République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Saad Dahlab de Blida

Mémoire de fin d'études

Pour l'obtention du diplôme de master en informatique

Option : Sécurité des Systèmes d'Information

Thème

Étude et Mise en place d'une solution SIEM Au sein de la société NAFTAL

Organisme d'accueil : NAFTAL

Réalisé par :

Bouraoua Mohamed Amine

Hallel Yacine Wassim

Présenté devant la commission d'examination composée de :

- *Mme.* ***Meskaldji*** *(Promotrice)*
- *M.* ***Mansouri Djamel*** *(Encadrant)*
- *Mme.* ***Oukid Lamia*** *(Présidente)*
- *Mme.* ***Hadj Henni*** *(Examinatrice)*

Année universitaire 2021/2022

Remerciements

Tout d'abord, nous remercions Dieu le Tout-Puissant, de nous avoir donné la volonté et le courage de réaliser ce travail.

Un grand merci à notre promotrice, ***Mme Kh.MESKALDJI****, qui nous a beaucoup aidé dans la correction et la rédaction de ce document, et aussi pour sa patience, sa disponibilité, et pour avoir accepté de mener ce travail.*

Nous exprimons notre gratitude à tout le personnel de la société ***NAFTAL*** *pour leur collaboration, en particulier les personnes qui nous ont fourni les éléments nécessaires à la réalisation de notre projet. Nous sommes particulièrement reconnaissants à monsieur* ***MANSOURI Djamel*** *et* ***GUELMAOUI Mohamed Amine*** *pour avoir accepté de nous confier cette mission et aussi pour leur aide et leurs précieux conseils ainsi que leur disponibilité durant le stage.*

Nous remercions sincèrement les membres du Jury de nous avoir fait l'honneur d'accepter et d'évaluer notre travail et tout le corps professoral du département d'informatique de l'Université Saad Dahlab Blida1 pour les efforts fournis dans notre formation.

Enfin, nous voudrions exprimer notre profonde gratitude et nos vrais sentiments à nos familles, qui nous ont toujours soutenus.

DÉDICACES Amine

Je dédie ce modeste travail :

À Mes parents, pour leur amour infini, leurs sacrifices, leur soutien et leurs encouragements,

À Ma chère grand-mère maternelle et à la mémoire de mon grand-père, que Dieu l'accueille dans son vaste paradis,

À la mémoire de mes chers Grands-parents paternels, que Dieu les accueille dans son vaste paradis,

À mon cher frère et ma chère sœur,

À tous les membres de ma famille ainsi que mes amis,

À tous ceux qui me sont chers et à toutes les personnes qui m'ont aidé à atteindre ce niveau.

AMINE

DÉDICACES Wassim

Je dédie ce modeste travail :

À mes chers parents, pour tous leurs sacrifices, leur amour, leur tendresse, leur soutien et leurs prières tout au long de mes études,

À mes chères sœurs lyna, sofia et nada pour leurs encouragements permanents, et leur soutien moral,

À mes chers frères Abdelhamid et Moumen, pour leur support et leurs encouragements,

À toute ma famille pour leur soutien tout au long de mon parcours universitaire, Que ce travail soit l'accomplissement de vos vœux, et le fruit de votre soutien infaillible,

Merci d'être toujours là pour moi.

Wassim

Résumé

Cette étude propose un système de collecte et d'analyse d'événements SIEM (Security Information and Event Management), un outil d'aide à la gestion centralisée des fichiers journaux, ces fichiers sont les parties que les entreprises analysent. Car ces derniers facilitent le suivi de l'activité des utilisateurs dans le SI (système d'information), ce système est équipé de tableau de bord qui permet un suivi en temps réel des journaux analysés et génère des alertes de sécurité en cas de détection d'anomalies afin de suivre l'utilisation d'Ubuntu ou de Windows à l'aide de différents agents SIEM installés.

Les journaux des systèmes d'exploitation Windows et Ubuntu sont collectés par un agent, et les données brutes sont ensuite envoyées à un serveur de traitement pour normalisation et analyse. Après traitement, les journaux sont indexés et enregistrés dans une base de données No-SQL afin que les outils de lecture puissent les afficher sous forme de graphiques et de tableaux intégrés dans des tableaux de bord. Après avoir examiné plusieurs journaux pertinents liés à la sécurité, une alerte est générée lorsqu'une tentative d'attaque (telle qu'une attaque par force brute ou par injection) est découverte.

Mots clés : SIEM, fichiers journaux, SI, événements, tableau de bord, alertes, anomalies, No-SQL, attaque, analyse d'événements, indexation.

Abstract

This study proposes a SIEM (Security Information and Event Management) event collection and analysis system, a tool to help with the centralized management of log files, these files are the parts that companies analyzed. Because these logs facilitate the monitoring of user activity in the IS (information system), this system is equipped with a dashboard that allows real-time monitoring of analyzed logs and generates security alerts when anomalies are detected in order to track the usage of Ubuntu or Windows using different installed SIEM agents.

Windows and Ubuntu operating system logs are collected by an agent, and the raw data are then sent to a processing server for normalization and analysis. After processing, the logs are indexed and saved in a No-SQL database so that reading tools can display them as graphs and tables integrated into dashboards. After reviewing several relevant security-related logs, an alert is generated when an attack attempt (such as a brute force or injection attack) is discovered.

Keywords: SIEM, log files, IS, events, dashboard, alerts, anomalies, No-SQL, attack, event analysis, indexing.

ملخص

تقترح هذه الدراسة نظام تجميع وتحليل أحداث SIEM (Security Information and Event Management)، وهي أداة للمساعدة في الإدارة المركزية لملفات السجل، وهذه الملفات هي الأجزاء التي قامت الشركات بتحليلها. نظرًا لأن هذه السجلات تسهل مراقبة نشاط المستخدم في IS (نظام المعلومات)، فقد تم تجهيز هذا النظام بلوحة معلومات تتيح المراقبة في الوقت الفعلي للسجلات التي تم تحليلها وإنشاء تنبيهات أمنية عند اكتشاف حوادث أمنية من أجل تتبع استخدام Ubuntu أو Windows باستخدام مختلف وكلاء SIEM مثبتين على النظامين.

يتم جمع سجلات نظام تشغيل Windows وUbuntu بواسطة وكيل، ثم يتم إرسال البيانات الأولية إلى خادم معالجة للتحليل. بعد المعالجة، يتم فهرسة السجلات وحفظها في قاعدة بيانات No-SQL بحيث يمكن لأدوات القراءة عرضها كرسومات بيانية وجداول مدمجة في لوحات المعلومات. بعد مراجعة العديد من السجلات ذات الصلة بالأمان، يتم إنشاء تنبيه عند اكتشاف محاولة هجوم.

الكلمات الرئيسية: SIEM، ملفات السجل، نظام المعلومات، الأحداث، لوحة التحكم، التنبيهات، حوادث أمنية، No-SQL، الهجوم، تحليل الأحداث، الفهرسة.

Liste des abréviations

- **MITM :** Man In The Middle
- **DOS :** Denial Of Service
- **DDOS :** Distributed Denial Of Service
- **IOT :** Internet Of Things
- **SOC :** Security Operations Center
- **SIEM :** Security Information and Event Management
- **MSSP :** Managed Security Service Provider
- **NOC :** Network Operations Center
- **CIRT :** Computer Incident Response Team
- **OT :** Operational Technology
- **CSV :** Comma Separated Values
- **JSON :** JavaScript Object Notation
- **XML :** Extensible Markup Language
- **ELF :** Extended Log Format
- **W3C :** World Wide Web Consortium
- **CLF :** COMMON Log Format
- **NCSA :** National Center for Supercomputing Applications
- **CNSS :** Committee on National Security Systems
- **SIM :** Security Information Management
- **SEM :** Security Event Management
- **IPS :** Intrusion Prevention System
- **IDS :** Intrusion Detection System
- **AlienVault OSSIM :** AlienVault Open Source Security Information and Event Management
- **UEBA :** User and Entity Behavior Analytics
- **ML :** Machine Learning
- **AI :** Artificial Intelligence
- **UTM :** Unified Threat Management

- **IP :** Internet Protocol
- **SPA :** Société Par Actions
- **GPL :** Gaz de Pétrole Liquéfié
- **DCSI :** Direction Centrale des Systèmes d'Information
- **IBM :** International Business Machines
- **VM :** Virtual Machine
- **ESX(I) :** Elastic Sky X (Integrated)
- **MozDef :** Mozilla Enterprise Defense Platform
- **RabbitMQ :** Rabbit Messaging Queue
- **MongoDB :** Humongous (énorme) Data Base
- **Http(s) :** Hypertext Transfer Protocol (Secure)
- **AMQP :** Advanced Message Queuing Protocol
- **Amazon SQS :** Amazon Simple Queue Service
- **ELK :** Elasticsearch, Logstash, Kibana
- **Inc :** incorporated
- **RESTful API :** Representational State Transfer Application Programming Interface
- **Apache Ni-Fi :** Apache Niagara Files
- **CentOS :** Community ENTerprise Operating System
- **RHEL :** Red Hat Enterprise Linux
- **Saas :** Software as a service
- **SLES :** SUSE Linux Enterprise Server
- **Amazon EC2 :** Amazon Elastic Compute Cloud 2
- **Mac :** Macintosh
- **AWS :** Amazon Web Services
- **HIDS :** Host-based Intrusion Detection System
- **SPL :** Search Processing Language
- **TF-IDF :** Term Frequency–Inverse Document Frequency
- **Nmap :** Network Mapper
- **UTF-8 :** 8-Bit Universal Character Set Transformation Format
- **OpenJDK :** Open Java Development Kit

- **Sysmon :** System Monitor
- **PGP :** Pretty Good Privacy
- **URL :** Uniform Resource Locator
- **MySQL :** My Structured Query Language
- **Redis :** Remote Dictionary Server
- **TLS :** Transport Layer Security
- **CA :** Certification Authority
- **CSR :** Certificate Signing Request
- **OSIF :** Open Information Security Foundation
- **PPA :** Personal Package Archive
- **SYN :** Synchronize
- **TCP :** Transmission Control Protocol
- **KQL :** Kibana Query Language
- **CISA :** Cybersecurity and Infrastructure Security Agency
- **Syslog :** System Logging Protocol
- **RBAC :** Role-Based Access Control
- **LIBPCAP :** Promiscuous Capture Library
- **SSL :** Secure Sockets Layer

Table Des Matières

Liste Des Figures

Liste Des Tableaux

INTRODUCTION GÉNÉRALE

Nous vivons à une époque où les technologies de l'information évoluent constamment. Si une entreprise ne s'alignait pas sur cette évolution, cela représenterait pour elle un sérieux handicap, face à une concurrence de plus en plus rude sur le marché d'une part, et aux risques de sécurité de l'information de plus en plus fréquents d'autre part, d'autant plus que le système d'information représente un patrimoine essentiel de l'entreprise, la sécurité de ce dernier est primordiale.

Pour faire face aux menaces de sécurité, la mise en place d'outils et de technologies permettant d'opérer avec le système et d'assembler les différentes informations en temps réel est essentielle. Pour ce faire, l'adoption du concept de gestion des informations et des événements de sécurité est le bon choix.

C'est dans ce sens que la grande entreprise nationale leader sur le marché algérien des produits pétroliers **NAFTAL Spa** nous a ouvert ses portes, pour la réalisation de notre stage de fin d'études.

Problématique

La difficulté de gérer d'énormes fichiers journaux sur plusieurs hôtes individuellement et le défi de comprendre les modèles d'attaque qui conduisent à un incident de sécurité sont deux des principales causes de l'échec de l'identification des attaques. En effet, le fait de n'avoir aucun moyen de détecter, de suivre et d'analyser le flux de données à travers le système peut gravement paralyser la capacité de réagir et de remédier aux incidents à temps.

