INF726

Manipulation de ELK Stack

Nicolas Gallay

-

TELECOM Paris

TELECOM
Paris

IP PARIS

Table des matières

Introduction

« ELK », aujourd'hui connu sous le nom de « Suite Elastic », est un acronyme faisant référence à trois projets distincts : Elasticsearch, Logstash et Kibana. Son but est d'ingérer simultanément des données provenant d'une multitude de sources, de les transformer, de les envoyer vers un système de stockage pour ensuite facilement les visualiser.

Le but de ce projet est d'implémenter ces outils et de les utiliser pour recréer cette chaine de traitement de bout en bout. Pour cela nous verrons tout d'abord l'environnement choisi et l'installation de ces outils. Nous verrons ensuite le traitement des données et sa visualisation.

Environnement

Afin de réaliser ce projet, le choix a été fait de partir d'un environnement Linux Ubuntu vierge installé sur une VM (VMWare fusion) selon le schéma suivant :

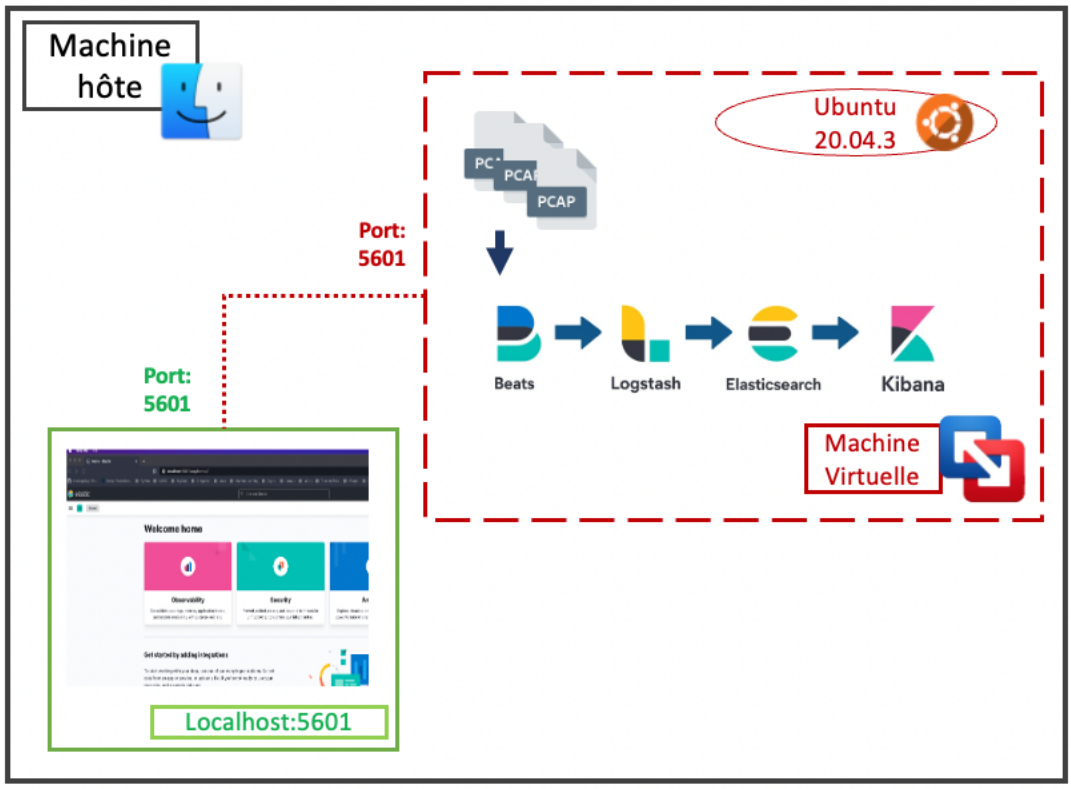


Figure 1: Architecture de notre environnement ELK

Cette VM abrite tous les outils nécessaires à la mise en place de la suite Elastic. Une redirection du port de l'UI de Kibana a été mise en place pour une meilleur visualisation de Kibana. Cette étape n'est pas obligatoire mais l'affichage de la VM VMWare fusion n'est pas optimal sur un Macbook.

