Rapport du projet

Mise en place d'un Pipeline du trafic réseau de mon ordinateur et visualisation sur un dashboard

Réalisé par : **MAHAMAT ATTEÏB Adoum**
Encadré par : **DREANO Julien**

Mastère spécialisé : Big Data
UE : Sécurité pour le Big Data

Année académique : 2023 - 2024

Table de matières

1. Contexte du projet

Ce projet consiste à créer une architecture qui englobe un pipeline permettant de capturer et d'acheminer les données provenant du trafic réseau de mon propre ordinateur, de les stocker sur la base de données NoSQL Elasticsearch et de les utiliser pour créer un dashboard sur Kibana. Le but du dashboard est de créer quelques indicateurs ou widgets permettant de le suivre en temps réel. Dans la partie qui suit, je présenterai les différentes étapes ainsi que les commandes Linux m'ayant permis d'installer les outils utilisés qui sont : Elasticsearch, Logstash, Kibana et Packetbeats. Pour chacun d'eux, c'est la dernière version qui a été installée : 8.12.

2. Les installations

J'ai procédé à une installation classique de ces 4 éléments cités ci-haut sur mon ordinateur portable sur lequel est installé le système d'exploitation Linux.

a. Elasticsearch

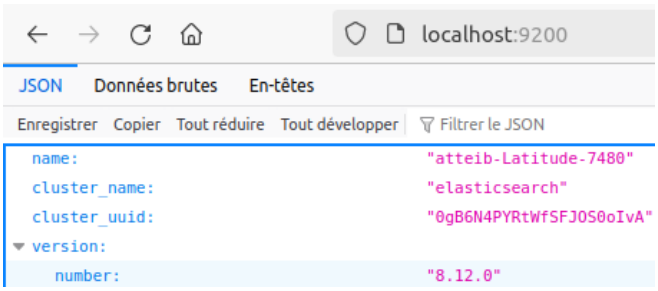
Elasticsearch est une base de données NoSQL. Elle nous permettra de stocker les données aspirées de mon trafic réseau. Les commandes bash utilisées pour son installation sont les suivantes :

```
wget https://artifacts.elastic.co/downloads/elasticsearch/elasticsearch-8.12.1-linux-x86_64.tar.gz
wget https://artifacts.elastic.co/downloads/elasticsearch/elasticsearch-8.12.1-linux-x86_64.tar.gz.sha512
shasum -a 512 -c elasticsearch-8.12.1-linux-x86_64.tar.gz.sha512
tar -xzf elasticsearch-8.12.1-linux-x86_64.tar.gz
cd elasticsearch-8.12.1/
```

Dans le fichier de configuration "**elasticsearch.yml**", j'ai mis "**network.host=0.0.0.0**" qui indique à Elasticsearch d'accepter les connexions sur toutes les interfaces réseau de la machine. En plus de cela, j'ai configuré "**http.host=0.0.0.0"** pour signifier que le service HTTP d'Elasticsearch acceptera les connexions de n'importe quelle adresse IP.

Une fois l'installation d'Elasticsearch terminée, je parviens à le lancer sans difficulté :

```
(base) atteib@atteib-Latitude-7480:~/elasticsearch-8.12.0$ ./bin/elasticsearch
warning: ignoring JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64; using bundled JDK
CompileCommand: exclude org/apache/lucene/util/MSBRadixSorter.computeCommonPrefixLengthAndBuildHistogram bool exclude = true
CompileCommand: exclude org/apache/lucene/util/RadixSelector.computeCommonPrefixLengthAndBuildHistogram bool exclude = true
févr. 10, 2024 10:36:51 AM sun.util.locale.provider.LocaleProviderAdapter <clinit>
WARNING: COMPAT locale provider will be removed in a future release
[2024-02-10T10:36:55,009][INFO ][o.a.l.i.v.PanamaVectorizationProvider] [atteib-Latitude-7480] Java vector incubator API enabled; uses preferr
edBitSize=256; FMA enabled
[2024-02-10T10:36:57,941][INFO ][o.e.n.Node               ] [atteib-Latitude-7480] version[8.12.0], pid[2612], build[tar/1665f706fd9354802c021
46c1e6b5c0fbcddfbc9/2024-01-11T10:05:27.953830042Z], OS[Linux/5.15.0-92-generic/amd64], JVM[Oracle Corporation/OpenJDK 64-Bit Server VM/21.0.1
/21.0.1+12-29]
[2024-02-10T10:36:57,943][INFO ][o.e.n.Node               ] [atteib-Latitude-7480] JVM home [/home/atteib/elasticsearch-8.12.0/jdk], using bun
```



Pour tester son bon fonctionnement, j'accède à Elasticsearch avec ce lien "http://localhost:9200/".