Les systèmes actuels utilisés pour l'extraction et l'analyse des événements de sécurité sont confrontés à des défis majeurs. Si l'on regarde du côté open source, la plupart des SIEM (Security Information and Event Management) open source nécessitent plus de compétences et un déploiement approprié que les solutions SIEM commerciales.

En outre, les solutions SIEM commerciales nécessitent des coûts de déploiement et de maintenance élevés.

Objectifs

Le but de ce travail est de déployer et configurer un SIEM (Security Information and Event Management) sous la forme d'une Appliance Virtuelle.

Le SIEM doit pouvoir effectuer les opérations suivantes :

- Regrouper divers appareils, services, serveurs et systèmes.
- La vérification des logs, des alertes antivirus et des événements des différents appareils, services, etc.
- Consolidation de plusieurs points de données.
- Stocker et indexer les événements collectés
- Effectuer des recherches rapides et filtrées sur les événements stockés.
- Corréler les événements de plusieurs ressources.
- Surveiller l'accès aux ressources critiques.
- Réagir à différent types d'attaques.

Les résultats seront présentés sous forme d'informations exploitables via des tableaux de bord.

Plan du mémoire

Pour bien organiser notre travail, nous avons divisé ce mémoire en cinq (05) chapitres :

- **Chapitre 1 :** Nous présentons des informations générales concernant la sécurité des systèmes d'information sur lesquels nous nous sommes basées tout au long de notre projet.
- **Chapitre 2 :** Nous commencerons par présenter l'organisme d'accueil, ses missions et ses objectifs. Ensuite, nous ferons une étude comparative des

différentes solutions disponibles et établirons une discussion pour choisir une solution. Nous présenterons ensuite l'architecture de la solution proposée et les techniques utilisées dans ce projet.

- **Chapitre 3 :** Ce chapitre couvrira les détails de conception de notre système.
- **Chapitre 4 :** Nous abordons dans ce chapitre l'implémentation de notre solution SIEM.
- **Chapitre 5 :** Dans ce chapitre, nous allons présenter les différents scénarios de test afin de valider l'efficacité et le bon fonctionnement de notre solution.
- **Conclusion générale** : Le mémoire se termine par une conclusion générale qui présente un bilan des travaux menés dans ce mémoire.

Chapitre 1

GÉNÉRALITÉS SUR LA SÉCURITÉ D'INFORMATION

1 Introduction

Avant de commencer tout projet, il est essentiel de documenter les concepts autour du sujet. Cette première étape nous permettra de nous familiariser avec les notions préalables à maîtriser et l'évolution des différents outils et méthodes proposés dans le domaine. Nous présentons ainsi dans ce premier chapitre les notions relatives à la Sécurité de l'Information, notamment la cybersécurité, les cyberattaques, le Security Operations Center, les Logs et le SIEM (Security Information and Event Management) sur lesquels nous nous sommes basées tout au long de notre projet.

2 Cybersécurité

La cybersécurité ou la sécurité informatique est la protection des réseaux et systèmes informatiques contre les fuites de données, le vol ou la détérioration du matériel, des logiciels ou des données électroniques, et aussi contre les coupures ou les détournements des services qu'ils fournissent [1]**.** Elle peut être répartie en différentes catégories [1] :

- **La sécurité réseaux** : Consiste à protéger les réseaux informatiques des intrus.
- **La sécurité des applications** : Elle vise à garantir la protection des programmes et équipements informatiques contre les menaces.
- **La sécurité des informations** : Son objectif est d'assurer l'exactitude, l'exhaustivité, la cohérence globales des données et la confidentialité des données stockées ou transférées.
- **La sécurité opérationnelle** : Son but principale est le traitement et la protection des données (qui a accès au réseau, où leurs données sont stockées et comment leurs données sont traitées).
- **La récupération après accident et la continuité des opérations** : Elle a pour objet de préciser la manière dont l'entreprise répondra aux cyberattaques ou autres incidents entraînant la perte d'opérations ou de données.

- **L'apprentissage de l'utilisateur final** : Elle est conçue pour informer les utilisateurs sur les dernières menaces en matière de cybersécurité pour prévenir les failles de sécurité (comment accéder en toute sécurité aux ressources de l'entreprise et quelles politiques informatiques sont en place).

2.1 Objectifs de la cybersécurité :

La sécurité informatique a plusieurs objectifs liés à la fois aux types de menaces et aux types de ressources. Les points les plus importants sont [2] :

- Empêcher les violations de données non autorisées
- Empêcher les modifications non autorisées des données
- Protection générale contre l'utilisation non autorisée du réseau ou des ressources informatiques

2.2 Principes de la cybersécurité :

Afin de caractériser ce que les utilisateurs des systèmes informatiques attendent au regard de la sécurité, la sécurité informatique s'appuie sur un certain nombre de piliers capables de guider les efforts de sécurité de l'information pour en assurer le succès [3].

2.2.1 Confidentialité

Le principe de confidentialité assure que seuls les utilisateurs autorisés ont accès à l'information. Si une personne non autorisée a accès à l'information, la confidentialité est compromise.

2.2.2 Authentification

Le principe d'authentification exige une preuve d'identité, il garantit que l'origine de l'information ainsi que les différents acteurs sont correctement identifiés. Peut-être explicite (e.g. mot de passe, empreintes, etc.) ou implicite (e.g. comportement, etc.)

2.2.3 Intégrité

Le principe d'intégrité garantit que l'information n'est altérée (modifiée) que par les acteurs autorisés. La signature électronique, par exemple, permet de vérifier l'existence d'une altération.

2.2.4 Disponibilité

Le principe de disponibilité garantit que les ressources soient disponibles à tout moment pour les personnes autorisées.

2.2.5 Non-répudiation

Le principe de non-répudiation ne permet pas au propriétaire d'une information (e.g. message) de réfuter la possession de cette information. Il y a des situations, par exemple, où un utilisateur envoie un message et nie par la suite qu'il a envoyé ce message

2.2.6 Traçabilité

S'assurer que l'accès et les tentatives d'accès aux éléments pertinents sont suivis, et que ces traces sont maintenues et utilisées.

3 Cyberattaques

Le Comité sur les systèmes de sécurité nationale (CNSS) est une organisation intergouvernementale des États-Unis qui établit des politiques, des instructions et des procédures opérationnelles de cybersécurité au niveau national pour le gouvernement des États-Unis pour la sécurité des systèmes de sécurité nationale (NSS) [4] **.**

Selon CNSS une cyberattaque est considérée comme « *des actions prises dans le cyberespace qui créent des effets de déni perceptibles (c'est-à-dire dégradation, perturbation ou destruction) dans le cyberespace ou manipulation qui conduit à un déni qui apparaît dans un domaine physique, et qui est considéré comme une forme de feu* ». [5]**.**

3.1 Différents types de cyberattaques les plus courants

La cybercriminalité augmente considérablement chaque année, à mesure que les cyberattaquants s'améliorent en termes d'efficacité et de sophistication. Les cyberattaques se produisent pour un certain nombre de raisons différentes et de différentes manières. Cependant, un fil conducteur est que les cybercriminels chercheront à exploiter les vulnérabilités des politiques, pratiques ou technologies de sécurité d'une organisation. [6]

Il existe plusieurs types de cyberattaques. La liste ci-dessous comprend les cyberattaques les plus communes que les cyberattaquants exploitent pour compromettre les systèmes des entreprises. [7]

- Logiciels malveillants (Malware)
- Phishing
- Man-in-the-Middle (MitM)
- Déni de service (DOS) / déni de service distribué (DDOS)
- SQL Injections
- Zero-day Exploitation
- Attaque par mot de passe (force brute)
- Cross-site Scripting
- Rootkits
- Attaques liées à l'Internet des objets (IoT)

3.2 Conséquences d'une cyberattaque

Les cyberattaques peuvent entraîner des cybercrises financières, informatiques et de réputation. Une cyberattaque peut avoir les conséquences suivantes :

- Compromettre l'intégrité
- Vol de données
- Redirection vers des sites malveillants
- Destruction de données

- Pannes matérielles
- Paralysie des systèmes

4 Security Operations Center (SOC)

« Un centre d'opérations de sécurité (SOC) est un lieu centralisé de surveillance et de gestion fréquente de la sûreté et de la sécurité du statut de l'entreprise. L'objectif principal du SOC est de permettre de meilleures capacités de détection, d'investigation et de réponse aux incidents en utilisant les données des terminaux, des journaux, des systèmes de sécurité et des flux réseau. De plus, un SOC efficace peut aider les organisations à améliorer leur capacité de connaissance de la situation et à accroître le déploiement des ressources de l'entreprise pour atténuer les problèmes de sécurité. » [8]. Ses responsabilités sont les suivantes :

- Service de détection d'incidents
- Service de réponse aux incidents
- Services de reprise en cas d'incident
- Analyse du système informatique suite à un incident
- Exécution du plan de récupération **SOC** (centre des opérations de sécurité) et **SIEM** (Security Information and Event Management) travaillent ensemble pour aider les entreprises à prévenir les violations de données et les alerter des cyber-événements potentiels en cours.

4.1 Modèles de SOC :

Il existe de nombreux modèles de SOC, notamment : [9]

4.1.1 Virtual SOC

Il s'agit d'un portail Web basé sur des technologies de sécurité décentralisées qui ne nécessite aucune installation dédiée. Il s'appuie sur une équipe à temps partiel pour s'activer lorsqu'une alarme ou un événement critique survient.

4.1.2 Dedicated SOC

C'est un SOC centralisé avec une installation, une équipe et des processus dédiés entièrement axés sur la sécurité, il offre à l'équipe la meilleure visibilité pour surveiller l'environnement et obtenir une image complète des menaces et de la sécurité.

4.1.3 Distributed/Co-managed SOC

Ce modèle a des membres d'équipe dédiés/semi-dédiés qui sont au moins cinq à huit experts en sécurité internes embauchés pour travailler aux côtés d'un fournisseur de services de sécurité gérés (MSSP) tiers.

4.1.4 Command SOC

Coordonner les différents SOC (centres des opérations de sécurité) en offrant des informations supplémentaires sur les menaces, une connaissance de la situation et d'autres expertises.

4.1.5 Network Operations Center (NOC)

Comme le SOC dédié, ce modèle de SOC dispose d'une installation, d'une équipe et de processus dédiés, assurant non seulement la sécurité, mais également une surveillance continue des performances et de la santé du réseau.

4.1.6 Fusion SOC

Les Fusion SOC sont considérés comme des SOC avancés, ils combinent des fonctions SOC de base et plus récentes, telles que les renseignements sur les menaces, l'équipe de réponse aux incidents informatiques (CIRT) et les fonctions de technologie opérationnelle (OT), intégrées dans une seule installation SOC.

4.2 Comment fonctionne un SOC :

Un centre des opérations de sécurité contient des systèmes de détection d'intrusion, des systèmes de prévention des intrusions et d'autres solutions de détection d'intrusion, des pares-feux et un SIEM (Security Information and Event Management)[10].

Des technologies doivent être déployées pour collecter des données dans des flux, des métriques, des captures de paquets, des journaux système et d'autres méthodes afin que les équipes SOC puissent corréler et analyser l'activité des données.

Le SOC inspecte aussi les terminaux et le réseau pour trouver des failles de sécurité dans le but de garantir la protection des données importantes et pour se conformer aux réglementations sectorielles ou gouvernementales.

