



**DATA**PROTECT  
Security is our **commitment**

ECOLE NATIONALE SUPÉRIEURE D'INFORMATIQUE ET D'ANALYSE DES  
SYSTÈMES - RABAT

---

SUJET : CONCEVOIR UNE ARCHITECTURE DE  
PIPELINE RÉSILIENTE POUR L'ANALYSE DES  
LOGS EN TEMPS RÉEL

---

*Réalisé par :*

KOUSSY AYOUB

*Encadré par :*

M. BOUZIANI OSSAMA





## *Remerciements :*

Avant d'entrer dans les détails, il est opportun de commencer ce rapport en exprimant ma gratitude envers ceux qui m'ont aidé tout au long de sa réalisation, ainsi qu'envers ceux qui ont rendu cette expérience enrichissante et profitable. Je tiens tout d'abord à remercier chaleureusement M. BOUZIANI Ossama, qui m'a formé et accompagné avec patience et pédagogie tout au long de cette expérience professionnelle. Je lui suis reconnaissant pour son assistance, son suivi et le temps qu'il a consacré à superviser le bon déroulement de mon sujet.

Je souhaite également exprimer ma gratitude envers tous les responsables de stage et le personnel de CyberSOC DATAPROTECT pour l'expérience captivante et enrichissante que j'ai vécue au sein de l'entreprise pendant ces deux mois .

En outre, je tiens à remercier sincèrement nos professeurs à l'École Nationale Supérieure d'Informatique et d'Analyse des Systèmes pour leur formation qui nous a permis de développer les compétences nécessaires à la réalisation de ce travail. Je suis également reconnaissant envers les membres du jury qui ont accepté de juger notre travail.



# Résumé

Le présent document constitue une synthèse de mon projet réalisé dans le cadre de mon stage de fin de 2ème année. L'objectif de ce projet était de mise en place d'une architecture de pipeline résilient pour la collecte, l'analyse et le stockage des logs. Nous avons pu mettre en oeuvre une solution complète qui permet aux entreprises de surveiller efficacement leurs systèmes informatiques, de détecter les anomalies et les cyberattaques potentielles, et de prendre des mesures appropriées pour y remédier.

ce projet a permis de mettre en place une infrastructure solide et résiliente pour la gestion des logs, offrant aux entreprises la possibilité de détecter rapidement les problèmes potentiels, de prendre des mesures appropriées et de renforcer la sécurité de leurs systèmes informatiques.

Ce projet nous a permis de comprendre l'importance de la gestion des logs dans la sécurité des systèmes d'information. En collectant, analysant et visualisant les logs en temps réel, nous pouvons détecter rapidement les cyberattaques, identifier les anomalies et prendre des mesures préventives pour protéger les systèmes.

.

# Abstract

This document is a summary of my project completed during my second-year internship. The objective of this project was to establish a resilient pipeline architecture for log collection, analysis, and storage. We were able to implement a comprehensive solution that enables businesses to effectively monitor their IT systems, detect anomalies and potential cyber attacks, and take appropriate measures to address them.

This project allowed us to establish a robust and resilient infrastructure for log management, providing businesses with the ability to quickly detect potential issues, take appropriate actions, and enhance the security of their IT systems.

Through this project, we gained an understanding of the importance of log management in information system security. By collecting, analyzing, and visualizing logs in real-time, we can swiftly detect cyber attacks, identify anomalies, and implement preventive measures to protect the systems.

# Table des figures

1.1	Fiche technique . . . . .	3
1.2	les pôles de DATAPROTECT . . . . .	3
1.3	CyberSOC DATAPROTECT . . . . .	4
1.4	Les avantages du Cybersoc . . . . .	5
2.1	Planning du projet . . . . .	9
3.1	Architecture Globale . . . . .	11
3.2	Architecture Globale Adjuster . . . . .	12
3.3	Elasticsearch . . . . .	12
3.4	Logstach . . . . .	13
3.5	Kibana . . . . .	13
3.6	Winlogbeat . . . . .	13
3.7	Apache Kafka . . . . .	14
3.8	Relation de composants Elasticsearch . . . . .	15
3.9	Le concept d'Apache Kafka . . . . .	15
3.10	Architecture Kafka Ajuster . . . . .	16
3.11	Architecture Kafka Utilise . . . . .	16
3.12	Exemples des logs . . . . .	17
3.13	Quelques informations de journal . . . . .	17
4.1	Composent de Dosiser Elasticsearch . . . . .	20
4.2	Yml Elasticsearch Configuration . . . . .	22
4.3	Yml Kibana Configuration . . . . .	22
4.4	Yml Winlogbeat Configuration . . . . .	24
4.5	Kafka Configuration . . . . .	24
4.6	Logstach Configuration . . . . .	25
4.7	Commandes pour lancer la pipeline . . . . .	25
4.8	Visualisez les journaux . . . . .	26
4.9	Anomaly Detection Jobs . . . . .	26
4.10	Anomaly Explorer . . . . .	27

# Table des matières

<b>1</b>	<b>Présentation de l'organisme d'accueil</b>	<b>1</b>
1.1	Présentation de l'organisme . . . . .	2
1.2	Les valeurs de l'entreprise . . . . .	2
1.3	Fiche technique . . . . .	2
1.4	Les activite de DATAPROTECT . . . . .	3
1.5	Le Pôle CyberSOC DATAPROTECT . . . . .	4
<b>2</b>	<b>Contexte général du projet</b>	<b>6</b>
2.1	Mise en situation . . . . .	7
2.2	Présentation du projet . . . . .	7
2.2.1	Introduction . . . . .	7
2.2.2	Problématique . . . . .	7
2.2.3	Solutions et objectifs . . . . .	7
2.2.4	Etude de faisabilité . . . . .	8
2.3	Planning du stage . . . . .	9
2.4	Conclusion . . . . .	9
<b>3</b>	<b>Conception d'Architecture de pipeline</b>	<b>10</b>
3.1	Architecture Globale de Pipligne & Technologies utilisés . . . . .	11
3.1.1	Architecture Globale de Pipligne . . . . .	11
3.1.2	Technologies utilisés . . . . .	12
3.2	Architecture de Base de Données . . . . .	14
3.3	Architecture d'Apache Kafka . . . . .	15
3.4	Analyse des logs . . . . .	17
3.5	Conclusion . . . . .	18
<b>4</b>	<b>Réalisation du projet</b>	<b>19</b>
4.1	Configuration de Pipeline . . . . .	20
4.1.1	Configuration d'Elasticsearch & Kibana . . . . .	20
4.1.2	Configuration de Winlogbeat & Apache Kafka et Logstach . . . . .	23
4.2	Visualisation des résultats . . . . .	25
<b>5</b>	<b>Conclusion et Perspectives</b>	<b>28</b>





# Chapitre 1

## Présentation de l'organisme d'accueil

---

Dans ce chapitre, nous allons nous intéresser tout d'abord à une présentation générale de DATAPROTECT : activités , valeurs et Pôles .

## 1.1 Présentation de l'organisme

DATAPROTECT est une entreprise spécialisée en sécurité de l'information. Fondée par Ali EL AZZOUZI, un expert en sécurité de l'information ayant mené plusieurs projets de conseil et d'intégration de solutions de sécurité, DATAPROTECT appuie son offre sur une vision unifiée de la sécurité de l'information. Dotée d'un réservoir de compétences pointues en sécurité lui permettant d'assurer une expertise unique sur le marché local et régional. Depuis sa création, DATAPROTECT ne cesse d'évoluer pour délivrer ses prestations d'excellence à travers une équipe d'experts pluridisciplinaires dotée d'un sens unique de l'intimité client. Aussi, son statut d'entité accréditée PCI QSA et PA QSA sur les zones CEMEA et EUROPE par le consortium Payment Card Industry Security Standards Council pour les certifications PCI DSS et PA DSS, fait d'elle un acteur unique dans la région.

Avec une centaine de clients en Europe, Afrique et Moyen Orient, DATAPROTECT est aujourd'hui capable de délivrer ses services en toute agilité, pour des multinationales comme pour des entreprises locales, avec à la clé une réputation établie de pionnier sur la thématique de la sécurité de l'information.

## 1.2 Les valeurs de l'entreprise

Nos valeurs forment le socle de la culture de DATAPROTECT. Elles fixent l'orientation de la stratégie de l'entreprise et donnent du sens à nos actions quotidiennes. Pérennes, ces valeurs participent à faire comprendre aux collaboratrices et collaborateurs la raison d'être de l'entreprise.

Écoute client : Une implication à tout niveau est portée auprès de nos clients, pour mieux anticiper, conseiller et accompagner.

Respect et diversité : Notre diversité est notre richesse. Le respect mutuel et la confiance font partie de nos convictions fondamentales.

Innovation : La créativité des collaborateurs véritable boost de notre dynamisme, est fortement encouragée.

Professionnalisme : Notre exigence de qualité et de transparence est un levier de développement et nos efforts d'engagement sont constants

## 1.3 Fiche technique

Logo :	<b>DATA</b> PROTECT Security is our <b>commitment</b>
Date de création :	05/2009
Forme juridique :	Société à Responsabilité Limitée à Associé Unique
Siège sociale :	Tour CFC Lot N° 57 10ème étage Quartier Casa-Anfa, Hay Hassani, Casablanca Maroc
Direction :	Ali El Azouzi
Activité(s) :	Développement et intégration de solutions informatiques et digitales, BPO et ITO
Effectif :	Entre 50 et 100
CHIFFRE D'AFFAIRES :	De 5,000,000 à 10,000,000 dh
Site web :	www.dataprotect.ma

FIGURE 1.1 – Fiche technique

## 1.4 Les activités de DATAPROTECT

DATAPROTECT propose une approche globale et met des experts certifiés pour répondre à tout besoin en matière de cybersécurité. Il se spécialise dans plusieurs domaines clés pour assurer la protection des données et la sécurité des cyberprojets.