Le "name" correspond au nom d'utilisateur de ma machine. On y voit aussi la version installée.

b. Kibana

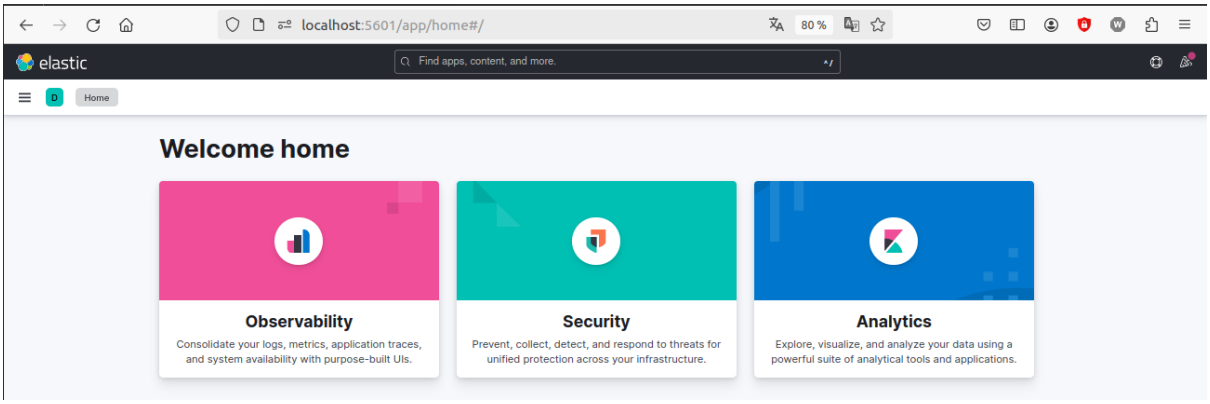
Kibana est un outil de visualisation, la création de dashboard. Les commandes bash utilisées pour son installation sont les suivantes :

```
curl -O https://artifacts.elastic.co/downloads/kibana/kibana-8.12.1-linux-x86_64.tar.gz
curl https://artifacts.elastic.co/downloads/kibana/kibana-8.12.1-linux-x86_64.tar.gz.sha512 | shasum -a 512 -c -
tar -xzf kibana-8.12.1-linux-x86_64.tar.gz
cd kibana-8.12.1/
```

A la fin de l'exécution des commandes ci-dessus, on parvient à démarrer Kibana :

```
(base) atteib@atteib-Latitude-7480:~/kibana-8.12.0$ ./bin/kibana
Kibana is currently running with legacy OpenSSL providers enabled! For details and instructions on how to disable see https://www.elastic.co/g
uide/en/kibana/8.12/production.html#openssl-legacy-provider
{"log.level":"info","@timestamp":"2024-02-10T14:32:45.071Z","log.logger":"elastic-apm-node","ecs.version":"8.10.0","agentVersion":"4.2.0","env
":{"pid":26675,"proctitle":"./bin/../node/bin/node","os":"linux 5.15.0-92-generic","arch":"x64","host":"atteib-Latitude-7480","timezone":"UTC+
0100","runtime":"Node.js v18.18.2"},"config":{"active":{"source":"start","value":true},"breakdownMetrics":{"source":"start","value":false},"ca
ptureBody":{"source":"start","value":"off","commonName":"capture_body"},"captureHeaders":{"source":"start","value":false},"centralConfig":{"so
urce":"start","value":false},"contextPropagationOnly":{"source":"start","value":true},"environment":{"source":"start","value":"production"},"g
lobalLabels":{"source":"start","value":[["kibana_uuid","88bf9daa-ed78-4115-bd0f-b0595f22c736"],["git_rev","e9092c0a17923f4ed984456b8a5db619b0a
794b3"]],"sourceValue":{"kibana_uuid":"88bf9daa-ed78-4115-bd0f-b0595f22c736","git_rev":"e9092c0a17923f4ed984456b8a5db619b0a794b3"}},"logLevel"
:{"source":"default","value":"info","commonName":"log_level"},"metricsInterval":{"source":"start","value":120,"sourceValue":"120s"},"serverUrl
":{"source":"start","value":"https://kibana-cloud-apm.apm.us-east-1.aws.found.io/","commonName":"server_url"},"transactionSampleRate":{"source
":"start","value":0.1,"commonName":"transaction_sample_rate"},"captureSpanStackTraces":{"source":"start","sourceValue":false},"secretToken":{"
source":"start","value":"[REDACTED]","commonName":"secret_token"},"serviceName":{"source":"start","value":"kibana","commonName":"service_name"
},"serviceVersion":{"source":"start","value":"8.12.0","commonName":"service_version"}},"activationMethod":"require","message":"Elastic APM Nod
e.js Agent v4.2.0"}
```