5 Log

Avant de parler du SIEM (Security Information and Event Management) et de ses fonctionnalités, nous devons parler des journaux et de leur importance. Plus précisément, il convient aux domaines de l'ingénierie et de la sécurité des systèmes d'information. Alors on a besoin de savoir dès le départ ce qu'est un fichier journal et quel est son objectif ?

5.1 Définition De Log

S. Al-Fedaghi et F. Mahdi définissent les logs comme celui-là : « *Une séquence ordonnée d'occurrences contenant la preuve de l'exécution d'un processus par des utilisateurs, des systèmes ou d'autres entités. Diverses sources et entités du système envoient des messages concernant leurs processus (par exemple, qui, quelles opérations, heure, etc.), ces derniers sont conservés dans plusieurs journaux ou fichiers logs* » [11].

K. Kent et M. Souppaya définissent un log de cette manière : « *Un log est un enregistrement des événements survenant dans les systèmes et les réseaux d'une organisation.*

Les logs sont composés d'entrées de journal ; chaque entrée contient des informations relatives à un événement spécifique qui s'est produit au sein d'un système ou d'un réseau. » [12]**.**

5.2 Formats De Journaux

Le format du journal est la structure et l'emplacement des différents champs dans chaque ligne du fichier journal. Il peut s'agir de documents CSV, JSON, XML ou simplement de texte brut. [13]

Pour que les journaux soient utiles et faciles à utiliser, ils doivent être conformes aux formats acceptés par les différentes solutions SIEM. Ce dernier dispose d'analyseurs spécifiques pour chaque type de journal. Si la solution SIEM utilisée n'accepte pas le format de journal qui lui a été envoyé, un analyseur spécial introduit par l'utilisateur peut être utilisé.

Les **formats de journaux communs** (CLF) du NCSA (National Center for Supercomputing Applications) et **les formats de journaux étendus** (ELF) du W3C (World Wide Web Consortium) sont l'une des structures ou formats de journaux les plus reconnus.

Le format d'un **journal commun** est le suivant :

"%h %l %u %t \" %r\" %>s %b"

Les champs du log sont les suivants :

- **%h :** L'adresse IP ou le nom DNS du client distant.
- **%l :** Le nom du journal, dans le cas d'un champ vide, on le remplace par "–".
- **%u :** Le nom de l'utilisateur distant, dans le cas d'un champ vide, on le remplace par "–".
- **%t :** Horodatage ou Le temps de l'événement envoyé.
- **\" %r\" :** La ressource de la requête envoyée.
- **%>s :** Le code d'état HTTP renvoyé par le serveur.
- **%b :** La taille de l'objet renvoyé en octets.

Un exemple d'un **journal commun** est le suivant :

```
192.168.1.7 user-identifier DELL [12/Jul/2022:13:55:36] "GET /example.gif HTTP/1.0" 200 3519
```

Figure 1 : Exemple d'un commun log.

Dans l'exemple ci-dessus, **192.168.1.7** représente l'adresse IP du client distant, **user-identifer** est le nom du journal, **DELL** représente le nom d'utilisateur distant, l'horodatage est **12/Jul/2022:13:55:36**, la ressource demandée avec la méthode **GET** est **/example.gif**, **200** représente le code d'état HTTP envoyé par le serveur et la taille de l'objet renvoyé est **97** octets.

Le **format de journal étendu** est le même format que **le journal commun**, mais les journaux ELF (Extended Log Format) fournissent plus d'informations, comme indiqué dans la figure suivante :

```
"%h %l %u %t \"%r\" %>s %b \"%{Referer}i\" \"%{User-agent}i\""
```

Figure 2 : Le format d'un EXTENDED Log.

Les champs ajoutés sont :

- **\" % {Referer} i \"** : L'adresse Web de la requête extraite de l'en-tête.
- **\" %{User-agent}i\"** : Le navigateur Internet utilisé.

Un exemple d'un **EXTENDED Log** est le suivant :

```
192.168.1.7 user-identifier DELL [12/Jul/2022:13:55:36] "GET /example.gif HTTP/1.0" 200 3519 "https://www.google.com" "Chrome/103.0.5060.134"
```

Figure 3 : Un exemple d'un EXTENDED Log

Dans l'exemple ci-dessus, **https://www.google.com/** représente l'URL de la requête extraite de l'en-tête et le navigateur Internet utilisé est **"Chrome/103.0.5060.134".**

5.3 Catégories De Journaux

Chaque composant d'un réseau ou d'un système génère aujourd'hui différents types de données, et chaque composant collecte ces données dans ses propres journaux. Pour cette raison, il existe de nombreuses catégories de journaux, telles que [13] :

- **Les événements système ou Syslog (System Logging Protocol) :** contiennent les journaux liés au système d'exploitation comme les notifications de démarrage, les modifications du système, les arrêts soudains, erreurs et avertissements, ainsi que d'autres processus de haut niveau.
- **Les événements de sécurité :** incluent les tentatives réussies et infructueuses d'accès aux systèmes, aux applications ou aux données importantes, journaux d'audit, etc.
- **Les événements des applications :** contiennent des informations sur les événements qui se sont produits dans les applications installées sur le système.
- **Journaux des menaces réseau :** contiennent des événements de trafic système, fichier ou application qui correspondent à une règle de sécurité dans un pare-feu.

5.4 Importance Des Fichiers Journaux

La gestion des journaux peut bénéficier aux entreprises de plusieurs façons, par exemple en examinant et en analysant périodiquement les fichiers journaux qui ont été conservés en détail pendant une période de temps définie par l'entreprise, cela aide à identifier les incidents de sécurité, les violations de politique, les activités frauduleuses

et les problèmes opérationnels et fournit des informations utiles pour résoudre ces problèmes.[12]

Les journaux peuvent également faciliter les enquêtes organisationnelles internes, établir des lignes de base et identifier les tendances de performances et les problèmes à long terme. [12]

6 SIEM

E. Kostrecová et H. Bínová définissent un SIEM (Security Information and Event Management) ainsi « *En général, le gestionnaire des informations et des événements de sécurité est une combinaison des catégories de produits autrefois disparates que sont SIM (gestionnaire des informations de sécurité) et SEM (gestionnaire des événements de sécurité). La technologie SIEM fournit une analyse en temps réel des alertes de sécurité générées par le réseau, le matériel et les applications. Les solutions SIEM se présentent sous la forme de logiciels, d'appliances ou de services gérés, et sont également utilisées pour enregistrer les données de sécurité et générer des rapports non seulement à des fins de conformité, mais également pour détecter une menace potentielle* » [14]**.**

M. Di Mauro and C. Di Sarno définissent un SIEM à leur tour comme ceci : « *un SIEM a deux fonctions principales : la collecte et l'agrégation des données de journalisation et des informations de sécurité provenant de divers périphériques réseau (routeurs, pare-feu, systèmes de détection d'intrusion, et autres) et l'analyse des données recueillies en implémentant un ensemble de règles de corrélation visant à détecter d'éventuels événements suspects comme la présence de trafic chiffré en temps réel*» [15] **.**

Sur la base de ces définitions, on peut définir un SIEM comme une Technique qui réunit à la fois la gestion des événements de sécurité et la gestion des informations de sécurité. Par conséquent, SIEM doit pouvoir analyser, gérer et stocker les logs (fichiers journaux) et les événements afin d'extraire des informations de sécurité pour détecter les menaces et les incidents de sécurité.

6.1 La différence entre SIM, SEM et SIEM :

Le SIEM assemble les fonctionnalités du Gestionnaire d'informations de sécurité et du Gestionnaire d'événements de sécurité. Les missions de SEM (Security Event Management) et SIM (Security Information Management) peuvent sembler similaires, mais chacun a des caractéristiques distinctes : [16]

6.1.1 SIM

- SIM collecte des fichiers journaux qui peuvent être constitués de différents types de données, les stocke dans un centre de données central pour les analyser ultérieurement.
- SIM est également appelé gestionnaire des journaux.
- Les solutions SIM sont souvent basées sur des agents avec des logiciels exécutés sur les serveurs et les ordinateurs surveillés. Ceux-ci relaient les fichiers journaux et d'autres informations liées à la sécurité vers un serveur SIM central.

6.1.2 SEM

- Un SEM identifie, regroupe, inspecte et met en corrélation les journaux et les signaux du système.
- Tout comme SIM, les données sont généralement relayées de l'ordinateur hôte vers un référentiel central qui garantit que les événements et les alertes sont conservés dans un stockage fiable et sécurisé.
- SEM analyse les événements avec des algorithmes de sécurité et des calculs statistiques pour identifier les menaces, les vulnérabilités et les risques.
- Le SEM peut analyser les entrées pour leur importance au fur et à mesure qu'elles arrivent et avertir immédiatement les personnes responsables chaque fois qu'une entrée mérite leur attention.
- L'objectif principal de l'outil SEM est d'identifier les alertes ou les événements méritant une enquête, tels que l'authentification suspecte, la connexion au compte ou l'accès au personnel administratif de haut niveau à certaines heures du jour ou

de la nuit. Les experts citent souvent les "événements de super-utilisateur" comme ce que recherchent les gestionnaires d'événements de sécurité.

6.1.3 SIEM

- Le SIEM assemble les fonctionnalités du SIM et du SEM.
- Un SIEM collecte, organise et analyse les activités liées à la sécurité à partir de nombreuses sources matérielles et logicielles dans l'infrastructure technologique d'une organisation.
- Un SIEM regroupe les données en temps réel et historiques des routeurs, commutateurs, serveurs, ordinateurs, logiciels antivirus, pare-feu, systèmes de prévention/détection d'intrusion (IPS/IDS), applications d'entreprise, bases de données, etc.
- Il applique des règles d'analyse prédéfinies aux données afin de détecter les menaces et les activités suspectes qui nécessitent une action ou une enquête de la part d'un administrateur système.

Le tableau suivant résume les différences entre SEM, SIM et SIEM.

	SIM	SEM	SIEM
Avis Général	Collecte et analyse de données liées à la sécurité à partir de journaux informatiques	Analyse des menaces en temps réel, visualisation et réponse aux incidents.	SIEM combine les capacités de SIM et SEM.
Caractéristiques	Facile à déployer, puissantes capacités de gestion des journaux	Plus complexe à déployer, supérieur dans la surveillance en temps réel.	Plus complexe à déployer, fonctionnalité complète.
Exemple	AlienVault OSSIM	NetIQ Sentinel	Elastic Stack

Tableau 1 : Comparaison entre SEM, SIM et SIEM [17].

6.2 Les Fonctions D'un SIEM

Les solutions SIEM, malgré leur diversité entre solutions open source et solutions commerciales, partagent les mêmes tâches de base. Les solutions SIEM offrent une vue globale de toutes les activités se déroulant dans une infrastructure informatique en surveillant l'activité du réseau et en utilisant les renseignements sur les menaces et l'analyse du comportement des utilisateurs pour détecter et atténuer les attaques.

6.2.1 Logs Management

La gestion des journaux implique la collecte, la normalisation et l'analyse des données des journaux pour comprendre l'activité du réseau, détecter les attaques et les événements de sécurité et répondre aux exigences réglementaires informatiques pour une analyse efficace des journaux, les solutions SIEM utilisent différents processus tels que la corrélation des journaux et l'investigation, qui aident à détecter les violations de

données et les attaques en temps réel. La gestion des journaux comprend également un archivage sécurisé des données des journaux pour conserver les journaux pendant des périodes personnalisables.

6.2.2 La gestion des incidents

Les incidents de sécurité sont des événements anormaux provoqués par une activité régulière sur le réseau. Les incidents peuvent mettre en danger les données sensibles d'une organisation et entraîner une violation ou une attaque de données.[18]

La résolution d'incident fait référence à la résolution d'une attaque dans un réseau et à la restauration du réseau à un état fonctionnel [19]**.** Les solutions SIEM fournissent divers flux de travail qui peuvent être automatisés lorsque des alertes sont déclenchées. Ces flux de travail contribuent grandement à empêcher les attaques de se propager latéralement au sein du réseau.