DATAPROTECT est une entreprise qui propose différents services dans le domaine de la sécurité des données. Elle est organisée en plusieurs pôles, chacun ayant des responsabilités spécifiques.

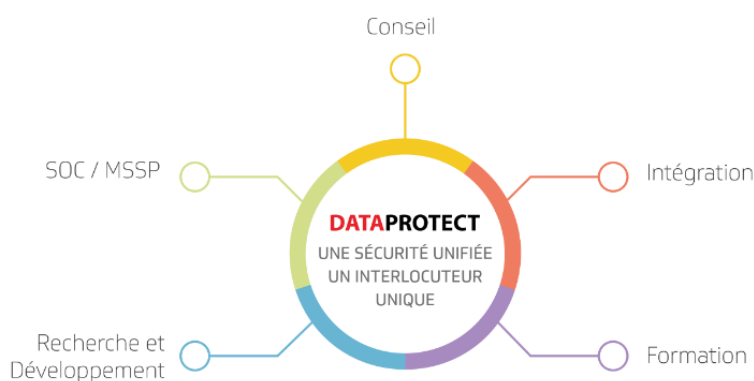


FIGURE 1.2 – les pôles de DATAPROTECT

**1.Pôle Conseil :** Ce pôle se concentre sur le conseil en matière de sécurité. Il propose des services tels que l'audit de sécurité, les tests d'intrusion récurrents, l'audit d'ingénierie sociale et l'audit de sécurité des infrastructures critiques. Il aide également les entreprises à se conformer aux normes de sécurité telles que l'ISO 27001 et l'ISO 22301. De plus, ce pôle accompagne les organisations dans la mise en place de plans de continuité d'activité et de certifications telles que PCI DSS et PA DSS. Il offre également des services d'analyse des risques IT et d'accompagnement à la protection des infrastructures critiques.

**2.Pôle Intégration :** Le pôle Intégration se concentre sur la mise en place de solutions de sécurité. Il offre des services tels que la sécurisation des postes de travail, le filtrage URL, le firewalling, le firewall applicatif,

la prévention des intrusions, la gestion des vulnérabilités, la prévention des fuites d'informations sensibles, le cryptage des données, le contrôle d'intégrité, la solution PKI (Infrastructure à clés publiques), la traçabilité des accès aux bases de données, la gestion des comptes à hauts privilèges, la gestion des événements de sécurité (SIEM), la protection contre les attaques Zéro Day et la synchronisation des horloges.

**3.Pôle SOC/MSSP :** Le pôle SOC/MSSP (Security Operations Center/Managed Security Service Provider) propose des services d'infogérance des solutions de sécurité. Il assure la maintenance, l'administration, le reporting et la supervision des solutions de sécurité telles que les pare-feu, les systèmes de prévention des intrusions, la gestion des logs, les pare-feu applicatifs, les SIEM, etc. Ce pôle offre également des services de gestion des vulnérabilités et opère un Security Operations Center (SOC) pour surveiller et répondre aux incidents de sécurité.



FIGURE 1.3 – CyberSOC DATAPROTECT

**4.Pôle Formation :** Le pôle Formation propose des formations sur différents aspects de la sécurité des données. Il offre des programmes de formation sur la protection des données personnelles, la sécurité de l'information, la gestion des risques IT, les services IT, la sécurité applicative, la continuité d'activité, la préparation aux certifications, la sensibilisation et l'audit. Ce pôle organise également des ateliers pratiques pour permettre aux participants d'acquérir des compétences pratiques en matière de sécurité.

**5.Pôle RD :** Le pôle RD (Recherche et Développement) se concentre sur le développement de solutions sécurisées. Il mène des recherches et une veille des vulnérabilités pour rester à jour avec les dernières menaces et technologies de sécurité. Ce pôle se spécialise également dans le développement sécurisé d'applications souveraines et effectue des tests de sécurité pour les applications critiques.

## 1.5 Le Pôle CyberSOC DATAPROTECT

Le CyberSOC (Centre Opérationnel de Sécurité) est une entité chargée de surveiller et de protéger les systèmes d'information d'une organisation contre les cyberattaques. Son rôle principal est de détecter, analyser et contrer les menaces de sécurité en temps réel, afin de minimiser les risques et de garantir la continuité des opérations.

Le CyberSOC travaille en mode 24/7, ce qui signifie qu'il assure une surveillance permanente des incidents de sécurité. Il dispose d'une équipe d'analystes certifiés et expérimentés, dotés d'une solide expertise dans le domaine de la sécurité informatique. Ces experts sont formés pour détecter les menaces en temps réel, analyser leur nature et prendre des mesures rapides et précises pour y répondre efficacement.



FIGURE 1.4 – Les avantages du Cybersoc

Le CyberSOC utilise des outils et des technologies avancés pour surveiller les événements de sécurité, tels que les journaux d'activité, les alertes de sécurité et les anomalies de trafic. Il dispose également d'une base de connaissances riche en "use cases" (scénarios d'attaques) qui lui permet de traiter les incidents en utilisant les meilleures pratiques et les modes opératoires les plus complexes.

En plus de la surveillance et de la réponse aux incidents, le CyberSOC joue également un rôle important dans la prévention des attaques. Il aide les organisations à identifier les failles de sécurité et à adopter les meilleures stratégies de protection en se basant sur les vulnérabilités identifiées. Le CyberSOC travaille en étroite collaboration avec une équipe CSIRT (Computer Security Incident Response Team), qui est responsable de la gestion et de la réponse aux incidents de sécurité. Cette équipe est mobilisée 24/7 et dispose des compétences nécessaires pour qualifier, remédier et enquêter sur les incidents de sécurité.

## Chapitre 2

# Contexte général du projet

---

Ce chapitre présentera une description générale de la problématique du sujet puis présentera notre objectif à atteindre

## 2.1 Mise en situation

Un stage est une période d'activité durant laquelle un étudiant met en application les enseignements théoriques suivis, dans le cadre d'un projet réalisé dans un organisme d'accueil. Il peut durer de quelques jours à plusieurs mois ce qui nous permet de distinguer 3 types de stages :

- Stage d'observation ou d'initiation : Souvent demandé en début du cursus, ce stage permet d'initier l'étudiant à l'entreprise et prendre connaissance de son mécanisme. Ce stage est généralement de courte durée ( 1 mois)
- Stage d'application : Comme son nom l'indique, ce stage a pour objectif de mettre en application ses connaissances acquises au cours d'une formation. Le stagiaire devrait travailler sur une problématique d'entreprise.
- Stage de fin d'études : Généralement de longue durée ( au moins de 3 mois), le stage de fin d'études permet au stagiaire de mettre en pratique l'ensemble de ses connaissances et savoir-faire acquis. C'est aussi une sorte de période d'essai où l'on teste les stagiaires pour une embauche éventuelle.

## 2.2 Présentation du projet

### 2.2.1 Introduction

Le CyberSOC est chargé de surveiller et de protéger les systèmes d'information d'une organisation contre les cyberattaques. Son rôle principal consiste à détecter, analyser et contrer les menaces de sécurité en temps réel. Une des méthodes utilisées pour atteindre cet objectif est l'analyse des logs provenant des serveurs. Cette analyse permet de prévenir ou de détecter toute cyberattaque visant ces serveurs. Afin de réaliser cette tâche, il est essentiel de concevoir une architecture de Pipeline Résilient qui permet l'analyse des données en temps réel.

### 2.2.2 Problématique

Le problème majeur qui bloque le travail du CyberSOC est la protection des logs provenant des serveurs. Pour cela, il faut créer une architecture de pipeline résiliente qui assure tout d'abord une transmission protégée des logs depuis les serveurs. En cas d'échec de l'un des serveurs, il faut s'assurer de conserver les logs. Ensuite, il est nécessaire d'avoir un stockage résilient qui permet la redondance des données avec une protection maximale. Enfin, il faut analyser les logs afin de les exploiter efficacement

### 2.2.3 Solutions et objectifs

L'objectif principal de ce projet est de concevoir et réaliser une architecture de pipeline de transformation et de stockage, ainsi que d'analyser des logs depuis la source. Pour atteindre cet objectif, nous devons mettre en place différentes étapes clés. Dans l'ensemble, la conception et la réalisation d'une architecture de pipeline de transformation et de stockage de logs nécessitent une compréhension approfondie des besoins du projet, ainsi

que des compétences en programmation, en gestion des données et en analyse. Il est également important de prendre en compte des aspects tels que la sécurité, la scalabilité et la maintenance à long terme de l'architecture.