Lors de l'installation d'Elasticsearch, un token a été généré qu'on a utilisé afin d'établir la communication entre Elasticsearch et Kibana. Dans le fichier de configuration "**kibana.yml**", j'ai également mis "**server.host = 0.0.0.0**". Enfin, je parviens à ouvrir Kibana sur le navigateur web (voir capture d'écran) via ce lien "http://localhost:5601/app/home#/".



c. Logstash

Logstash permet l'envoi des données et aussi leur traitement ou transformation. En revanche, il nécessite la mobilisation de beaucoup de ressources contrairement aux beats. Dans le répertoire "**/etc/logstash**", on y trouve le fichier "**pipelines.yml**" dans lequel on liste les pipelines au cas où on en a plusieurs.

Les commandes Linux permettant de démarrer, de redémarrer et d'arrêter Logstash sont :

```
sudo systemctl restart logstash
sudo systemctl status logstash
sudo systemctl stop logstash
```

d. Packetbeats

Les commandes utilisées pour installer Packetbeats sont les suivantes :

```
curl -L -O https://artifacts.elastic.co/downloads/beats/packetbeat/packetbeat-8.12.1-linux-x86_64.tar.gz
tar xzvf packetbeat-8.12.1-linux-x86_64.tar.gz
```

Dans le fichier de configuration "**packetbeat.yml"** on désigne Logstash comme point d'arrivée de nos données provenant de Packetbeats. Les commandes Linux permettant de démarrer, redémarrer et d'arrêter Packetbeats sont similaires à celles présentées plus haut pour Logstash.

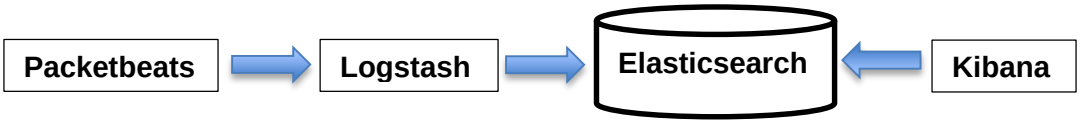
```
(base) atteib@atteib-Latitude-7480:/etc/packetbeat$ sudo systemctl status packetbeat
● packetbeat.service - Packetbeat analyzes network traffic and sends the data to Elasticsearch.
     Loaded: loaded (/lib/systemd/system/packetbeat.service; enabled; vendor preset: enabled)
     Active: active (running) since Tue 2024-02-06 20:09:22 CET; 8s ago
       Docs: https://www.elastic.co/beats/packetbeat
   Main PID: 94296 (packetbeat)
      Tasks: 9 (limit: 18954)
     Memory: 30.3M
     CGroup: /system.slice/packetbeat.service
             └─94296 /usr/share/packetbeat/bin/packetbeat --environment systemd -c /etc/packetbeat/packetbeat.yml --path.home /usr/share/pack
```

Une fois l'installation terminée, j'exécute également la commande ci-dessous pour s'assurer que la configuration de Packetbeats est parfaite.

```
(base) atteib@atteib-Latitude-7480:/etc/packetbeat$ sudo packetbeat test config
Config OK
```