6.2.3 Analyse du comportement des utilisateurs et des entités (UEBA)

L'UEBA (User and Entity Behavior Analytics) dans une solution SIEM est souvent basée sur le **ML** (Machine Learning) ou l'**AI** (Artificial Intelligence) et analyse les schémas de travail normaux des utilisateurs, ou la manière dont un utilisateur particulier accède généralement au réseau au quotidien. Il peut détecter les écarts par rapport au comportement normal, déclencher des alertes et avertir immédiatement les administrateurs de sécurité.

6.2.4 Protection des données.

L'un des principaux objectifs des professionnels de la sécurité est d'empêcher la perte ou la fuite de données sensibles. Les solutions SIEM aident à détecter, atténuer et prévenir les violations de données en surveillant en permanence le comportement des utilisateurs. Les solutions SIEM suivent l'accès aux données critiques et identifient les accès non autorisés ou les tentatives d'accès. Il surveille également l'élévation des privilèges dans les comptes d'utilisateurs et toutes les modifications de données

apportées par ces comptes. Lorsque ces capacités de détection sont associées à la gestion des flux de travail, les administrateurs de sécurité peuvent configurer des solutions SIEM pour empêcher les activités malveillantes sur le réseau.

6.2.5 Le renseignement sur les menaces

Le renseignement sur les menaces est un élément clé pour les systèmes UTM (Unified threat management) et pour les solutions SIEM. Ces derniers doivent être réglés pour recueillir des événements sur les menaces dans le but de détecter les nouvelles sources de spam, les malwares et d`autres vulnérabilités. Ces renseignements sont exploités pour des vérifications automatisées pour neutraliser les menaces sur les réseaux d'entreprises.

6.3 Cycle de vie des événements dans SIEM

Lorsqu'un événement se produit, le SIEM signale immédiatement l'événement et collecte toutes les données brutes. Ces données sont ensuite transformées dans un format habituel, où le SIEM peut d'une manière automatique deviner et attribuer un degré de risque à l'évènement. Dans cette section, nous couvrons chaque étape du cycle de vie des événements dans une solution SIEM :

6.3.1 L'analyse (Parsing)

L'analyse est la première étape par laquelle chaque journal doit passer lorsqu'il est capturé par SIEM. L'analyse est le processus de décomposition des données en éléments d'information plus faciles à manipuler et à stocker. En ce qui concerne l'analyse du fichier journal, chaque journal contient plusieurs informations stockées sous forme de texte, et le but de l'analyse est de les identifier et de les regrouper de manière significative (par exemple, regrouper tous les ID utilisateur contenus dans le fichier journal) [20].

On peut diviser l'analyse en deux étapes principales :

- Primaire - allocation et remplissage des structures de données.

- Secondaire - Exécution (de la logique, des appels d'API, etc.) basée sur les données des structures.

6.3.2 L'agrégation

L'agrégation de journaux est le processus de collecte, de normalisation et de combinaison des données de journaux à partir d'une gamme de sources (applications, services, etc.) pour organiser ces données dans un emplacement centralisé et de rendre les données consultables [21] **.**

Une mise en œuvre efficace de l'agrégation des journaux permet à l'équipe de sécurité de gagner un temps précieux lors des pannes et d'identifier de manière proactive les modèles d'activité du système. Avec la bonne mise en œuvre, les équipes peuvent rechercher et filtrer les journaux, résoudre les problèmes en réponse aux différents incidents, effectuer une surveillance en temps réel, etc.

6.3.3 La corrélation

La corrélation d'événements est un processus qui connecte une masse d'événements de réseau, de système, de service, etc. à des modèles identifiables [22].

L'objectif principal de la corrélation est de transformer les données brutes en alertes et rapports grâce aux règles définies par l'équipe de sécurité.

6.3.4 Les Rapports et les alertes

Les outils d'alerte sont le plus souvent utilisés par les centres d'opérations de sécurité pour protéger une organisation, ce qui en fera un outil indispensable dans chaque solution SIEM.

Les alertes SIEM sont déclenchées en cas de détection d'un incident de sécurité en vérifiant les règles de corrélation. En bref, si un ensemble d'événements correspond à une règle de corrélation, une alerte sera déclenchée.

6.3.5 Cas d'utilisation (Use cases)

Après avoir reçu des alertes, l'analyste de sécurité crée un cas d'utilisation et commence à enquêter sur l'alerte. Cela peut prendre un moment ou plusieurs jours pour résoudre cet incident.

La solution SIEM aide l'analyste à trouver l'incident de sécurité, et il peut aller plus loin en fournissant plus d'informations sur cet incident, par exemple en cas d'attaque de malware, le SIEM fournit plusieurs informations dont l'hôte infecté et les fichiers infectés, ce qui aide l'analyste dans l'investigation car il peut savoir à quel point ce logiciel malveillant s'est propagé dans le réseau de l'organisation, et commencer à isoler les machines affectées du réseau de l'organisation pour limiter la propagation de ce malware.

La figure ci-dessous résume le cycle de vie des journaux dans SIEM :

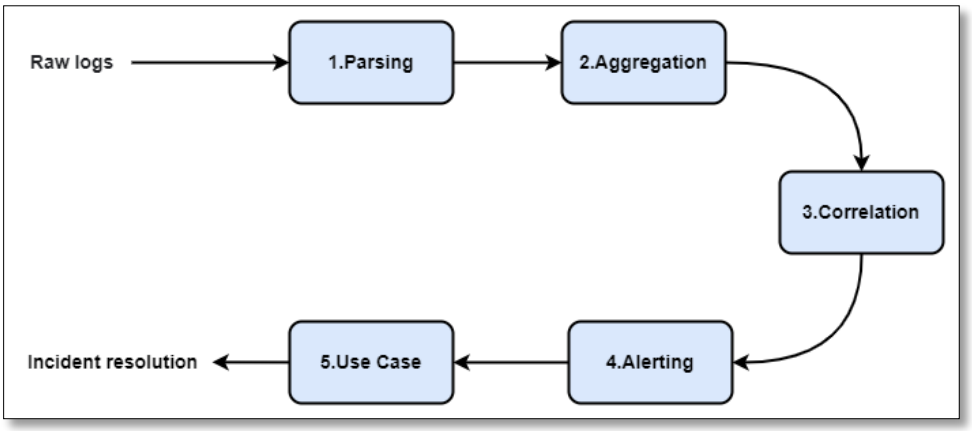


Figure 4 : Cycle de vie d'un log dans le SIEM.

7 Conclusion

Ce chapitre nous a permis de développer nos connaissances en sécurité des systèmes d'information et de comprendre les concepts de base et l'évolution des différentes technologies et méthodes proposées dans ce domaine et leur importance dans la sécurisation des systèmes d'information.

En prenant un aperçu de la cybersécurité et des menaces auxquelles elle est confrontée, et en parlant des centres d'opérations de sécurité (SOC) et du fait que SOC et SIEM (Security Information and Event Management) travaillent ensemble pour prévenir les violations de données, puis en passant à la compréhension de la technologie SIEM et de son fonctionnement et du cycle de vie des journaux dans SIEM nous a permis d'établir le plan d'étude comparatif qui sera traité dans le chapitre suivant pour choisir les solutions les plus adaptées à notre projet.

Chapitre 2

ETAT DE L'ART

1 Introduction

L'étude de l'existant est une étape du projet d'une importance majeure, puisqu'une étude incomplète ou mal menée implique la réalisation d'un projet non adapté aux besoins de l'entreprise et conduit donc inévitablement à un échec du projet. Cela aidera à mieux comprendre le problème afin de proposer la solution idéale pour l'entreprise.

Dans ce chapitre, nous présenterons à la fois l'organisme d'accueil et la structure d'accueil, leurs missions et objectifs. Ensuite, nous passerons en revue les différentes solutions disponibles et ferons une étude comparative de ces solutions et de nos choix et raisons de les choisir. Nous présenterons ensuite l'architecture de la solution proposée et les techniques utilisées dans ce projet.

2 Présentation de l'organisme d'accueil

2.1 Société Naftal

Fondée en 1982 et filiale à 100% du Groupe Sonatrach, Naftal est une société par actions (SPA) au capital social de 15 650 000 000 DA. Elle a pour mission principale, la distribution et la commercialisation des produits pétroliers et dérivés sur le marché national. Elle distribue et commercialise les carburants terre, aviation et marine, les GPL (Gaz de Pétrole Liquéfié), les lubrifiants, les bitumes, les pneumatiques et les produits spéciaux. L'entreprise siège à rue des dunes BP 73, Chéraga, Alger.

La figure suivante présente l'organigramme global de NAFTAL :

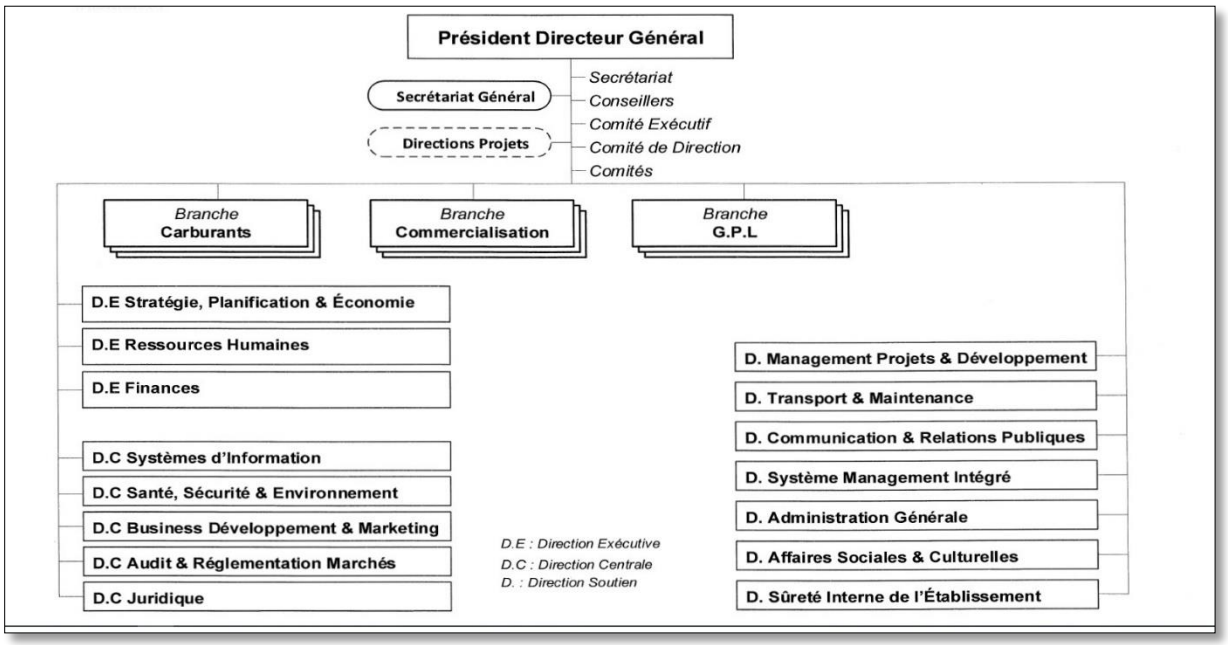


Figure 5 : Schéma de la Macrostructure de NAFTAL S.P.A

2.2 Missions et objectifs

A travers ses programmes de développement socio-économique, NAFTAL vise un double objectif :

- Poursuivre sa mission fondamentale qui est la distribution des produits pétroliers.
- Perfectionner sa qualité de service.