#### 2.2.4 Etude de faisabilité

L'objectif de cette étude de faisabilité est d'évaluer la possibilité de réaliser cette architecture résiliente qui permet l'analyse des données en temps réel. Voici les étapes clés de cette étude :

1. **Collecte des logs** : Le premier défi consiste à collecter et à analyser efficacement les logs générés par les serveurs. Pour cela, nous avons choisi **Winlogbeat**, qui est un outil permettant de gérer ces logs de manière efficace.
2. **Transmission des données en temps réel** : Parmi les meilleures solutions pour la transmission des données en temps réel, nous recommandons l'utilisation d'**Apache Kafka**. Kafka est une plateforme de streaming de données en temps réel qui permet également de sauvegarder les données en cas d'indisponibilité. En complément de Kafka, nous utiliserons **Logstash** pour connecter Kafka à la base de données Elasticsearch.
3. **Stockage des données** : **Elasticsearch** est l'une des meilleures bases de données pour stocker et protéger les logs. Elle permet de créer une base de données distribuée sur plusieurs serveurs, offrant ainsi une meilleure protection des données grâce à la redondance des données.
4. **Analyse des logs** : Pour l'analyse des logs, nous recommandons l'utilisation de **Kibana**. Kibana est un outil simple à intégrer avec Elasticsearch et permet une meilleure analyse des données grâce à des tableaux de bord et des modèles de machine learning.
5. **Évaluation de la faisabilité technique et financière** : Pour évaluer la faisabilité technique et financière de ce projet, il est important de prendre en compte plusieurs facteurs tels que les ressources matérielles nécessaires, les compétences techniques requises, les coûts associés à la mise en place et à la maintenance de cette architecture, ainsi que les bénéfices attendus en termes d'analyse des données en temps réel.



## 2.3 Planning du stage

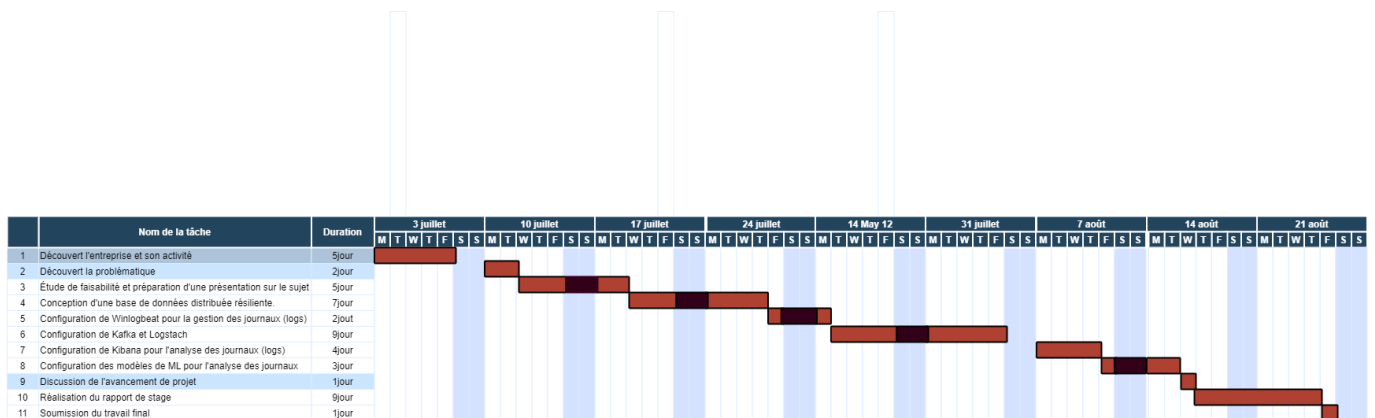


FIGURE 2.1 – Planning du projet

## 2.4 Conclusion

La conception et la réalisation d’une architecture de pipeline résilient pour l’analyse des logs provenant des serveurs sont essentielles pour assurer la sécurité des systèmes d’information d’une organisation. Ce projet vise à résoudre le problème majeur du CyberSOC en garantissant une transmission protégée des logs, une conservation des logs en cas d’échec des serveurs, un stockage résilient avec redondance des données, et une analyse efficace des logs. Pour atteindre ces objectifs, il est crucial de comprendre les besoins du projet, d’avoir des compétences en programmation, en gestion des données et en analyse, tout en tenant compte de la sécurité, de la scalabilité et de la maintenance à long terme de l’architecture. La mise en place de solutions telles que Winlogbeat, Apache Kafka, Logstash, Elasticsearch et Kibana permettra de créer un pipeline robuste pour la surveillance et la protection des systèmes d’information contre les cyberattaques.

## Chapitre 3

# Conception d'Architecture de pipeline

---

Après avoir présenté le contexte général du projet et énoncé ses différents objectifs dans le chapitre précédent, nous entamerons dans le présent chapitre la phase de Conception d'Architecture de pipeline. Cette phase a pour but d'expliquer les étapes de conception d'une telle architecture et de comprendre les raisons derrière ces choix.

## 3.1 Architecture Globale de Pipligne & Technologies utilisés

### 3.1.1 Architecture Globale de Pipligne

L'architecture globale se compose comme suit, tout d'abord, nous avons configuré Winlogbeat, un agent léger conçu pour transférer les logs d'événements vers Apache Kafka, une plateforme distribuée de diffusion de données en continu. Ensuite, Kafka permet de streamer les logs en temps réel vers Logstash, un outil informatique de collecte, analyse et stockage de logs. Logstash permet de connecter Kafka à la base de données Elasticsearch, un logiciel utilisant Lucene pour l'indexation et la recherche de données. Logstash dispose également d'un filtre qui permet de parser les logs afin de les exploiter au niveau de Kibana, un greffon de visualisation de données pour Elasticsearch. Ensuite, les logs sont stockés dans une base de données distribuée entre 3 nuds avec une redondance dans les trois nuds. Enfin, Kibana nous permet d'exploiter ces logs pour prédire s'il y a une cyberattaque.

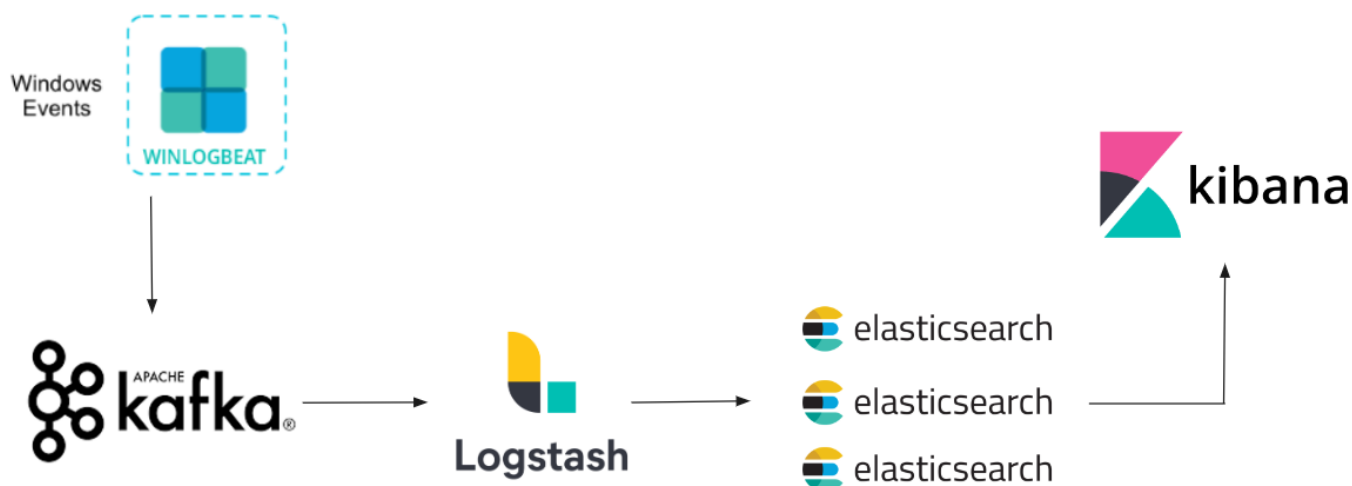


FIGURE 3.1 – Architecture Globale

Pour implémenter l'architecture précédente, des changements se posent sur l'architecture pour l'adapter au mieux. Tout d'abord, nous n'avons pas un seul serveur, mais nous pouvons avoir un cluster de serveurs ou même un seul serveur. Dans ce cas, il faut adapter Kafka pour supporter plusieurs sources, ainsi qu'Elasticsearch pour supporter le stockage des journaux de plusieurs machines.