Il est donc maintenant possible, après démarrages des trois outils, Elasticsearch, kibana et Logstash d'accéder, depuis le Localhost :5601 de la machine hôte, à l'interface kibana et aux différents dashboards.

Installation

L'installation de la stack ELK est un processus rapide et bien documenté. Une fois la machine virtuelle bien configurée la première étape a été d'installer les outils de la suite ELK : Elasticsearch, Logstash, Kibana et Packetbeat.

1. Téléchargement et installation (décompression) du tar.gz d'Elasticsearch trouvé à l'adresse suivante (Version Linux aarch64) :
   https://www.elastic.co/fr/downloads/elasticsearch

2. Lancer Elasticsearch depuis le dossier d'installation

```
$ bin/elasticsearch
```

3. Téléchargement et installation du tar.gz de kibana trouvé à l'adresse suivante (Version Linux aarch64) :
   https://www.elastic.co/fr/downloads/kibana

4. Lancer kibana depuis le dossier d'installation

```
$ bin/kibana
```

5. Faire une redirection du port 5601 de la VM vers le port 5601 de la machine hôte. Cette étape n'est pas nécessaire mais l'interface de la VM s'affiche mal. Cela permet juste un meilleur affichage de l'interface de Kibana.

```
$ sudo ssh -L  5601:localhost:5601 nicolasgallay@192.XXX.X.XX
```

6. Aller à l'url http://localhost:5601/ pour voir si l'application s'est bien lancée

7. Téléchargement et installation (décompression) du tar.gz de Logstash trouvé à l'adresse suivante (Version Linux aarch64) :
   https://www.elastic.co/fr/downloads/logstash

8. Création du fichier de configuration de base de Logstash : config/logstash.conf

```
input {
  beats {
    port => 5044
  }
}

output {
  elasticsearch {
    hosts => ["http://localhost:9200"]
    index => "%{[@metadata][beat]}-%{[@metadata][version]}-%{+YYYY.MM.dd}"

  }
```

```
}
```

Ce fichier de configuration de Logstash va permettre de créer des filtres ce qui permettra de traiter les fichiers de log avant de les envoyer dans Elasticsearch.

9. Lancer Logstash avec la conf créée précédemment

```
$ bin/logstash -f config/logstash.conf
```

Afin d'envoyer nos logs en format PCAP, nous allons utiliser Packetbeat.

10. Téléchargement et installation (décompression) du tar.gz de Packetbeat trouvé à l'adresse suivante (Version Linux aarch64) : https://www.elastic.co/fr/downloads/beats/packetbeat

11. Modification de packetbeat.yml, il faut commenter le bout de code concernant « output.elasticsearch » et dé-commenter « output.logstash ». Nous voulons que Packetbeat envoie les informations à Logstash pour qu'il puisse y appliquer les filtres.

```
# ------------------------------ Logstash Output -------------------------------
output.logstash:
  hosts: ["localhost:5044"]
```

12. Il est maintenant possible de lire des fichier PCAP. Soit en lançant Packetbeat en super utilisateur grâce à la commande suivante :

```
packetbeat-7.16.3-linux-arm64 ~ $ sudo ./packetbeat -e
```

Ceci va permettre la lecturee des paquets reçus et émis par l'ordianateur (VM). C'est cette option qui sera utilisé pour la suite.

Une autre solution est de télécharger un fichier PCAP depuis https://www.netresec.com/?page=PcapFiles et demander à Packetbeat de l'envoyer à Logstash. Des options permettent de l'envoyer en une seul fois ou en suivant les timestamps présents dans le fichier.

```
packetbeat-7.16.3-linux-arm64 ~ $ sudo ./packetbeat -e -c packetbeat.yml -t -I /home/nicolasgallay/Documents/pcap/4SICS-GeekLounge-151022.pcap -d "publish"
```

13. Il est maintenant possible de voir les paquets envoyés par Packetbeat sur Kibana dans l'onglet « Discover ». La prochaine étape est de mettre en place des filtres et créer des dashboard afin de mieux visualiser les données.