NAFTAL opère avec trois (03) branches d'activités opérationnelles :

2.2.1 La Branche CARBURANTS

A pour missions :

- L'approvisionnement et le ravitaillement en carburants des centres de dépôts carburants terre, aviation et marine à partir des raffineries et la commercialisation des produits aviation et marine.

- Le respect des exigences d'Hygiène, de Sécurité, de l'Environnement et de la Qualité.

2.2.2 La Branche GPL

A pour missions :

- La Satisfaction des besoins de la clientèle en **GPL vrac** et conditionné en tous lieux et en toute circonstances.
- Le rapprochement et le maintien de l'écoute des attentes et des exigences de la clientèle.
- Le respect des exigences d'Hygiène, de Sécurité, de l'Environnement et de la Qualité.

2.2.3 La Branche COMMERCIALISATION

A pour missions :

- La mise à la disposition de la clientèle, l'ensemble de produits pétroliers à travers son réseau station-service et par vente directe aux gros consommateurs sur tout le territoire national.
- La généralisation de la carte puce comme moyen de paiement électronique moderne sur tout le territoire national.
- Le respect des exigences d'Hygiène, de Sécurité, de l'Environnement et de la Qualité.

2.3 Présentation de la structure d'accueil DCSI

DCSI (Direction Centrale des Systèmes d'Information) est la nouvelle structure de NAFTAL, créée dans le cadre du processus de restructuration de la société en 2004. Le projet présenté dans ce mémoire a été proposé par la Direction Centrale Des Systèmes d'Information (DCSI) dans le cadre d'un plan d'action visant à déceler les failles de sécurité du système informatique et des systèmes de communications de l'entreprise.

La figure suivante présente l'Organigramme de la DCSI :

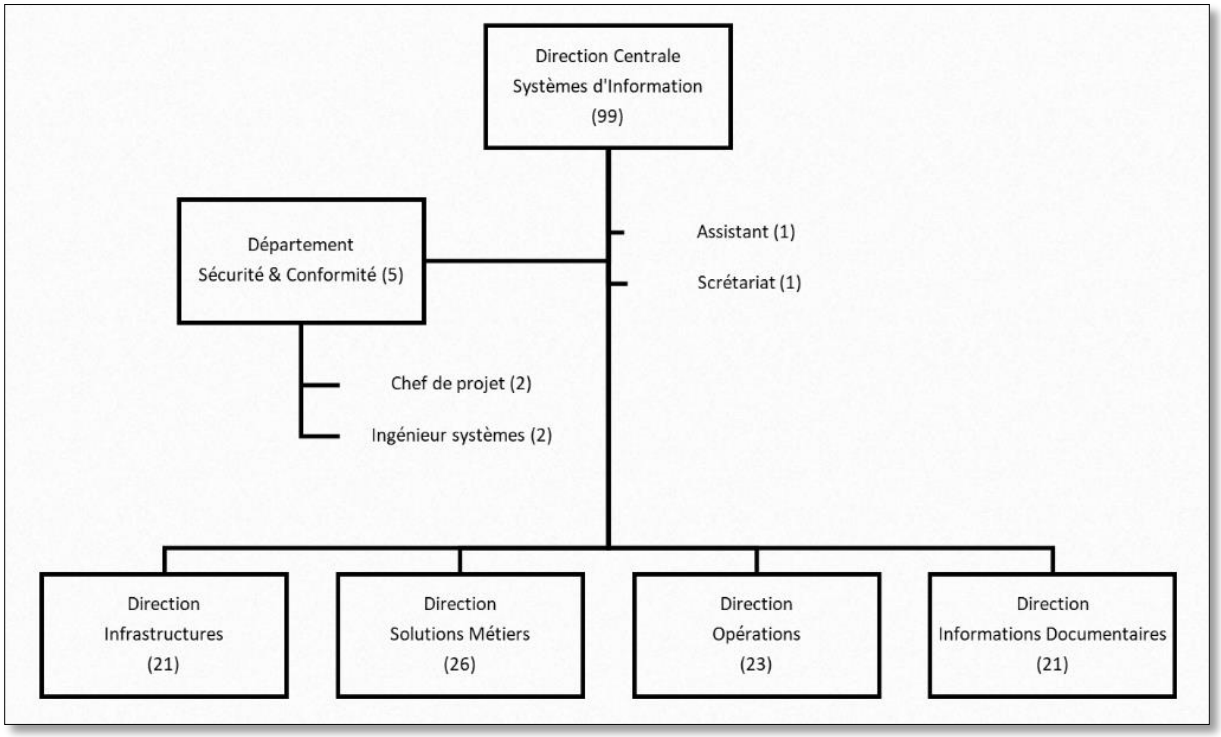


Figure 6 : Organigramme Direction Centrale Systèmes d'Information

2.4 Missions de la DCSI

Les missions des cinq (05) directions de la DCSI sont :

2.4.1 Direction INFRASTRUCTURES

- Définir et mettre en œuvre l'architecture des systèmes, des bases de données et réseaux d'infrastructure du Système d'Information.
- Doter l'entreprise d'une infrastructure de communication sous forme de réseau étendu qui intègre toutes les structures de la Société

2.4.2 Direction SOLUTIONS MÉTIERS

- Concevoir et réaliser des solutions informatiques qui répondent aux besoins opérationnels de l'ensemble des Structures de la Société.

2.4.3 Direction OPÉRATIONS

- Veiller au bon fonctionnement des plateformes monétiques et décisionnelles.
- Garantir la disponibilité du matériel informatique dédié aux utilisateurs finaux du système d'information de la société.

2.4.4 Département SÉCURITÉ & CONFORMITÉ

- Concevoir et mettre en place un dispositif permettant la sécurité et la pérennité des systèmes d'information mis en place.
- Mettre en place des règles de conformité et de contrôle interne dans l'établissement.

2.4.5 Direction INFORMATIONS DOCUMENTAIRES

- Construire une banque de données documentaire de l'information règlementaire interne et externe.
- Mettre en couvre les solutions adéquates pour la conservation du patrimoine documentaire et informationnel de la société.

3 Solutions disponibles

3.1 Les Solutions SIEM Sous Licence

De nombreux produits SIEM (Security Information and Event Management) commerciaux sont également disponibles, Gartner fournit des rapports annuels sur une variété de technologies et de principes, y compris SIEM Et afin de choisir les produits à présenter dans cette section, on va servir de Gartner

« *Gartner, est une société de recherche et de conseil qui fournit aux chefs d'entreprise des informations, des conseils et des outils qui aident à créer des organisations. Elle opère à travers les segments suivants : Recherche, Conseil et Conférences. Le segment Recherche fournit aux dirigeants des informations et des conseils objectifs par le biais de rapports, de*

briefings, d'outils propriétaires, d'un accès aux experts en recherche de l'entreprise, de services de réseautage entre pairs et de programmes d'adhésion. » [23].

La figure suivante montre le Gartner Magic Quadrant de Security Information and Event Management publier en avril 2021

Figure 7 : Gartner Magic Quadrant de SIEM avril 2021 [24] .

3.1.1 IBM Security QRadar

Au cours des dernières années, la réponse d'IBM (International Business Machines) au SIEM a évolué pour devenir l'une des meilleures du marché. La plate-forme fournit une variété de fonctionnalités de gestion des journaux, d'analyse, d'acquisition de données et de détection d'intrusion pour aider à maintenir les systèmes critiques en fonctionnement. Toute la gestion des journaux dans un seul outil : **QRadar Log Manager**. En matière d'analyse, QRadar est une solution presque complète.

a. Principales caractéristiques

- Gestion des journaux
- Détection d'intrusion
- Fonctions analytiques

Le système dispose d'une analyse de modélisation des risques qui peut simuler des attaques potentielles. Cela peut être utilisé pour surveiller divers environnements physiques et virtuels sur le réseau. [25]

3.1.2 Splunk Enterprise Security.

Splunk est l'une des solutions de gestion SIEM les plus populaires au monde. Ce qui le distingue de ses concurrents, c'est qu'il place l'analyse au cœur de son SIEM. Les données du réseau et de la machine peuvent être surveillées en temps réel pendant que le système recherche les vulnérabilités potentielles et identifie même les comportements

anormaux. La fonctionnalité Notables d'Enterprise Security affiche des alertes qui peuvent être affinées par l'utilisateur.

a. Principales caractéristiques

- Surveillance du réseau en temps réel
- Enquêteur d'actifs
- Analyse historique

L'interface utilisateur est très simple lorsqu'il s'agit de faire face aux menaces de sécurité. Lors de l'examen d'un événement, les utilisateurs peuvent commencer par un aperçu de base et cliquer pour accéder aux notes sur les événements passés. De même, les enquêteurs sur les actifs font un excellent travail pour signaler les comportements malveillants et prévenir les dommages futurs. [26]

3.1.3 LogRhythm NextGen SIEM Platform

LogRhythm est depuis longtemps un pionnier des solutions SIEM. De l'analyse comportementale à la corrélation des journaux et à l'intelligence artificielle pour l'apprentissage automatique, cette plateforme a tout pour plaire.

a. Principales caractéristiques

- Basé sur l'IA
- Gestion des fichiers journaux
- Analyse guidée

Le système est compatible avec un grand nombre d'appareils et de types de journaux. En termes de configuration de paramètres, la plupart des activités sont gérées via le gestionnaire de déploiement. Par exemple, On peut utiliser l'assistant

d'hébergement Windows pour filtrer les journaux Windows. Cela permet de mieux cibler ce qui se passe sur le Web [27].

3.1.4 McAfee Enterprise Security Manager

Système d'exploitation : Windows. VMWare ESX/ESXi, and Cloud

McAfee Enterprise Security Manager est considéré comme l'une des meilleures plates-formes SIEM pour l'analyse. Les utilisateurs peuvent collecter divers journaux sur divers appareils via le système Active Directory.

a. Principales caractéristiques

- Log consolidation
- Live monitoring

En termes de standardisation, le moteur de corrélation de McAfee peut facilement compiler des sources de données disparates. Cela facilite grandement la détection des incidents de sécurité lorsqu'ils surviennent.

Pour l'assistance, les utilisateurs ont accès au support technique McAfee Enterprise et au support technique McAfee Business. S'ils le souhaitent, les utilisateurs peuvent choisir de demander à leur responsable de compte d'assistance de visiter leur site deux fois par an. La plate-forme de McAfee s'adresse aux moyennes et grandes entreprises à la recherche d'une solution complète de gestion des incidents de sécurité. [28] .

3.2 Les Solutions SIEM Open Source

3.2.1 Mozilla Enterprise Defense Platform (MozDef)

Développé par Mozilla en 04 août 2014, ***MozDef*** est un logiciel open source basé sur d'autres technologies open source, y compris **Nginx, RabbitMQ, MongoDB, Firefox, Elasticsearch** qui vise à automatiser le processus de traitement des incidents de sécurité et à faciliter les activités en temps réel des gestionnaires d'incidents.

La plate-forme de défense Mozilla (MozDef) peut être vu comme un ensemble de micro-services qui vise à fournir des fonctionnalités SIEM traditionnelles :

- Analyse, stockage des événements et fichiers journaux.
- Faciliter la gestion des logs.
- Système d'alerte.

Le traitement frontal pour MozDef consiste à recevoir un événement/journal (en **JSON**) via **HTTP(S), AMQP(S) ou Amazon SQS** en effectuant la transformation des données, y compris la normalisation, l'ajout de métadonnées, etc. et en transmettant les données à Elasticsearch.

En interne, **MozDef** utilise **RabbitMQ** pour mettre en file d'attente les événements qui doivent encore être traités. Pour plus d'infos voir [29].