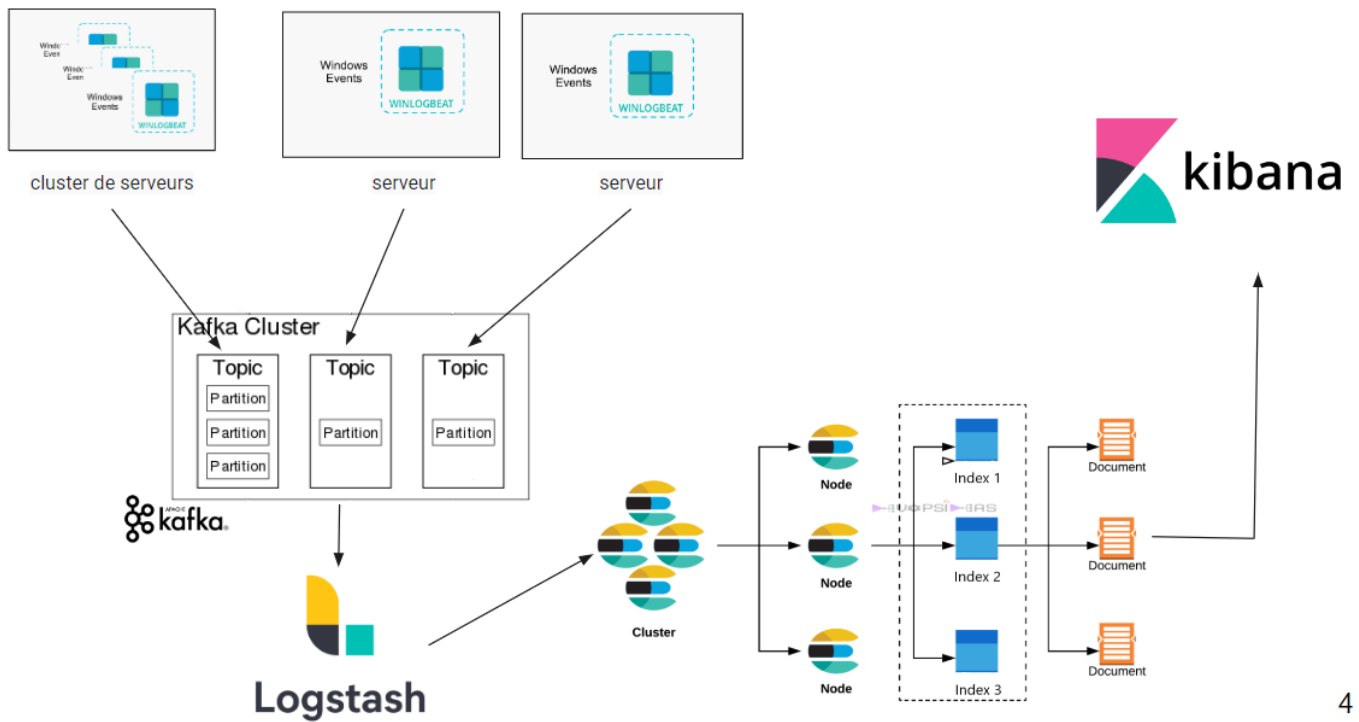


FIGURE 3.2 – Architecture Globale Adjuster

### 3.1.2 Technologies utilisés

**1.Elasticsearch :** Elasticsearch est un moteur de recherche et d'analyse de données puissant et open-source largement utilisé dans le domaine du génie logiciel. Il est construit sur Apache Lucene et offre une solution de recherche et d'analyse distribuée, évolutive et en temps réel.



FIGURE 3.3 – Elasticsearch

L'un des principaux avantages d'Elasticsearch est sa capacité à gérer de grandes quantités de données et à effectuer des recherches complexes de manière efficace. Il utilise une architecture distribuée, ce qui vous permet de le mettre à l'échelle horizontalement en ajoutant plus de nuds au cluster. Cela garantit une disponibilité élevée et une tolérance aux pannes.

**2.Logstash :** Logstash est un outil open-source de collecte, de traitement et d'expédition de données. Il fait partie de la suite Elastic Stack, aux côtés d'Elasticsearch et de Kibana, et est souvent utilisé en combinaison avec Elasticsearch pour l'ingestion de données. Logstash utilise une architecture modulaire et flexible. Il est livré



FIGURE 3.4 – Logstash

avec un large éventail de plugins qui permettent de configurer facilement les différentes étapes du pipeline de traitement des données. Ces étapes comprennent l'ingestion des données, la transformation, l'enrichissement et la sortie vers des destinations telles qu'Elasticsearch, des bases de données ou d'autres systèmes.

**3.Kibana :** Kibana est une plateforme open-source de visualisation et d'exploration de données qui fait partie de la suite Elastic Stack. Kibana permet aux utilisateurs de visualiser et d'analyser les données stockées dans Elasticsearch de manière intuitive et conviviale. Il offre une interface graphique basée sur le web qui permet de créer des tableaux de bord personnalisés, des graphiques, des cartes géographiques, des métriques et bien plus encore.



FIGURE 3.5 – Kibana

**4.Winlogbeat :** Winlogbeat est un agent léger open-source développé par Elastic. Il fait partie de la suite Elastic Stack et est utilisé pour collecter et envoyer des événements de journaux (logs) à Elasticsearch ou à Logstash pour l'analyse et la visualisation ultérieure. Winlogbeat est conçu pour être facile à configurer



FIGURE 3.6 – Winlogbeat

et à déployer. Il utilise une configuration basée sur des fichiers YAML, où vous pouvez spécifier les journaux Windows à surveiller, les filtres à appliquer et les destinations d'envoi des événements de journaux.

**5.Apache Kafka :** Apache Kafka est une plateforme de streaming de données open-source, distribuée et hautement évolutive. Elle est conçue pour gérer efficacement le flux de données en temps réel entre les

applications et les systèmes.



FIGURE 3.7 – Apache Kafka

Kafka est basé sur une architecture de journal de transactions distribué. Il utilise un modèle de publication-souscription, où les producteurs envoient des messages à des sujets (topics) et les consommateurs s'abonnent à ces sujets pour recevoir les messages. Les messages sont stockés de manière persistante dans des partitions, ce qui permet une distribution parallèle et une haute disponibilité.

### 3.2 Architecture de Base de Données

Pour construire une base de données résiliente qui permet de stocker les logs sans les perdre, nous avons choisi de travailler avec une base de données Elasticsearch distribuée sur plusieurs nuds. Chaque nud peut être un serveur indépendant, ce qui offre une meilleure disponibilité et une meilleure tolérance aux pannes.

En utilisant la distribution sur plusieurs nuds, nous répartissons la charge de stockage et de traitement des données, ce qui permet d'obtenir une meilleure performance et une capacité de montée en charge plus élevée. De plus, en cas de défaillance d'un nud, les autres nuds continuent de fonctionner, assurant ainsi la disponibilité des données. Elasticsearch permet également la redondance des données grâce à la réplication. Cela signifie qu'un document peut être stocké plusieurs fois sur plusieurs nuds, assurant ainsi la durabilité des données même en cas de défaillance d'un nud ou d'une partition du réseau.

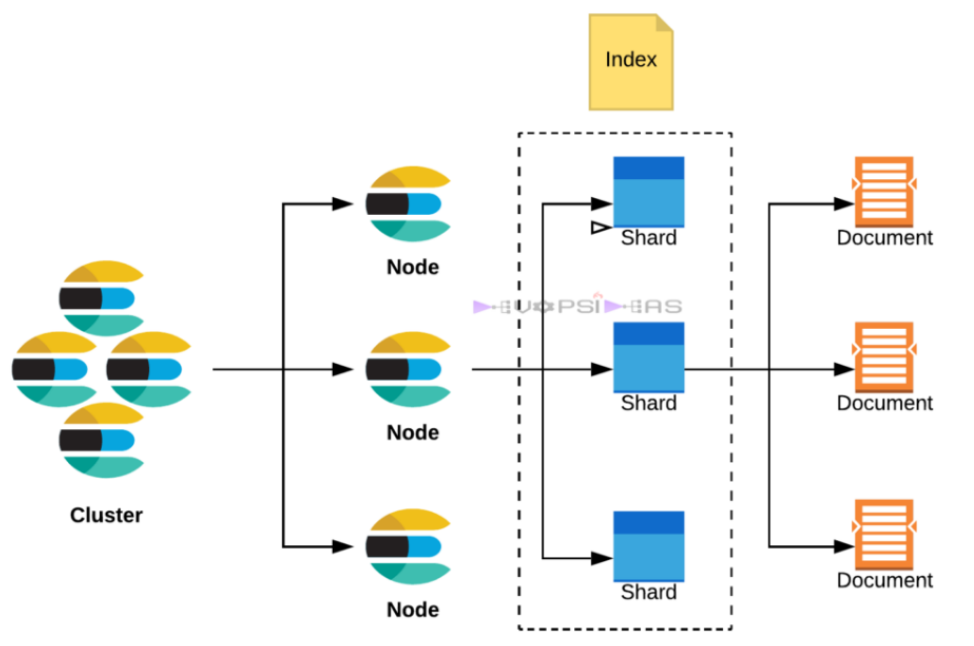


FIGURE 3.8 – Relation de composants Elasticsearch

L'indexation des données sur Elasticsearch est organisée de manière efficace. Les données sont divisées en shards, qui sont des fragments de l'index, et chaque shard est répliqué sur plusieurs nœuds. Cela permet une répartition équilibrée des données et une meilleure parallélisation des opérations de recherche et d'écriture.

### 3.3 Architecture d'Apache Kafka

Apache Kafka est une plateforme de streaming distribuée qui permet de gérer et de traiter des flux de données en temps réel. Il est conçu pour être hautement évolutif, durable et tolérant aux pannes. Kafka est largement utilisé dans les architectures de microservices, les systèmes de traitement des données en continu et les applications de streaming.