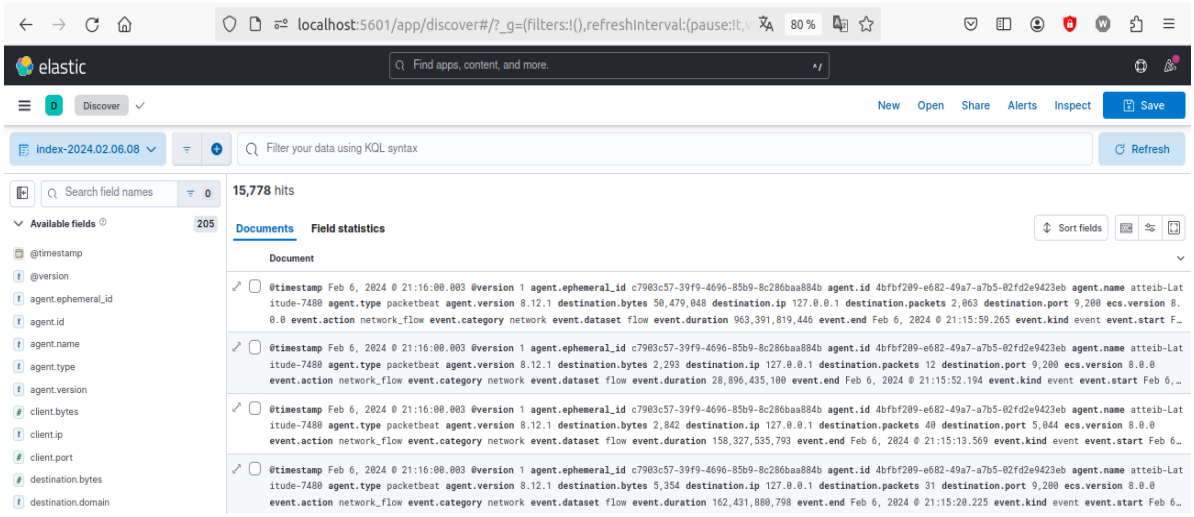
3. Architecture

Ci-dessous, l'architecture de mon pipeline. Mon pipeline comporte 3 structures qui sont « **inputs** », « **filter** » et « **output** ». Sur Input, on spécifie la provenance de données ; sur filter, les transformations voulues et enfin sur output la destination de données.



4. Donnée : Trafic réseau

Dès l'envoi des données sur Elasticsearch, on peut visualiser leur disponibilité en cliquant sur "Management" puis sur "index management". Sur la capture d'écran ci-dessous, on peut s'assurer de la disponibilité des données capturées de mon trafic réseau sur la base de données NoSQL Elasticsearch.



Sur Kibana, j'ai ensuite créé un index pattern qui porte le même nom que celui disponible sur Elasticsearch, les deux doivent matchés afin de pouvoir créer des graphiques.

5. Dashboard

Le dashboard que j'ai construit à partir du trafic réseau de mon ordinateur comporte 7 widgets. Sur la première ligne, on y trouve :

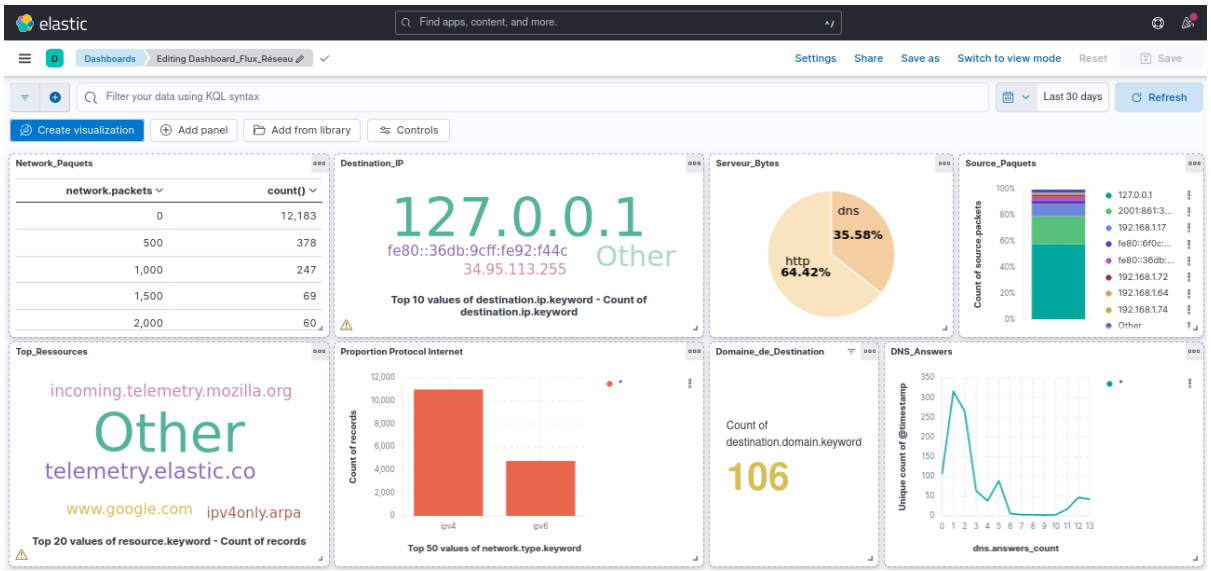
- un tableau sur le nombre de paquets réseau ;
- les IP de destination le plus majoritaire ;

- un camembert sur la proportion de serveurs (dns et http) ;
- un diagramme en barre empilé sur les sources de paquets ;

Sur la deuxième ligne, on y trouve :

- Le top ressources ;
- un diagramme en barre sur le type de protocol ;
- l'effectif du domaine de destination ;
- une courbe sur l'évolution de réponse de dns.