3.2.2 Elastic Stack (ELK Stack)

Développé par la société américano-néerlandaise Elastic NV le 12 février 2014, Elastic Stack appelé avant ELK Stack est une combinaison de quatre projets open source : **Elasticsearch, Logstash, Kibana et Beats**.

« *Elasticsearch est un moteur de recherche et d'analyse. Logstash est un pipeline de traitement de données côté serveur qui ingère simultanément des données provenant de plusieurs sources, les transforme, puis les envoie à une ‘ réserve ’ telle que Elasticsearch.* » [30].

Kibana donne aux utilisateurs la capacité d'afficher les données d'Elasticsearch sous forme de tableaux et des graphiques. Beats sont des plateformes Open-source et sous licence Apache. Ils transmettent les événements de diverses machines et de systèmes vers Logstash ou Elasticsearch. Ces quatre outils unis nous donnent une solution adéquate de rassemblement, traitement et examination des logs. [31].

3.2.3 Wazuh

Wazuh est une plateforme gratuite et open source pour la détection des menaces, le contrôle de sécurité, la réponse aux incidents et la conformité réglementaire. Développé par Wazuh Inc. le 23 Novembre 2015.

Wazuh agrège, inspecte et corrèle les données, cet outil offre également la possibilité de détection des menaces, de gestion de la conformité et de résolution d'incidents. Il peut être installé localement ou dans des clouds hybrides.

Wazuh n'est pas une solution SIEM monobloc mais un ensemble d'outils y compris Wazuh manager, Service d'inscription, RESTful API, Wazuh app, Rootcheck, Syscheck …etc.

En plus d'être une solution SIEM, Wazuh fournit d'autres fonctionnalités comme :

- Vérification de sécurité dans le cloud ou localement
- Détection des attaques
- Moniteur d'intégrité des fichiers
- Détection des failles de sécurité

Pour plus d'informations consulter [32].

3.2.4 SIEMonster

SIEMonster est une solution logicielle open source accessible aux petites et moyennes entreprises, elle permet de surveiller (monitoring) la sécurité personnalisable et évolutive et de surveiller tous les réseaux à un coût minime par rapport aux autres SIEM

SIEMonster Deep Learning détecte et élimine automatiquement les attaques grâce à l'apprentissage automatique et l'analyse du comportement basé sur l'homme. Il offre aussi des options de corrélation du comportement humain pour enrichir les alertes et minimiser les faux positifs. Il fournit également des renseignements sur les menaces en temps réel pour répondre aux cyberattaques [33].

SIEMonster utilise une collection d'outils de sécurité open source comprenant les modules Wazuh, Apache Ni-Fi, Cortex et The Hive pour fournir des rapports d'incidents, une corrélation avancée avec renseignements sur les menaces (threat Intelligence) et une réponse active fonctionnant tous ensemble [33].

3.3 Comparatif des solutions disponibles

Les tableaux suivants établissent un comparatif et résument les avantages et inconvénients des différentes solutions SIEM présentées auparavant.

	Plateformes	Déploiement
Mozdef	CentOS 6, RHEL 6 et Ubuntu 14	Déployé en tant que logiciel comme service (SaaS) dans des conteneurs Docker dans le cloud ou bien localement
ELK Stack	Windows, RHEL 7, CentOS 7, Oracle Linux 7, Ubuntu 14, SLES/openSUSE 15	Environnements virtuels, matériel physique, cloud privé ou cloud public
Wazuh	Amazon Linux 2, CentOS 7, Debian 8, Oracle Linux 6, RHEL 6	Déployé en tant qu'instance Amazon EC2 dans le cloud ou bien localement dans des environnements virtuels ou physique
SIEMonster	Mac, Ubuntu, CentOS et Debian	Sur le cloud à l'aide de conteneurs Docker, et sur des machines virtuelles et physiques
IBM QRadar	RED HAT, LINUX	Peut être déployé sur cloud, SaaS et sur site
SPLUNK	WINDOWS, LINUX, MAC, SOLARIS	Peut être déployé sur, SaaS et sur site
LogRhythm	Windows, Appliance, ou Cloud	Peut être déployé sur Cloud, SaaS, web, Windows
MCAFEE	WINDOWS ET MAC	Peut être déployé sur site, cloud ou hybride

Tableau 2 : Comparaison des solutions disponibles (Plateformes, Déploiement).
[25]–[33]

	Avantages	Inconvénients
Mozdef	Prend en charge les sources de données basées sur le cloud, notamment Amazon Web Services (AWS) CloudTrail et GuardDuty Il a d'excellentes capacités de tableau de bord car il utilise Kibana Simple et précis	Ne peut pas fournir une telle capacité d'un système à fonctionner en continu sans défaillance pendant une période de temps donnée La création de contenu n'est pas simple. Compétences en programmation nécessaires Pas d'options de clustering (regroupement)
ELK Stack	Support plusieurs langages de programmation (12) Analyse et visualisation des données en temps réel grâce au tableau de bord de visualisation des données "Kibana" Traiter, filtrer, corréler et améliorer les données de journal (Log) qu'il collecte	Configuration complexe Difficile de faire des analyses complexes sur les données (pas de fonctions statistiques avancées) Fonctionnalités d'alerte et de rapports basiques.
Wazuh	Variété d'options de recherche et de visualisation dans les fichiers journaux Surveillance de l'intégrité des fichiers Facile à mettre en œuvre et à intégrer avec d'autres solutions	Processus difficile pour couvrir certaines sources d'événements L'utilisation des ressources côté serveur est élevée Nécessite une amélioration dans l'investigation des cas d'utilisation car Il ne détecte pas toutes les attaques

Tableau 3 : Comparaison des solutions disponibles (Avantages, Inconvénients).
[29]–[33]

	Avantages	Inconvénients
IBM QRadar	Peut être utilisé sur une large gamme de systèmes d'exploitation, Linux, Windows, Unix et Mac Peut fonctionner comme une combinaison SIEM et HIDS (Host-based Intrusion Detection System) L'interface est facile à personnaliser et très visuelle	Nécessite des outils secondaires comme Graylog et Kibana pour une analyse plus approfondie En raison des capacités limitées de tri et de filtrage, des exportations CSV et des manipulations dans Excel sont fréquemment nécessaires.
SPLUNK	Peut utiliser l'analyse du comportement pour détecter les menaces qui ne sont pas découvertes via les journaux Excellente interface utilisateur, très visuelle avec des options de personnalisation faciles Priorisation facile des événements Axé sur l'entreprise	La tarification n'est pas transparente, Nécessite un devis du fournisseur Plus adapté aux grandes entreprises Utilise le langage de traitement de recherche (SPL) pour les requêtes
LogRhythm	Prise en charge efficace de la surveillance des données réseau, avec un grand nombre de signatures de flux applicatifs pour analyser les données de flux. Une interface moderne, hautement personnalisable et facile à utiliser qui tire parti de l'intelligence artificielle et de l'apprentissage automatique pour analyser le comportement	Absence d'option d'essai gratuit Configuration assez complexe Manque de fonctionnalité de support multiplateforme

Tableau 4 : Comparaison des solutions disponibles (Avantages, Inconvénients).
[25]–[28]

	Avantages	Inconvénients
SIEMonster	Fournit des renseignements sur les menaces en temps réel qui peuvent être utilisés pour arrêter les attaques au fur et à mesure qu'elles se produisent Construit sur des projets open source éprouvés Fournit des options de corrélation basées sur le comportement humain pour enrichir les alertes et minimiser les faux positifs	Version gratuite : 100 protections des terminaux et n'offre pas d'analyse du comportement des utilisateurs Absence de documentation en ligne Interfaces hétérogènes
MCAFEE	Utilise un puissant moteur de corrélation pour aider à trouver et à éliminer les menaces plus rapidement S'intègre bien dans les environnements Active Directory	L'interface est encombrée et souvent écrasante Est assez gourmand en ressources

Tableau 5 : Comparaison des solutions disponibles (Avantages, Inconvénients). [25]–[28]

3.4 Discussion

- **Mozdef** est un ensemble de micro-services qui fournissent un SIEM (Security Information and Event Management) open-source, simple et précis. Les principaux points faibles de cet outil sont que sa mise en place et son fonctionnement peuvent prendre du temps et être techniquement exigeants. Il manque également des options de haute disponibilité et des fonctionnalités clés en matière de rapports, regroupement (clustering) et de conformité.
- **ELK** est facile à utiliser et très intuitif et répond à un besoin clé dans l'espace SIEM. Il offre la possibilité de recueillir et d'examiner des données provenant de diverses sources, stocker ces données dans un centre de données centralisé Ce qui peut augmenter à mesure que de nouvelles données continuent d'arriver, et une collection d'outils dans le but d'analyser les logs. Cependant, le déploiement de la pile est un processus en plusieurs étapes qui peut conduire à une configuration complexe et nécessite des connaissances approfondies pour connaître l'ensemble de la pile élastique (elastic stack) et la déployer correctement.
- **Wazuh** est certainement l'outil le moins complet lorsqu'il s'agit de couvrir certaines sources d'événements car les fichiers journaux (logs) sont transmis uniquement par des agents Wazuh, mais il est simple d'utilisation et de configuration et facile à mettre en œuvre et à intégrer avec d'autres solutions.
- **SIEMonster** est l'outil le plus complet en ce qui concerne les fonctionnalités qu'il offre mais sa version gratuite est très limitée, le principal inconvénient de la version gratuite est qu'elle n'offre pas d'analyse comportementale des utilisateurs, d'apprentissage automatique et, surtout, de support. De plus, sa capacité de rapport est limitée à seulement deux rapports.
- **IBM QRADAR** est une solution ouverte et complète de détection et de réponse aux menaces qui élimine les menaces avancées, cette solution supporte une large gamme de systèmes d'exploitation. IBM Security QRadar permet aux entreprises d'automatiser l'enrichissement, la corrélation et l'investigation des menaces avec une IA spécialement conçue et des playbooks prédéfinis. Autrement, Pour faire des analyses complexes sur les données il faut ajouter des outils secondaires comme

Graylog et Kibana, L'un des principaux inconvénients d'IBM Security QRadar est sa valeur élevée, c'est un logiciel coûteux mais avec de grands avantages.

- **SPLUNK** est un logiciel puissant et efficace pour surveiller et analyser les données et les événements, avec une excellente interface utilisateur, Splunk permet à l'utilisateur d'avoir rapidement un aperçu des aspects les plus importants de ses données, de fournir des informations en temps réel et de les utiliser pour identifier les menaces qui nécessitent une action urgente. SPLUNK fournit également une analyse de comportement pour détecter les menaces qui ne sont pas découvertes via les journaux. Cependant, la dépendance de Splunk au langage de traitement de recherche (SPL) dans les requêtes complexes peut empêcher le système de fournir la bonne réponse, il existe un risque d'erreur dans sa prédiction et ses résultats.
- **LogRhythm** offre une plate-forme polyvalente et étendue avec des fonctionnalités SIEM de base ainsi que des capacités complémentaires de surveillance de l'hôte et du réseau. Il fournit également une interface moderne, hautement personnalisable et facile à utiliser pour bénéficier de toutes les fonctionnalités des tableaux de bord. Les principaux points faibles de cette solution est qu'elle est entièrement payante, elle ne propose pas de versions d'essai gratuites, de plus elle nécessite un énorme investissement de temps initial pour la configurer et continuera à nécessiter du temps pour surveiller en permanence l'utilisation des systèmes, contrôler les sources, créer de nouvelles règles d'alerte pour rester au courant des tendances des menaces émergentes et gérer le bruit supplémentaire généré par ces alertes (faux positifs).
- **MCAFEE** SIEM fournit des informations et des intégrations exploitables pour hiérarchiser, enquêter et répondre aux menaces grâce à son puissant moteur de corrélation, comparée à d'autres solutions, elle est conviviale et facile à utiliser et à déployer. Mais cette solution manque certaines fonctionnalités que l'on retrouve dans les SIEM commerciales actuelles telles que l'analyse du comportement des utilisateurs, l'automatisation, la capture de paquets ou l'analyse de réseau. De plus, il manque une optimisation des performances du système en général car il est assez gourmand en ressources.