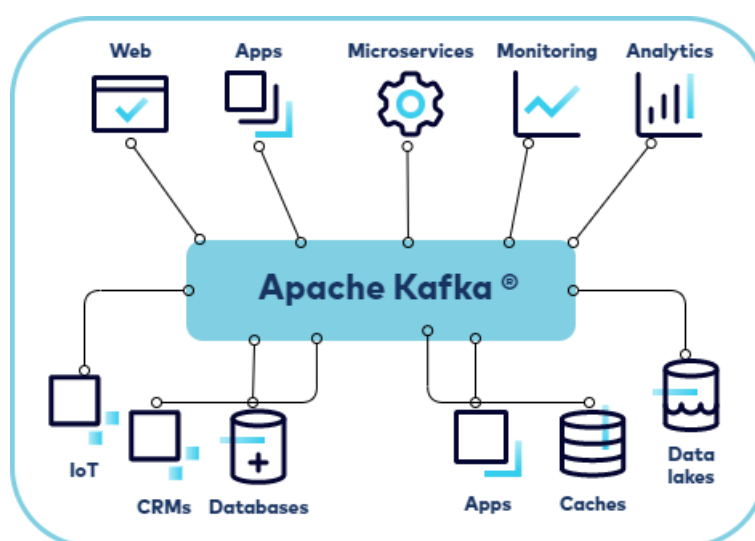


FIGURE 3.9 – Le concept d'Apache Kafka

Le fonctionnement d'Apache Kafka repose sur quelques concepts clés :

1. Producteurs (producers) : Les producteurs sont responsables de l'envoi des données vers les topics de Kafka. Ils peuvent être des applications ou des systèmes qui génèrent des données en temps réel.
2. Topics : Les topics sont des catégories ou des canaux de données dans Kafka. Les producteurs envoient des messages vers des topics spécifiques, et les consommateurs lisent ces messages à partir des topics.
3. Consommateurs (consumers) : Les consommateurs sont des applications ou des systèmes qui lisent les données à partir des topics de Kafka. Ils peuvent être configurés pour lire les données depuis le début ou à partir d'un point spécifique dans le temps.
4. Groupes de consommateurs (consumer groups) : Les consommateurs peuvent être regroupés en groupes de consommateurs. Chaque groupe de consommateurs lit les messages d'un topic donné de manière parallèle, ce qui permet d'augmenter le débit de traitement des données.
5. Brokers : Les brokers sont les serveurs qui stockent et gèrent les données dans Kafka. Ils sont responsables de la réplication des données, de la gestion des partitions et de la coordination entre les producteurs et les consommateurs.

Lorsqu'un producteur envoie un message à un topic, Kafka le stocke dans un journal d'entrées (log) persistant. Les messages sont ensuite répartis en partitions, qui sont distribuées sur les brokers. Chaque partition est ordonnée et chaque message reçoit un offset unique pour permettre un traitement séquentiel.

Les consommateurs peuvent lire les messages à partir des partitions en utilisant des offsets. Ils peuvent également spécifier leur position de lecture dans le journal d'entrées, ce qui leur permet de reprendre la lecture à partir d'un point spécifique en cas de panne ou de redémarrage.

Sur le plan technique, nous avons travaillé avec deux architectures différentes. La première consiste en un seul topic avec une seule partition, puisqu'il n'y a qu'un seul serveur qui envoie les logs. Pour un déploiement client, Apache Kafka peut être configuré de manière à comporter plusieurs topics. Chaque topic peut contenir un ou plusieurs partitions, en fonction du nombre de serveurs et de l'architecture de ces serveurs.

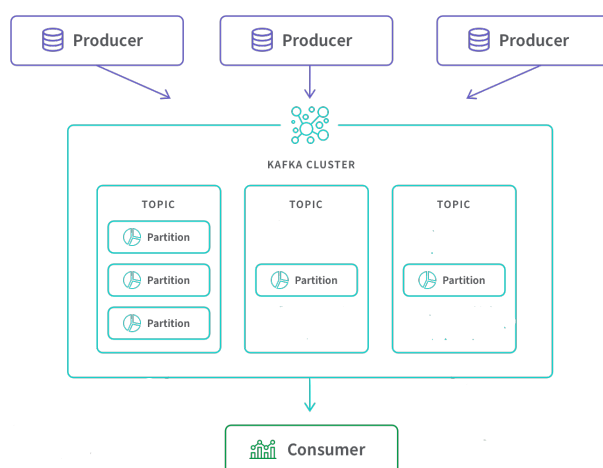


FIGURE 3.10 – Architecture Kafka Ajuster

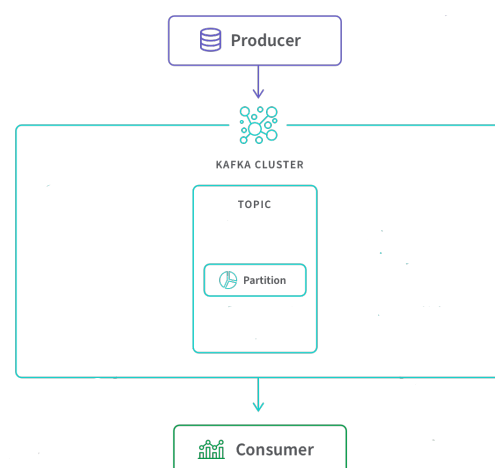


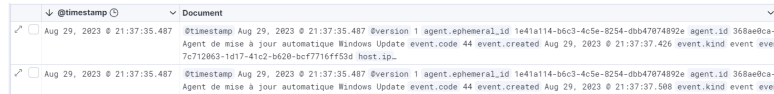
FIGURE 3.11 – Architecture Kafka Utilise



### 3.4 Analyse des logs

Les logs, ou journaux d'événements, sont des enregistrements chronologiques des activités qui se produisent dans un système informatique, telles que les erreurs, les avertissements, les informations de débogage, les actions effectuées par les utilisateurs, etc.

Les logs peuvent être générés par différents composants d'un système, tels que les applications, les serveurs, les bases de données, les services réseau, etc. Ils sont généralement stockés dans des fichiers texte ou des bases de données, et peuvent être consultés à l'aide d'outils spécifiques.



Timestamp	Document
Aug 29, 2023 @ 21:37:35.487	@timestamp Aug 29, 2023 @ 21:37:35.487 @version 1 agent.ephemeral_id 1e41a114-b6c3-4c5e-8254-dbb47074892e agent.id 368ae0ca-e000-47ae-8253-754790db6f27 event.action Agent de mise à jour automatique Windows Update event.code 44 event.created Aug 29, 2023 @ 21:37:37.426 event.kind event event.level information event.name Agent de mise à jour automatique Windows Update event.version 1
Aug 29, 2023 @ 21:37:35.487	@timestamp Aug 29, 2023 @ 21:37:35.487 @version 1 agent.ephemeral_id 1e41a114-b6c3-4c5e-8254-dbb47074892e agent.id 368ae0ca-e000-47ae-8253-754790db6f27 event.action Agent de mise à jour automatique Windows Update event.code 44 event.created Aug 29, 2023 @ 21:37:37.588 event.kind event event.level information event.name Agent de mise à jour automatique Windows Update event.version 1

FIGURE 3.12 – Exemples des logs

```
1 {
2   "_index": "winlogbeat_kafka_logstash_index",
3   "_id": "TJfNQooB0wVHvg34Mtng",
4   "_version": 1,
5   "_score": 0,
6   "_source": {
7     "ecs": {
8       "version": "8.0.0"
9     },
10    "log": {
11      "level": "information"
12    },
13    "agent": {
14      "id": "368ae0ca-e000-47ae-8253-754790db6f27",
15      "type": "winlogbeat",
16      "ephemeral_id": "1e41a114-b6c3-4c5e-8254-dbb47074892e",
17      "version": "8.8.2",
18      "name": "DESKTOP-4RQK3GG"
19    },
20    "@version": "1",
21    "event": {
22      "action": "Agent de mise à jour automatique Windows Update",
23      "kind": "event",
24      "code": "44".
```

FIGURE 3.13 – Quelques informations de journal

Dans les journaux, qui sont des documents au format JSON, plusieurs informations sont enregistrées. Par exemple, l'identifiant, la version du journal, le score, et l'événement. Toutes ces informations servent à identifier les erreurs ou les cyberattaques, que ce soit manuellement ou en utilisant des algorithmes de machine learning. Ces algorithmes sont principalement utilisés pour détecter les anomalies, telles que :

- **High Mean** : Cet algorithme identifie les anomalies en détectant les valeurs qui sont significativement plus élevées que la moyenne.
- **Low Mean** : Il détecte les anomalies en identifiant les valeurs qui sont significativement plus basses que la moyenne.
- **Rare** : Cet algorithme se concentre sur l'identification des occurrences rares ou des événements qui s'écartent de la norme.
- **Multivariate** : Il analyse plusieurs champs simultanément pour détecter les anomalies en fonction de leur comportement combiné.

Ces algorithmes de détection d'anomalies sont essentiels pour surveiller et protéger les systèmes informatiques contre les comportements indésirables ou malveillants. Ils permettent de repérer rapidement les problèmes potentiels et de prendre des mesures appropriées pour y remédier.

### **3.5 Conclusion**

l'architecture décrite dans ce chapitre offre une solution complète pour la collecte, l'analyse et le stockage des logs d'événements. En utilisant des technologies telles que Winlogbeat, Apache Kafka, Logstash, Elasticsearch et Kibana, cette architecture permet de traiter les logs en temps réel, de les stocker de manière distribuée et de les visualiser de manière conviviale. Grâce à cette architecture, les entreprises peuvent surveiller efficacement leurs systèmes informatiques, détecter les anomalies et les cyberattaques potentielles, et prendre des mesures appropriées pour y remédier. La capacité d'évolutivité et de tolérance aux pannes offerte par des technologies telles que Kafka et Elasticsearch garantit une disponibilité élevée des données et une performance optimale, même dans des environnements à grande échelle.

## Chapitre 4

# Réalisation du projet

---

Après avoir mené les phases précédentes, passant par la phase de la spécification et d'analyse, suivies par les phases de la conception détaillée et de l'étude technique, étape suivante sera consacrée à la réalisation du projet.