Parsing et indexation

Les données à notre disposition sont au format PCAP (« packet capture »). Packetbeat va se charger d'envoyer ces fichiers à Logstash. La combinaison de ces deux outils va permettre de parser et indexer automatiquement les différents champs de notre fichier de log.

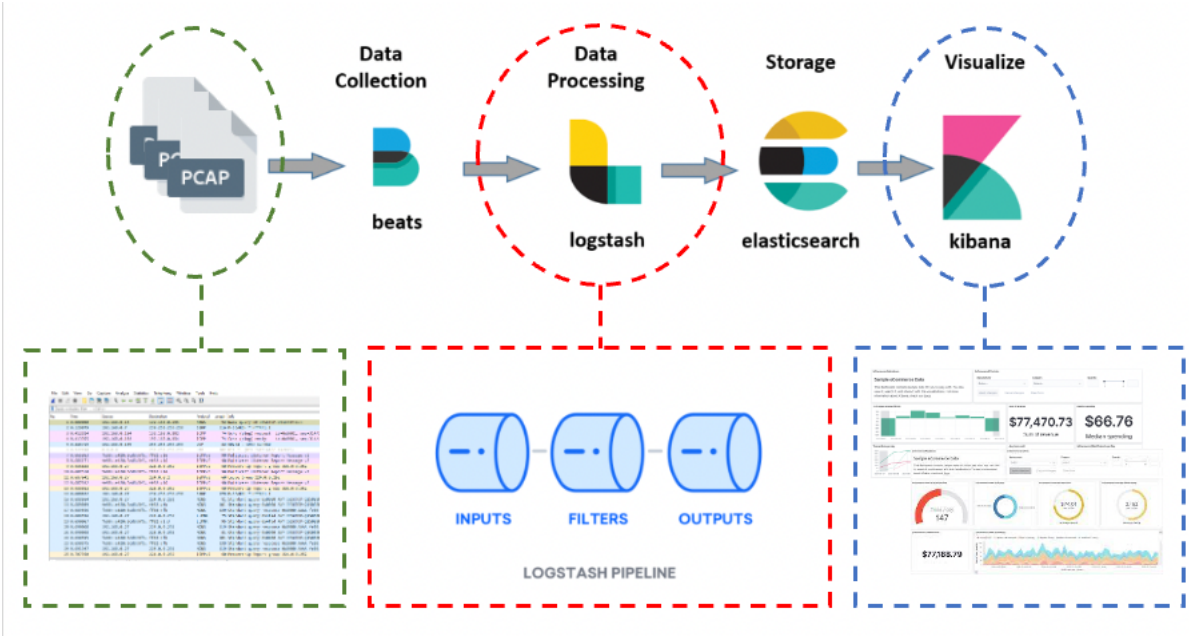


Figure 2: Workflow des données

Nous avons paramétré Logstash afin qu'il accepte en « input » les informations de packetbeat et qu'il envoie les informations traitées à Elasticsearch. Afin d'avoir plus d'information, l'ajout de filtre est obligatoire.

L'ajout de filtre se fait de la façon suivante : config/logstash.conf

```
input {
  beats {
    port => 5044
  }
}

filter {
        geoip {
              source => "[destination][ip]"
              }
        geoip {
              source => "[source][ip]"
              }
}

output {
  elasticsearch {
    hosts => ["http://localhost:9200"]
    index => "%{[@metadata][beat]}-%{[@metadata][version]}-%{+YYYY.MM.dd}"

  }
}
```

Données et résultats

Comme dit précédemment, nous allons lire les paquets envoyés et reçus par notre VM. En utilisant l'indexation par défaut et l'ajout de filtre Geoid, voici un exemple des informations qu'il a été possible de visualiser.