3.5 Choix de la solution et justification

Suite à une étude des différentes solutions disponibles mentionnées dans les tableaux 2, 3, 4, 5 et en raison d'une étude comparative en termes de systèmes d'exploitation pris en charge, de déploiement, de type, de fonctionnalités principales et de limitations, nous avons choisi de mettre en œuvre ce projet en utilisant Elastic Stack / ELK. Cet ensemble de produits open source est davantage utilisé par de nombreuses entreprises populaires telles que LinkedIn, Netflix, Uber et StackOverflow.

4 Architecture de la solution

Elastic Stack précédemment appelé ELK Stack est une combinaison de quatre projets open source : Elasticsearch, Logstash, Kibana et Beats, qui fournit une solution complète pour analyser les fichiers journaux de tout type de source dans n'importe quel format en temps réel, Elle permet de rechercher, analyser et visualiser en profondeur le fichier journal généré à partir de différentes machines [34]

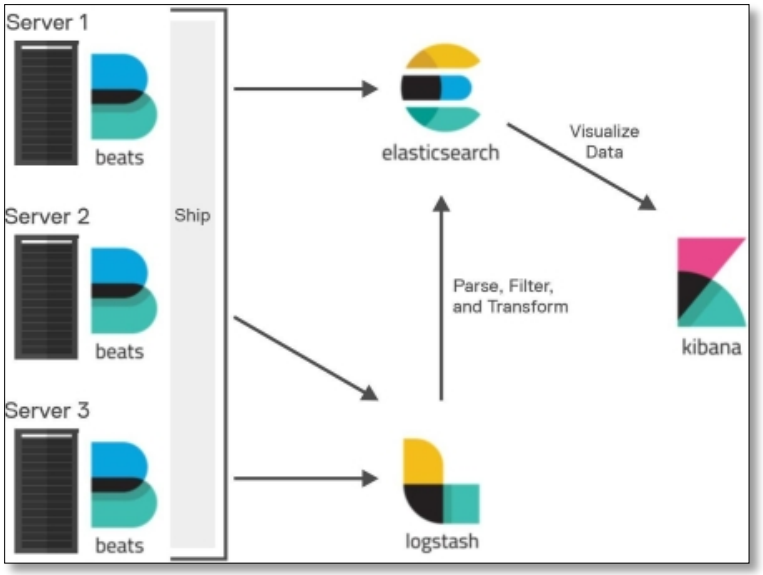


Figure 8 : Architecture de la pile elastic [35]

4.1.1 Elasticsearch

Elasticsearch est un outil open source lancé par Elasticsearch N.V. en 2010, cet outil accepte plusieurs formats de données, notamment textuelles, numériques, références spatiales, etc. Ses API REST peu compliquées, sa nature (exécution de recherche répartie), sa capacité à parcourir une grande quantité de données en peu de temps et sa capacité à évoluer font d'Elasticsearch le composant central de la Pile Elastique (un ensemble d'outils gratuits et open source pour ingérer, enrichir, stocker, analyser et visualiser des données) [36].

Elasticsearch fournit une recherche de langue basée sur des résultats de pertinence à l'aide d'algorithme TF/IDF (Term Frequency–Inverse Document Frequency) (Pour plus d'informations consulter l'**Annexe XI**). Il comprend également un processus de saisie semi-automatique. Il enregistre les données au format JSON et fournit une indexation et une recherche basées sur des attributs individuels [37].

4.1.2 Logstash

Logstash est un moteur de collecte de données côté serveur léger, et open source qui capture les données de différentes sources, les convertit au format correct et les envoie à la destination souhaitée. Il s'agit du pipeline de données le plus populaire pour Elasticsearch, car il offre des capacités de traitement de logs efficaces et plus de 200 modules d'extension open source pré-construits pour indexer simplement les journaux.

Logstash facilite l'ingestion de données non structurées à partir de diverses sources, notamment les logs du système, les logs du site Web, les logs du serveur d'applications, etc. Il fournit pareillement des filtres prédéfinis pour convertir facilement les types de données courants en index dans Elasticsearch sans créer de pipeline de transformation de données personnalisé [38].

4.1.3 Beats

Beats sont des expéditeurs de données open source qui fonctionnent comme des agents sur les serveurs pour envoyer des données opérationnelles directement à Elasticsearch ou via Logstash, où on peut poursuivre le traitement et l'amélioration des données, avant de les visualiser dans Kibana. [39]

Elastic fournit Beats pour la capture de :

- Les données d'audit en utilisant Auditbeat
- Fichiers journaux en utilisant Filebeat
- Données cloud en utilisant Functionbeat
- Surveillance de la disponibilité en utilisant Heartbeat
- Métriques en utilisant Metricbeat
- Trafic réseau en utilisant Packetbeat
- Événements Windows en utilisant Winlogbeat

4.1.4 Kibana

Kibana est une application frontale gratuite et ouverte qui se trouve au-dessus de la Pile Elastic, offrant des capacités de recherche et de visualisation des données pour les données indexées dans Elasticsearch. Kibana agit également comme interface utilisateur pour la surveillance, la gestion et la sécurisation d'un cluster Elastic Stack (Un cluster se compose de nombreux éléments travaillant ensemble de telle manière qu'ils peuvent être considérés comme un seul système). [40]

En raison de son appartenance à la Pile Elastique, Kibana est idéal pour les cas d'utilisation suivants [40] :

- Rechercher, visualiser et analyser les informations indexées dans Elasticsearch, notamment :
  - Journalisation et analyse
  - Indicateurs des infrastructures.
  - Examen et visualisation de données géographiques

  - L'analyse de la sécurité
- Surveillance, administration et sécurité des instances Elastic Stack via l'interface web
- Accès centralisé aux solutions déployées sur la Pile Elastique.

5 Conclusion

Après avoir pris toutes les notions nécessaires à la compréhension du projet dans le chapitre précédent et suite à une présentation de l'organisme d'accueil et de ses différentes missions et objectifs, Nous avons réalisé une étude comparative entre les différentes solutions SIEM (Security Information and Event Management) open source et commerciales afin de choisir les outils qui sont les mieux adaptés à notre projet.

Chapitre 3

CONCEPTION

1 Introduction

Ce chapitre décrit un modèle conceptuel du fonctionnement du système. Nous commencerons par présenter les diagrammes de séquence, le diagramme de cas d'utilisation puis nous expliquerons chaque cas d'utilisation de ce diagramme.

2 Conception du système

La solution SIEM, comme le montre la figure ci-dessous, collecte les journaux via des agents élastiques. Ces journaux proviennent de différents services, applications, composants système, etc. et de différents systèmes d'exploitation. Une fois ces journaux stockés, le SIEM commence à analyser les données pour détecter et découvrir les menaces et identifier les vulnérabilités de sécurité (failles), et alerter l'administrateur système lorsque des événements suspects se produisent.

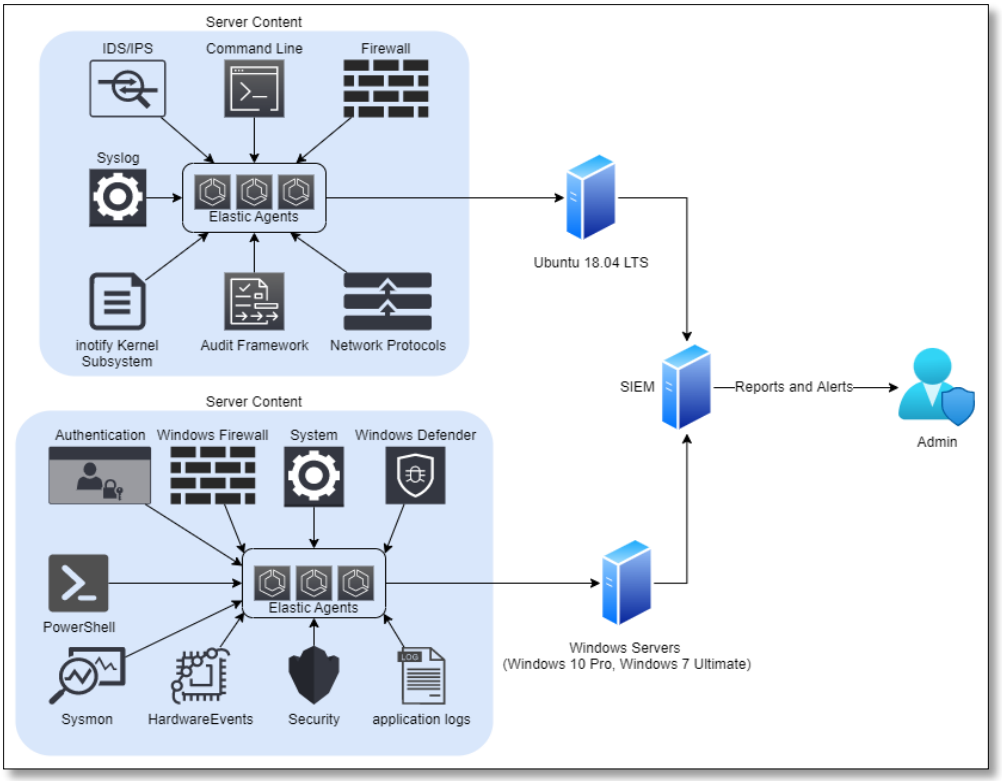


Figure 9 : Schéma descriptif du projet.

3 Diagramme de cas d'utilisation

Dans le langage UML, les diagrammes de cas d'utilisation modélisent le comportement d'un système et capturent les exigences du système.

Les diagrammes de cas d'utilisation décrivent la fonctionnalité générale et la portée d'un système. Les cas d'utilisation et les acteurs dans un diagramme de cas d'utilisation donnent un aperçu de ce que fait le système et comment les acteurs l'utilisent, mais ne montrent pas le fonctionnement interne du système [41]**.**

Dans cette section, nous présenterons le diagramme de cas d'utilisation global de notre solution. Par la suite nous détaillerons ce diagramme en expliquant chaque cas d'utilisation afin de mieux comprendre le fonctionnement du système en adoptant ce modèle, qui prend des informations communes liées au cas d'utilisation :

- **Libellé** du cas d'utilisation.
- **Acteur** lié au cas d'utilisation.
- **Description** : brève définition du cas d'utilisation.
- **Précondition(s)** : les conditions qui doivent être remplies pour que le cas d'utilisation commence à se produire.
- **Postcondition(s)** : conditions qui doivent être remplies lorsque le cas d'utilisation se termine.
- **Scénarios :**
  - **Scénario nominal** : Représente la série d'étapes qui conduisent le plus souvent au cas d'utilisation actuel.
  - **Scénario alternatif** : Représente la série d'étapes qui conduisent également au cas d'utilisation actuel.
  - **Scénario d'erreur :** Représente la série d'étapes qui ne permettent pas de terminer le cas d'utilisation actuel (les postconditions ne peuvent pas toutes être vraies).