## 4.1 Configuration de Pipeline

### 4.1.1 Configuration d'Elasticsearch & Kibana

Pour commencer, nous devons installer Elasticsearch sur notre machine. En effet, nous avons besoin d'installer 3 dossiers Elasticsearch, chacun se comportant comme un nud indépendant. Dans chaque dossier, vous trouverez plusieurs dossiers importants dans le répertoire d'installation. Voici une brève description de ces dossiers :



node-1	08/08/2023 00:33	Dossier de fichiers	
node-2	08/08/2023 00:40	Dossier de fichiers	
node-3	08/08/2023 00:47	Dossier de fichiers	

bin	26/06/2023 05:23	Dossier de fichiers	
config	08/08/2023 13:37	Dossier de fichiers	
data	30/08/2023 12:27	Dossier de fichiers	
jdk	26/06/2023 05:23	Dossier de fichiers	
lib	26/06/2023 05:23	Dossier de fichiers	
logs	30/08/2023 00:15	Dossier de fichiers	
modules	26/06/2023 05:23	Dossier de fichiers	
plugins	26/06/2023 05:17	Dossier de fichiers	
LICENSE	26/06/2023 05:15	Document texte	4 Ko
NOTICE	26/06/2023 05:17	Document texte	2 199 Ko
README.asciidoc	26/06/2023 05:15	Fichier ASCIIDOC	8 Ko

FIGURE 4.1 – Composent de Dossier Elasticsearch

1. **Le dossier "bin"** : Ce dossier contient les fichiers binaires exécutables d'Elasticsearch. Vous y trouverez des scripts pour démarrer, arrêter et gérer votre cluster Elasticsearch.
2. **Le dossier "config"** : Ce dossier contient les fichiers de configuration d'Elasticsearch. Vous pouvez y trouver le fichier `elasticsearch.yml`, qui est le fichier principal de configuration où vous pouvez spécifier des paramètres tels que le port d'écoute, les chemins des fichiers de données, etc.
3. **Le dossier "data"** : Ce dossier est utilisé par Elasticsearch pour stocker les données indexées. Chaque nud Elasticsearch aura son propre sous-dossier dans "data". Il est important de sauvegarder régulièrement ce dossier pour éviter toute perte de données.
4. **Le dossier "logs"** : Comme son nom l'indique, ce dossier contient les fichiers journaux d'Elasticsearch. Ces fichiers enregistrent les activités du cluster, les erreurs, les avertissements, etc. Ils sont utiles pour le débogage et la surveillance du cluster.
5. **Le dossier "plugins"** : Ce dossier est utilisé pour stocker les plugins Elasticsearch. Les plugins sont des extensions qui ajoutent des fonctionnalités supplémentaires à Elasticsearch, comme l'intégration avec d'autres systèmes ou l'ajout de nouveaux types de champs.
6. **Le dossier "scripts"** : Ce dossier est utilisé pour stocker les scripts Elasticsearch. Les scripts peuvent être utilisés pour effectuer des opérations avancées sur les données, comme les requêtes complexes ou les mises à jour conditionnelles.

Ces dossiers sont essentiels pour le bon fonctionnement d'Elasticsearch. Il est important de comprendre leur rôle et de les gérer correctement pour garantir la stabilité et les performances de votre cluster Elasticsearch."

Ensuite, il faut configurer les trois nuds pour qu'ils se comportent comme un seul cluster. Pour ce faire, il faut accéder au fichier `elasticsearch.yml` dans le dossier `config` et modifier certains paramètres, à savoir :

- **cluster.name : my-application** : Cette configuration définit le nom du cluster Elasticsearch. Chaque cluster doit avoir un nom unique pour éviter toute confusion avec d'autres clusters.
- **node.name : node-1** : Cette configuration spécifie le nom du nud Elasticsearch. Chaque nud dans un cluster doit avoir un nom unique pour faciliter l'identification et la gestion.
- **network.host : [local]** : Cette configuration spécifie l'adresse IP ou l'interface réseau sur laquelle Elasticsearch écoute les connexions entrantes. La valeur [local] signifie que le nud Elasticsearch n'accepte que les connexions locales.
- **http.port : 9200** : Cette configuration définit le port sur lequel Elasticsearch écoute les requêtes HTTP. Par défaut, Elasticsearch utilise le port 9200 pour les requêtes REST.
- **transport.port : 9300** : Cette configuration spécifie le port utilisé pour les communications internes entre les nuds Elasticsearch. Par défaut, Elasticsearch utilise le port 9300 pour les communications de cluster.
- **discovery.seed\_hosts : ["127.0.0.1 :9300"]** : Cette configuration spécifie les adresses IP des noeuds du cluster à utiliser pour la découverte initiale des autres nuds. Dans cet exemple, seul le noeud local est utilisé pour la découverte.
- **xpack.security.enabled : false** : Cette configuration désactive la fonctionnalité de sécurité X-Pack d'Elasticsearch, qui fournit des fonctionnalités telles que l'authentification et l'autorisation.
- **xpack.security.enrollment.enabled : true** : Cette configuration active la fonctionnalité d'inscription de sécurité X-Pack, qui permet d'enregistrer des nuds Elasticsearch auprès d'un cluster sécurisé.
- **xpack.security.http.ssl** : Ces configurations permettent d'activer le chiffrement SSL pour les connexions client HTTP, telles que Kibana, Logstash et les agents. Dans cet exemple, le chiffrement SSL est désactivé.
- **xpack.security.transport.ssl** : Ces configurations permettent d'activer le chiffrement SSL pour les communications internes entre les nuds Elasticsearch. Le mode de vérification est défini sur "certificate" et les chemins des certificats sont spécifiés pour le keystore et le truststore.
- **cluster.initial\_master\_nodes : ["node-1"]** : Cette configuration spécifie les nuds qui peuvent être élus comme maîtres initiaux lors de la formation du cluster. Dans cet exemple, seul le nud "node-1" est éligible.
- **http.host : 0.0.0.0** : Cette configuration spécifie l'adresse IP sur laquelle Elasticsearch écoute les connexions HTTP. La valeur "0.0.0.0" signifie que le nud Elasticsearch accepte les connexions de toutes les adresses IP.

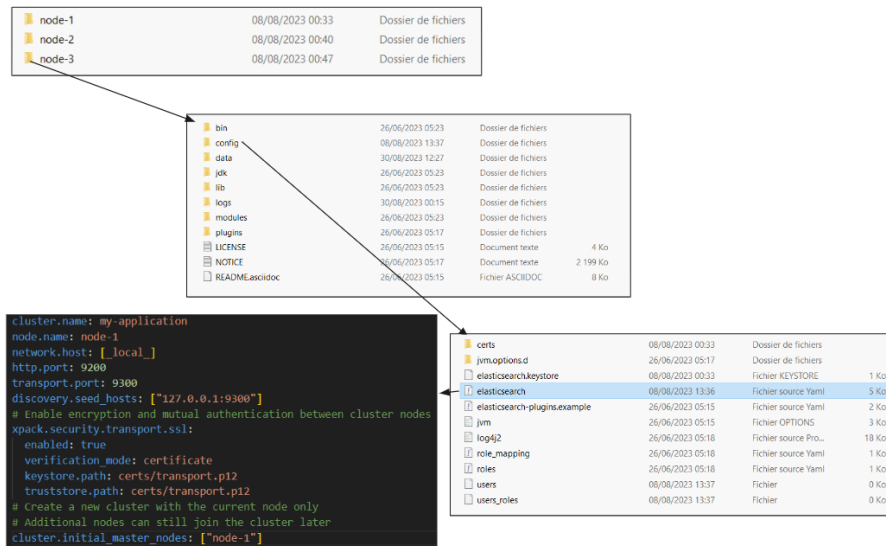


FIGURE 4.2 – Yml Elasticsearch Configuration

Après avoir configuré Elasticsearch, nous passons à la configuration de Kibana. Tout d'abord, nous devons configurer quelques paramètres, à savoir :

- **server.port : 5601** : Ce paramètre spécifie le port sur lequel Kibana sera accessible. Dans cet exemple, Kibana sera accessible via le port 5601. Vous pouvez modifier ce paramètre si vous souhaitez utiliser un port différent.
- **elasticsearch.hosts ["http://localhost:9200"]** : Ce paramètre spécifie les hôtes Elasticsearch auxquels Kibana se connectera. Dans cet exemple, Kibana se connectera à Elasticsearch exécuté localement sur la machine où Kibana est installé, via l'URL textbf"http://localhost:9200". Vous pouvez spécifier plusieurs hôtes Elasticsearch en utilisant une liste, par exemple : textbf["http://localhost:9200", "http://autre-hote-elasticsearch:9200"].
- **http.host : 0.0.0.0** : Cette configuration spécifie l'adresse IP sur laquelle Elasticsearch écoute les connexions HTTP. La valeur "0.0.0.0" signifie que le nud Elasticsearch accepte les connexions de toutes les adresses IP.

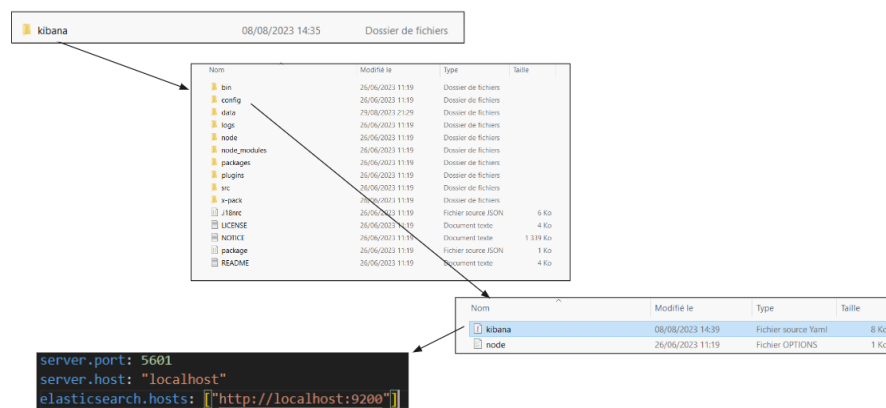


FIGURE 4.3 – Yml Kibana Configuration

Ensuite, passez à la dernière étape qui consiste à utiliser des tokens pour intégrer Elasticsearch et Kibana.

