement réaliser les actions suivantes :

II.1. Redéployer l'infrastructure du TP4

Récupérer les données du TP4 et rejouer le docker-compose pour construire l'infrastructure. Ainsi on pourra collecter les logs de l'ensemble des conteneurs de l'application.

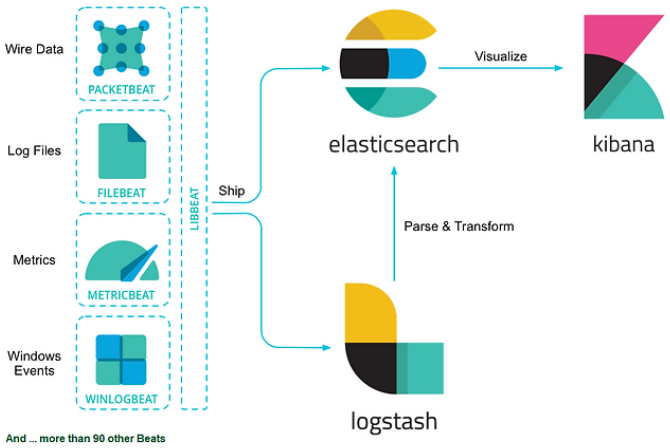
II.2. Déployer la stack de gestion des logs

Pour notre stack, on va déployer séquentiellement les briques suivantes :

- elasticsearch
- Kibana
- Filebeat

Ces derniers vont collecter/afficher les logs de l'ensemble des conteneurs présents sur notre infrastructure Docker.

On va donc être dans l'architecture **BEATS -> ELASTICSEARCH <- Kibana**



Dans un second temps on va commencer la manipulation de ces derniers et on va se rendre compte que tout est mis en vrac et qu'il n'y a aucun "**enrichissement**".

Pour cela on va ajouter un intermédiaire en déployant une brique logstash

On va donc être dans l'architecture **BEATS -> LOGSTASH -> ELASTICSEARCH <- Kibana**

Dans un troisième temps on va voir qu'il y a moyen de le faire plus simplement avec le concept de **module** dans BEATS qu'on pourra exploiter dans le cas de Docker avec les **LABEL "co.elastic.logs"**. On fera donc l'enrichissement automatique pour les conteneurs NGINX avec les labels.

On sera donc finalement dans un modèle hybride :

- **BEATS -> ELASTICSEARCH <- Kibana**
- **BEATS -> LOGSTASH -> ELASTICSEARCH <- Kibana**

Mais pour diverses raisons : ouverture de flux, gestion, … certaines entreprises simplifient en faisant tout passer par logstash avec un pipeline sans action s'il n'a pas besoin de manipuler les logs.

Pour ceux qui veulent aller plus loin vous pouvez vous abonner à la chaîne https://youtube.com/@duconceptalexpertise pour plus de cas pratiques et de configurations.