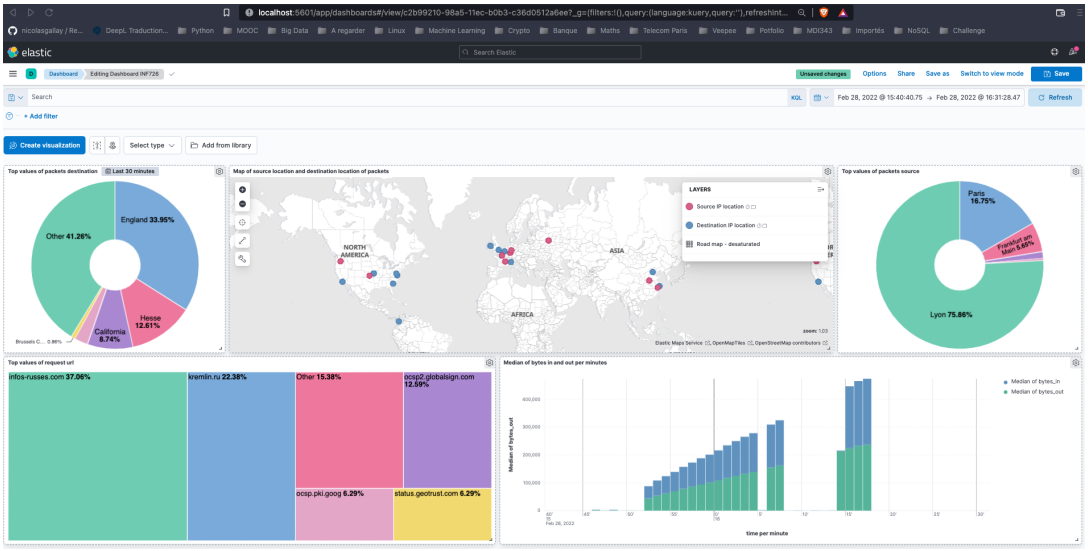


Figure 3: Aperçu du Dashboard Kibana

Après avoir laissé tourner Packetbeat quelques minutes, nous avons à notre disposition plus de 130 champs non vide (127 pouvant être visualisés) et plus de 40000 enregistrements. Tous ne sont pas forcements intéressent, cette sélection dépends de ce que l'on cherche à montrer. Les champs utiles ne sont pas les mêmes si l'on cherche à connaitre l'état d'un réseau, le statut des requêtes, le statut de certains serveurs, le type de requêtes…ou si l'on cherche à prévenir de certaines attaques et intrusion.

Il est possible par exemple de voir que le simple fait de lancer un navigateur, ici Firefox, de nombreuses informations sont échangées avec des serveurs un peu partout dans le monde.

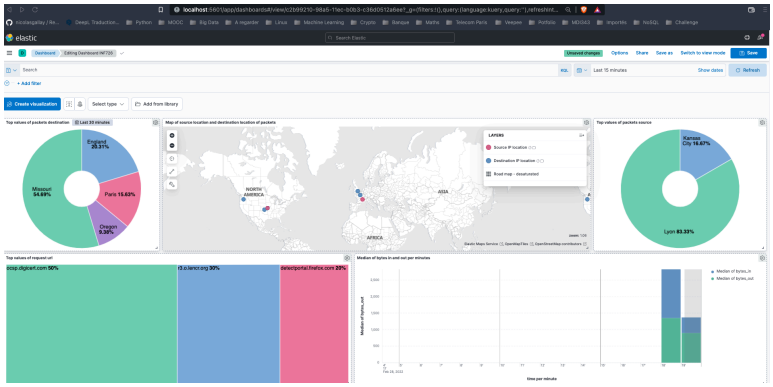


Figure 4: Capture d'écran du dashboard avant visite du site russe

En regardant de plus près, il s'agit notamment de « ocsp.digicert.com », de « r3.o.lencr.org » qui selon mes recherches permettent la validation des certificats en ligne des sites visités.

Un test a été fait sur la map du dashboard en se connectant au site du Kremlin. Une fois la connexion établie il est alors possible de voir des connections à un serveur situé à Moscou.