1. Générer un jeton d'accès : Tout d'abord, vous devez générer un jeton d'accès qui sera utilisé pour l'authentification entre Kibana et Elasticsearch. Le jeton peut être généré en utilisant différentes méthodes, telles que l'utilisation de l'API Elasticsearch ou un fournisseur d'authentification tiers.
2. Configurer Elasticsearch : Dans le fichier de configuration d'Elasticsearch (**elasticsearch.yml**), vous devez activer l'authentification basée sur les jetons en définissant la propriété **xpack.security.authc.api\_key.enabled** sur **true**. Cela permet à Elasticsearch d'accepter l'authentification basée sur les clés API.
3. Configurer Kibana : Dans le fichier de configuration de Kibana (**kibana.yml**), vous devez spécifier l'URL d'Elasticsearch et le jeton d'accès. Définissez la propriété **elasticsearch.hosts** sur l'URL de votre cluster Elasticsearch, et définissez les propriétés **elasticsearch.username** et **elasticsearch.password** sur le jeton d'accès généré à l'étape 1.
4. Redémarrer Elasticsearch et Kibana : Après avoir effectué les modifications de configuration nécessaires, redémarrez à la fois Elasticsearch et Kibana pour appliquer les nouveaux paramètres.
5. Vérifier la configuration : Une fois les services redémarrés, vous pouvez vérifier la configuration en accédant à Kibana via votre navigateur web. Si la configuration est correcte, vous devriez pouvoir vous connecter à Kibana en utilisant le jeton d'accès et accéder aux données Elasticsearch.

#### 4.1.2 Configuration de Winlogbeat & Apache Kafka et Logstash

Pour configurer Winlogbeat, vous devez accéder au fichier winlogbeat.yml et configurer certains paramètres. Cette configuration de Winlogbeat est utilisée pour collecter des journaux d'événements à partir de différentes sources sur un système Windows. Permettez-moi de vous expliquer chaque partie de cette configuration :

1. **winlogbeat.event\_logs** : Cette section spécifie les journaux d'événements à collecter. Dans cet exemple, nous collectons les journaux d'événements des applications, du système, de la sécurité, de Sysmon, de Windows PowerShell et de PowerShell opérationnel. Certains journaux d'événements spécifiques sont filtrés en utilisant les identifiants d'événement.
2. **setup.template.settings** : Cette section définit les paramètres du modèle d'index utilisé pour stocker les données collectées. Dans cet exemple, nous définissons le nombre de fragments de l'index sur 1.  
**setup.dashboards.enabled** : Cette option active la création des tableaux de bord Kibana pour visualiser les données collectées.
3. **processors** : Cette section spécifie les processeurs qui seront appliqués aux événements collectés. Dans cet exemple, nous ajoutons des métadonnées d'hôte et des métadonnées cloud aux événements. Les métadonnées d'hôte fournissent des informations sur l'hôte à partir duquel les événements sont collectés, tandis que les métadonnées cloud fournissent des informations sur l'environnement cloud dans lequel l'hôte est déployé.
4. **output.kafka** : Cette section spécifie la sortie des événements collectés vers un cluster Kafka. Dans cet exemple, nous spécifions les brokers Kafka à utiliser, le sujet dans lequel les événements seront publiés, et d'autres paramètres tels que le mode de partitionnement et la compression des messages.

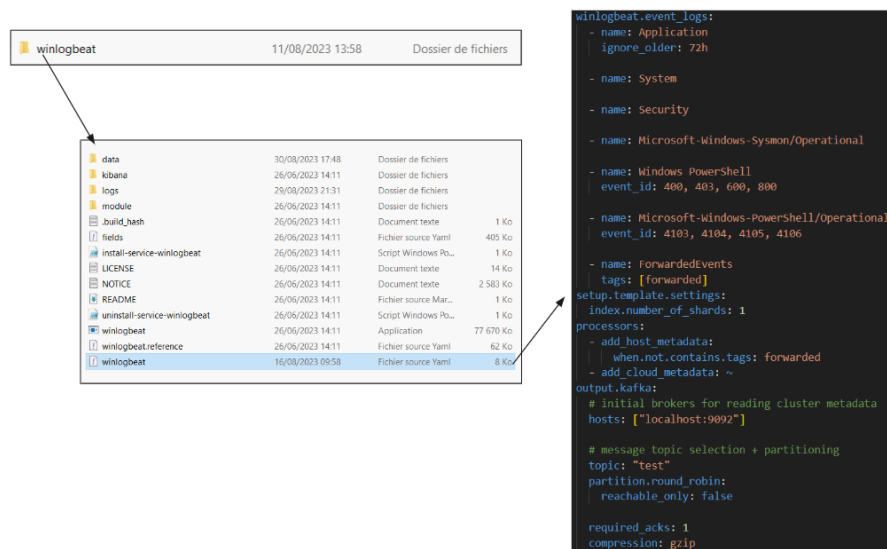


FIGURE 4.4 – Yml Winlogbeat Configuration

Dans le cadre de la configuration de Kafka, il est nécessaire de créer un topic. Dans notre cas, nous l'appelons 'test' et il contient une seule partition. Pour réaliser cette opération, vous pouvez utiliser la commande suivante :

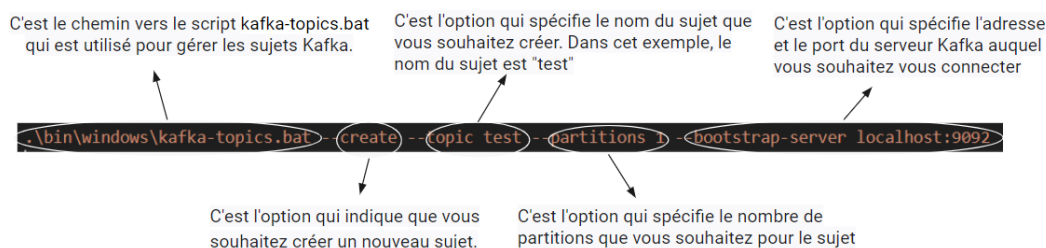


FIGURE 4.5 – Kafka Configuration

Pour lier Kafka à Elasticsearch, nous devons configurer Logstash en définissant trois éléments principaux : l'input, l'output et le filtre.

- **L'input** : Cette section spécifie la source des données que Logstash collectera. Dans notre exemple, nous utilisons l'input Kafka pour récupérer les messages à partir du serveur Kafka local (**localhost :9092**) et du topic **"test"**.
- **Le filtre** : Cette section permet de traiter les données collectées avant de les envoyer à l'output. Nous utilisons le filtre JSON pour extraire les données JSON du champ **"message"**. Cela nous permet de structurer les logs avant de les envoyer à Elasticsearch.
- **L'output** : Cette section définit la destination des logs traités. Dans notre exemple, nous utilisons l'output Elasticsearch pour envoyer les logs vers un cluster Elasticsearch local. Les adresses des hôtes Elasticsearch sont spécifiées avec les URLs **"http ://localhost :9200"**, **"http ://localhost :9201"** et **"http ://localhost :9202"**. L'index utilisé est **"winlogbeat\_kafka\_logstash\_index"**. De plus, la vérification du certificat SSL est désactivée pour simplifier la configuration.



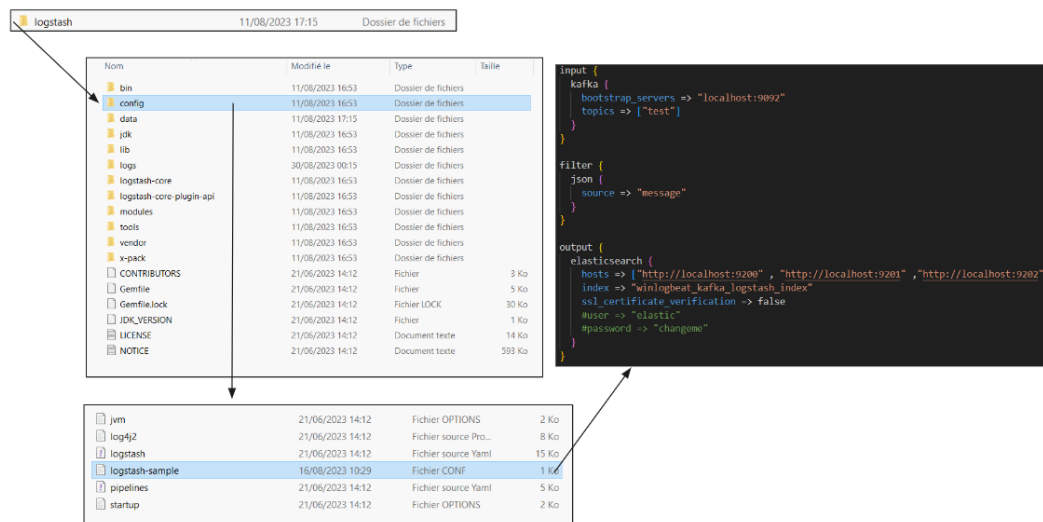


FIGURE 4.6 – Logstash Configuration

## 4.2 Visualisation des résultats

Après la configuration du projet, nous passons à la phase expérimentale. Pour lancer l'ensemble du projet, il faut suivre ces commandes. Tout d'abord, nous commençons par lancer la base de données Elasticsearch, nous lançons les 3 nuds qui forment le cluster. Ensuite, nous lançons Kibana. Ensuite, nous lançons Winlogbeat pour générer les logs. Enfin, nous lançons Apache Kafka et Logstash pour la transmission des données.

```

1- start elasticsearch :
  1.1- Node1 :
    cd C:\elstic\node-1\bin
    elasticsearch.bat
  1.2- Node2
    cd C:\elstic\node-2\bin
    elasticsearch.bat
  1.3- Node3
    cd cd C:\elstic\node-3\bin
    elasticsearch.bat
2- Start kibana :
  cd C:\elstic\kibana\bin
  kibana.bat
3- Start Winlogbeat :
  cd C:\elstic\winlogbeat
  .\winlogbeat.exe -c .\winlogbeat.yml
4- Start kafka :
  cd C:\elstic\kafka
  .\bin\windows\zookeeper-server-start.bat .\config\zookeeper.properties
  .\bin\windows\kafka-server-start.bat .\config\server.properties
5- Test kafka :
  .\bin\windows\kafka-console-consumer.bat --topic test --bootstrap-server localhost:9092 --from-beginning
6- Start logstash :
  cd C:\elstic\logstash
  bin\logstash -f .\config\logstash-sample.conf