Ci-joint une capture d'écran de mon dashboard. Vous pouvez importer le format « **ndjson** » de mon dashboard pour le visualiser.



Vous avez la possibilité d'importer ce dashboard que vous avez reçu sous format « ndjson ».

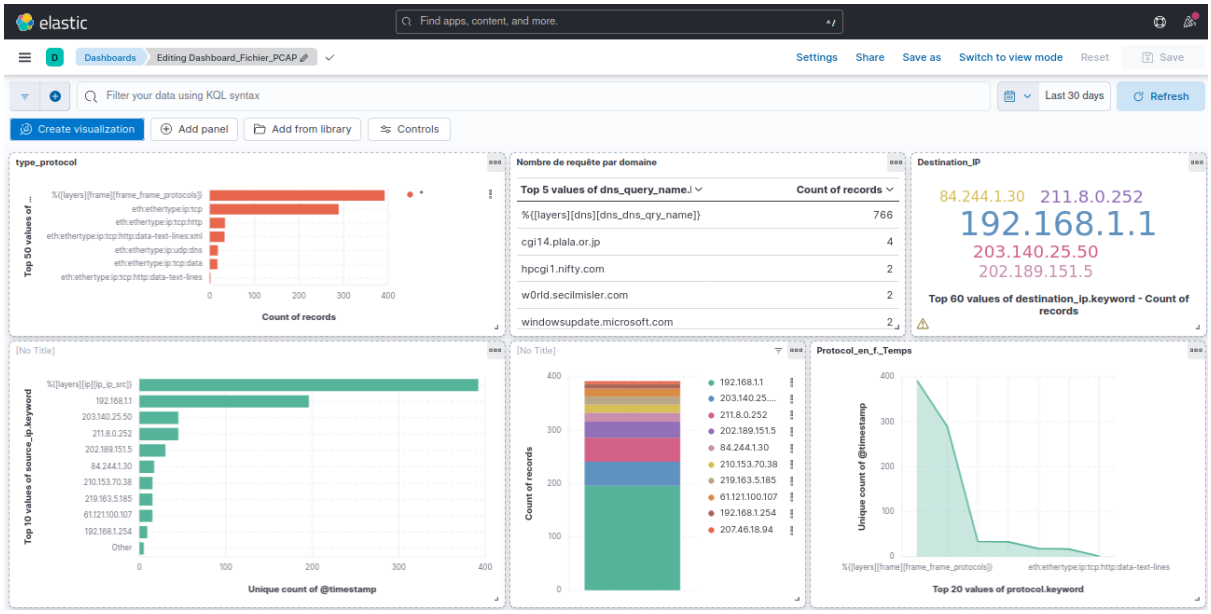
6. Conclusion

N'ayant jamais utilisé le Stack ELK, ce projet a été pour moi l'occasion de le découvrir et de le manier pour la réalisation d'une architecture incluant la collecte des données, leur transformation et leur stockage ainsi que la création de deux dashboards, un premier sur les données du trafic réseau de mon ordinateur et le second mis en annexe en utilisant le fichier "toolsmith.pcap".

Ce projet a été pour moi l'occasion de me familiariser davantage avec les commandes Linux. En guise de perspective, j'envisage de réaliser un projet en utilisant la Stack ELK et Apache Kafka ou encore avec X-pack, l'extension payante pour Kibana et Logstash, offrant des fonctionnalités de sécurité, de monitoring, d'alerting et de reporting.

7. Annexe : 2ème dashboard réalisé avec le fichier “toolsmith.pcap”

Dans l’onglet “Single PCAP files” ce cette page https://www.netresec.com/?page=pcapfiles, J’ai téléchargé ce fichier “https://holisticinfosec.io/toolsmith/files/nov2k6/toolsmith.pcap”. Il concerne les machines infectées par W32/Sdbot. Ci-joint, un deuxième dashboard créé en utilisant ce fichier.



8. Bibliographie

- https://www.elastic.co/
- https://www.elastic.co/fr/beats/packetbeat
- https://www.netresec.com/?page=pcapfiles
- https://fr.wikipedia.org/wiki/Conteneur_(informatique)
- https://fr.wikipedia.org/wiki/Virtualisation
- https://fr.wikipedia.org/wiki/Pcap