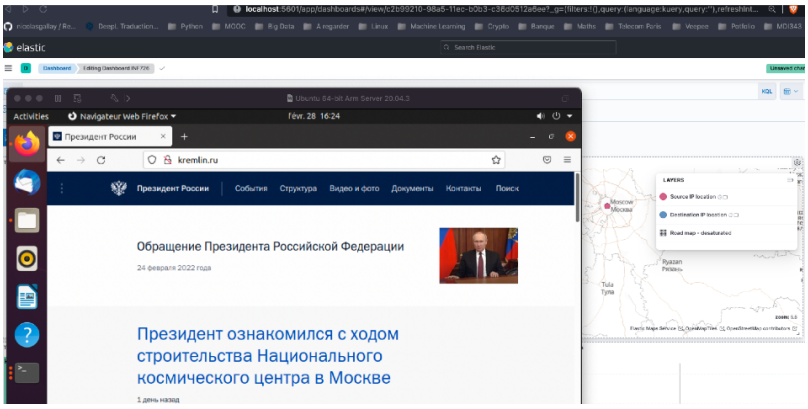
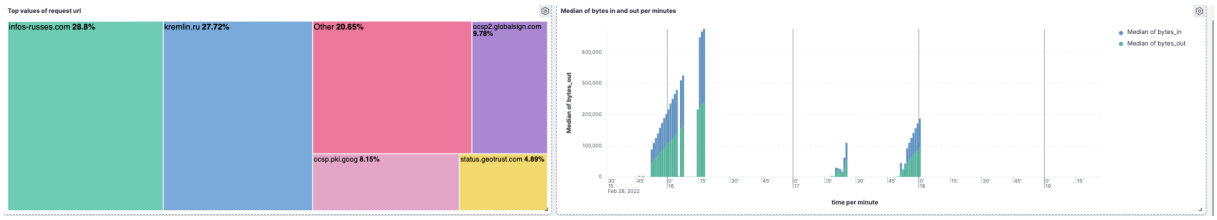


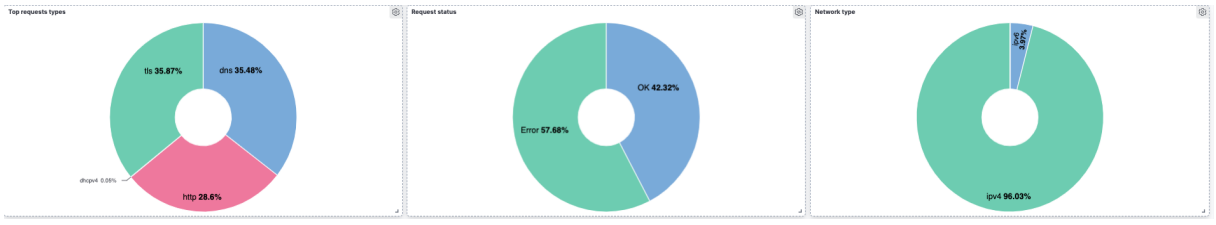
Figure 5: Capture d'écran après visite d'un site russe

En résumé, le Dashboard est composé d'une carte avec les localisations des ip source et ip de destination, de deux camemberts montrant le pourcentage de ces enregistrements et leur localisation plus précise.

Nous avons ensuite des informations sur les URL les plus visitées et la quantité d'information envoyée et reçue toutes les minutes.



Enfin nous avons des informations sur le type de requêtes, les statuts des requêtes et le type réseau.



Conclusion

ELK est outils pratique permettant de mettre en place rapidement des outils de monitoring et surveillance de logs. À notre échelle (un seul ordinateur, petit volume de log), il n'y a pas forcément d'utilité. On comprend cependant le rôle qu'ELK peut jouer au sein d'un entreprise.

ELK peut permettre en un simple coup d'œil de visualiser l'information en temps réel de ce qui se passe sur le réseau de l'entreprise et ainsi prendre conscience rapidement de potentielles menaces.

Le travail effectué ici ne fait bien sûr qu'effleurer la surface de ce qu'il est possible de faire à l'aide d'ELK mais a permis de s'informer sur les différents outils existants et d'en apprendre un peu plus sur la constitution et le fonctionnement du trafic réseau.