```

FIGURE 4.7 – Commandes pour lancer la pipeline

Une fois la pipeline lancée, vous pouvez accéder à Kibana pour visualiser vos logs sur le port suivant : **"localhost :5601"**. Kibana vous permet de visualiser vos logs sous forme de JSON ou de tableau, triés par ordre chronologique. De plus, vous avez la possibilité de filtrer les logs en fonction de n'importe quel champ.

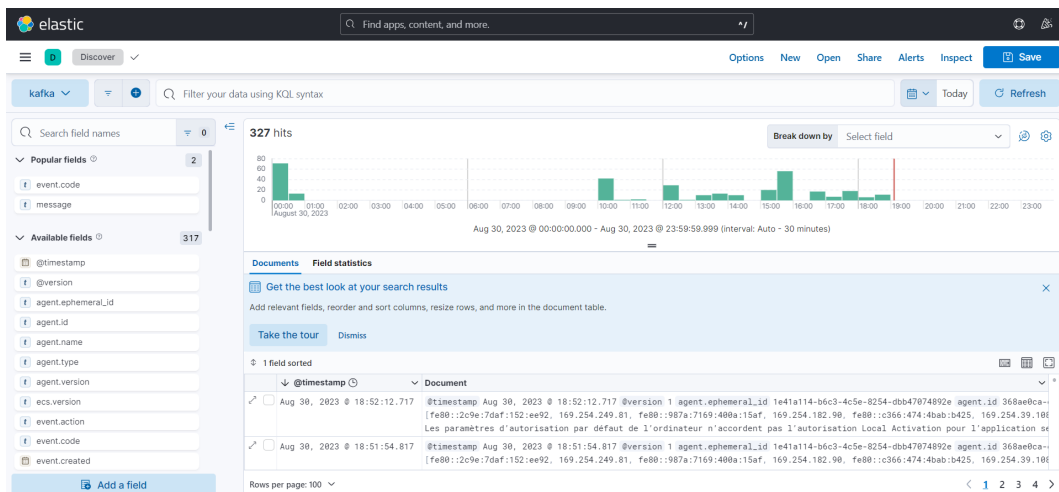


FIGURE 4.8 – Visualisez les journaux

Sur Kibana, vous avez la possibilité de configurer des jobs spécifiques, tels que "**Anomaly Detection Job**", qui permet d'analyser les logs à la recherche d'indices anormaux pouvant être liés à des cyberattaques. Pour cela, nous utilisons des modules de machine learning. Sur la figure suivante, vous pouvez voir quelques modules avec une description de leur rôle.

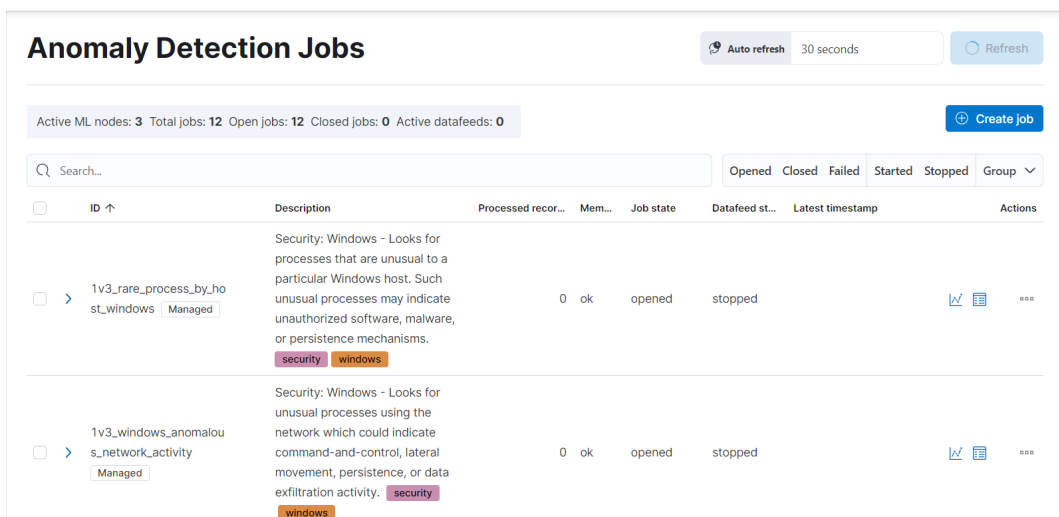


FIGURE 4.9 – Anomaly Detection Jobs

Après la configuration des jobs, Kibana offre la possibilité d'explorer les résultats et de visualiser tous les indices ou les logs qui pourraient être suspects et provenir d'une attaque. Veuillez consulter la figure suivante (il n'y a pas d'attaque ou d'indice d'une cyberattaque).

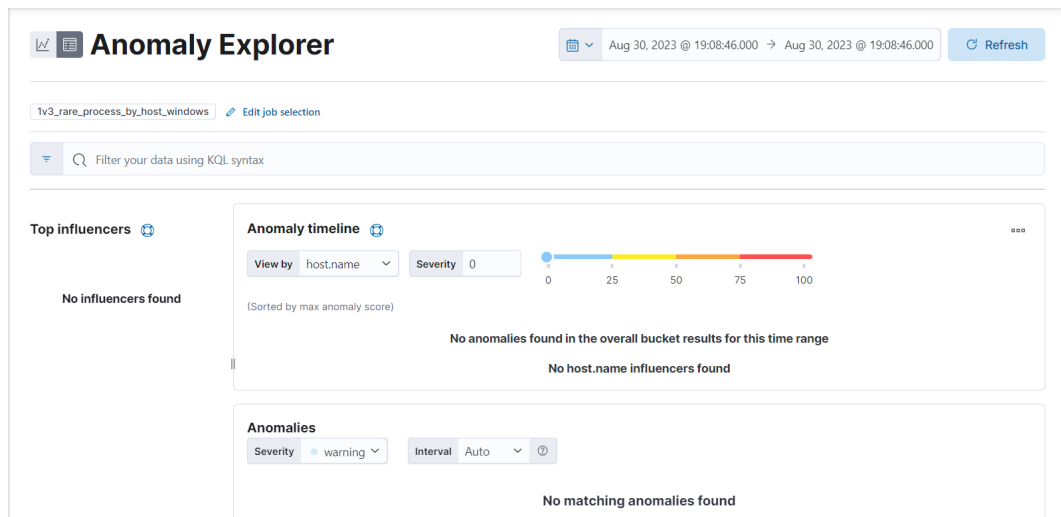


FIGURE 4.10 – Anomaly Explorer

## Chapitre 5

# Conclusion et Perspectives

En conclusion, la conception et la mise en place d'une architecture de pipeline résilient pour l'analyse des logs ont permis d'améliorer la sécurité des systèmes d'information. Grâce à l'utilisation d'outils tels qu'Elasticsearch, Logstash et Kibana, les entreprises peuvent collecter, stocker, analyser et visualiser leurs logs de manière efficace et en temps réel. Cela leur permet de détecter rapidement les anomalies, les cyberattaques et de prendre les mesures nécessaires pour protéger leurs systèmes.

Ce projet a démontré l'importance d'une approche proactive en matière de sécurité de l'information. En mettant en place une surveillance continue des logs et en utilisant des techniques d'analyse avancées, les entreprises peuvent renforcer leur posture de sécurité et réduire les risques liés aux cybermenaces. Il est essentiel de rester à jour avec les dernières technologies et les meilleures pratiques en matière de sécurité pour garantir une protection efficace des systèmes informatiques.

Ce stage m'avait non seulement introduit dans l'univers professionnel, surtout sur le plan technique, dans la mesure où j'ai eu l'occasion d'approfondir mes connaissances théoriques acquises tout au long de l'année scolaire, mais encore sur le plan relationnel et humain.

# Bibliographie

- [1] <https://www.elastic.co/guide/en/elasticsearch/reference/8.2/index.html>.
- [2] <https://www.elastic.co/guide/en/beats/winlogbeat/8.2/index.html>.
- [3] <https://www.elastic.co/guide/en/logstash/current/logstash-8-8-2.html>.
- [4] <https://www.elastic.co/guide/en/kibana/current/release-notes-8.8.2.html>.
- [5] <https://kafka.apache.org/documentation/>.