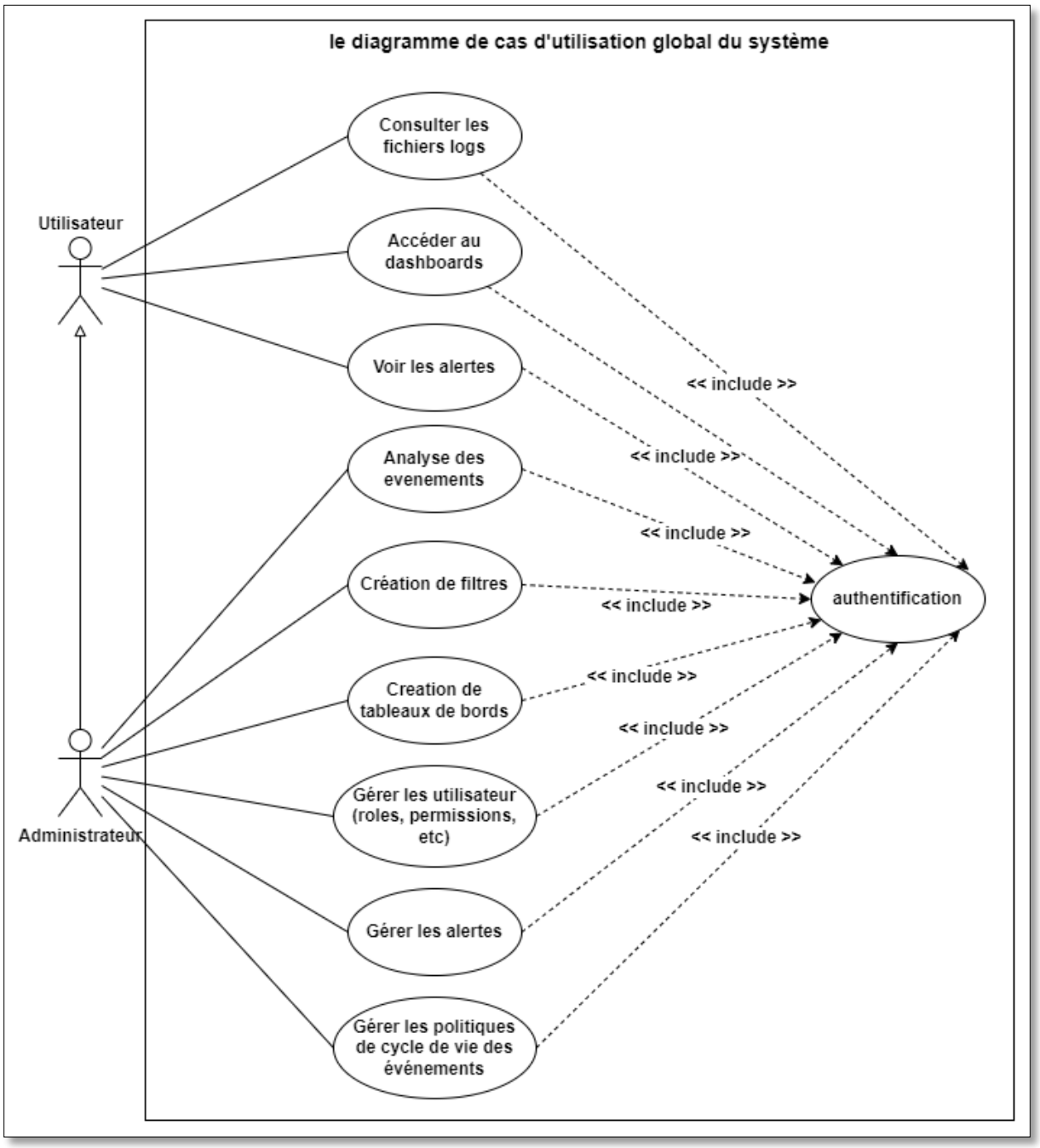


Figure 10 : Diagramme de cas d'utilisation.

3.1 Cas d'utilisation Analyse des évènements

Libellé	Analyse des événements
Acteur	Administrateur
Description	Permettre à l'administrateur d'observer tous les champs d'événement
Précondition(s)	• Le système doit avoir reçu des journaux • Les évènements doivent être structurés et indexées
Postcondition(s)	• Les résultats seront exposés sous forme de table ou sous format JSON

Scénarios	• **Scénario nominal**
	  ○ L'utilisateur sélectionne un événement et observe ses différents champs
	• **Scénario alternatif**
	  ○ L'utilisateur sélectionne un événement et trouve des champs vides ou manquants
	• **Scénario d'erreur**
	  ○ L'utilisateur ne trouve aucun événement

Tableau 6 : Description textuelle de cas d'utilisation « Analyse des événements ».

3.2 Cas d'utilisation Création des filtres

Libellé	Création des filtres
Acteur	Administrateur
Description	Permettre à l'administrateur de créer une recherche filtrée pour capturer des événements spéciaux
Précondition(s)	• Le système doit avoir reçu des journaux • Les évènements doivent être structurés et indexées • L'utilisateur doit écrire une requête pour que le système effectue la recherche correcte
Postcondition(s)	• Le résultat de la recherche sera exposé à l'utilisateur sur l'interface utilisateur de Kibana

Scénarios	• **Scénario nominal**
	  ○ L'administrateur donne une requête de recherche dans la barre de recherche et reçoit des événements filtrés
	• **Scénario alternatif**
	  ○ L'administrateur donne une requête de recherche dans la barre de recherche et ne reçoit aucun événement.
	• **Scénario d'erreur**
	  ○ L'administrateur donne une requête de recherche non valide dans la barre de recherche.

Tableau 7 : Description textuelle de cas d'utilisation « Création des filtres ».

3.3 Cas d'utilisation Création de tableaux de bords

Libellé	Création de tableaux de bords
Acteur	Administrateur
Description	Permettre à l'acteur de créer/ajouter une collection de visualisations dans un tableau de bord
Précondition(s)	• Le système doit avoir reçu des journaux • Les évènements doivent être structurés et indexées • La présence de visualisations déjà créées ou importées
Postcondition(s)	• Le résultat de la création sera exposé à l'utilisateur sur l'interface utilisateur de Kibana

Scénarios	• **Scénario nominal**
	  ○ L'utilisateur ajoute des visualisations et crée un tableau de bord
	• **Scénario alternatif**
	  ○ Un tableau de bord vide.
	• **Scénario d'erreur**
	  ○ Échec de la création du tableau de bord (Error on saving 'dashboard')

Tableau 8 : Description textuelle de cas d'utilisation « Création de tableaux de bords ».

3.4 Cas d'utilisation Gérer les utilisateurs

Libellé	Gérer les utilisateurs (rôles, permissions, etc.)
Acteur	Administrateur
Description	Permettre à l'acteur de créer, mettre à jour, supprimer des rôles et des privilèges accordés aux utilisateurs du système
Précondition(s)	• La présence des rôles, privilèges ou utilisateurs dans le cas de mettre à jour ou de suppression • Le privilège de cluster **manage security** est requis pour accéder
Postcondition(s)	• Le résultat sera présenté à l'admin lors du filtrage de la liste des utilisateurs.

Scénarios	• **Scénario nominal**
	  ○ L'administrateur crée un nouveau rôle et l'attribue à un utilisateur existant
	• **Scénario alternatif**
	  ○ L'administrateur crée un nouvel utilisateur avec des champs manquants (email, numéro de téléphone, etc.)
	• **Scénario d'erreur**
	  ○ L'administrateur crée un nouvel utilisateur avec des champs invalides ou manquants (mot de passe, Nom d'utilisateur, rôles, etc.)

Tableau 9 : Description textuelle de cas d'utilisation « Gérer les utilisateurs ».

3.5 Cas d'utilisation Gérer les alertes

Libellé	Gérer les alertes
Acteur	Administrateur
Description	Permettre à l'acteur de créer, mettre à jour, supprimer des alertes basées sur des règles de corrélation
Précondition(s)	• Le système doit avoir reçu des journaux • Les évènements doivent être structurés et indexées • Les règles de corrélation doivent être claires et définie • La sécurité de la couche de transport (TLS) entre Kibana et Elasticsearch doit être activée
Postcondition(s)	• Les résultats seront présentés dans la page Sécurité de l'interface utilisateur de Kibana

Scénarios	• **Scénario nominal**
	  ○ L'acteur met à jour une alerte afin de réduire les faux positifs
	• **Scénario alternatif**
	  ○ L'acteur crée une alerte avec des règles de corrélation non valides
	• **Scénario d'erreur**
	  ○ Echec de création d'une alerte

Tableau 10 : Description textuelle de cas d'utilisation « Gérer les alertes ».

3.6 Cas d'utilisation Gérer les politiques de cycle de vie des évènements

Libellé	Gérer les politiques de cycle de vie des évènements
Acteur	Administrateur
Description	Permettre à l'administrateur de créer et de gérer des politiques de cycle de vie des événements en fonction d'exigences de conformité prédéfinies.
Précondition(s)	• La présence d'index de données • La présence d'un modèle d'index • Elasticsearch et Kibana doivent être en cours d'exécution
Postcondition(s)	• Création de nouveaux index • Changement d'état de l'index à travers différentes phases

Scénarios	• **Scénario nominal**
	  ○ L'acteur met à jour la durée de la phase de suppression de l'index en fonction des nouvelles exigences de conformité.
	• **Scénario alternatif**
	  ○ L'acteur crée une politique de cycle de vie sans aucune phase (l'index sera stocké en permanence)
	• **Scénario d'erreur**
	  ○ L'index ne passe pas à la phase suivante (problème de santé de l'index)

Tableau 11 : Description textuelle de cas d'utilisation « Gérer les politiques de cycle de vie des évènements ».

3.7 Cas d'utilisation Consulter les fichiers logs

Libellé	Consulter les fichiers log
Acteur	Utilisateur
Description	Permettre à l'utilisateur de consulter les fichiers journaux et d'observer leur contenu.
Précondition(s)	• Le système doit avoir reçu des journaux • Les évènements doivent être structurés et indexées • L'utilisateur doit écrire une requête pour que le système effectue la recherche correcte
Postcondition(s)	• Les résultats seront exposés sous forme de table ou sous format JSON sur l'interface utilisateur de Kibana.

Scénarios	• **Scénario nominal**
	  ○ L'utilisateur sélectionne un fichier journal et observe ses différents champs
	• **Scénario alternatif**
	  ○ L'utilisateur sélectionne un fichier journal et trouve des champs vides ou manquants
	• **Scénario d'erreur**
	  ○ L'utilisateur ne trouve aucun fichier

Tableau 12 : Description textuelle de cas d'utilisation « Consulter les fichiers logs ».

3.8 Cas d'utilisation Accéder au dashboards

Libellé	Accéder au dashboards
Acteur	Utilisateur
Description	Permettre à l'utilisateur d'accéder aux différents tableaux de bord créés par l'administrateur.
Précondition(s)	• Le système doit avoir reçu des journaux • Les évènements doivent être structurés et indexées • La présence de visualisations déjà créées ou importées
Postcondition(s)	• Le résultat sera exposé à l'utilisateur sur l'interface utilisateur de Kibana.

Scénarios	• **Scénario nominal**
	  ○ L'utilisateur accède au tableau de bord et commence à visualiser les données
	• **Scénario alternatif**
	  ○ Un tableau de bord vide.
	• **Scénario d'erreur**
	  ○ Erreur de visualisation (erreur : esaggs) lors de la tentative de chargement d'un tableau de bord

Tableau 13 : Description textuelle de cas d'utilisation « Accéder au dashboards ».

3.9 Cas d'utilisation Voir les alertes

Libellé	Voir les alertes
Acteur	Utilisateur
Description	Permettre à l'utilisateur de visualiser les alertes et leurs règles de corrélation déployées par l'administrateur du système
Précondition(s)	• Le système doit avoir reçu des journaux • Les évènements doivent être structurés et indexées • Les règles de corrélation doivent être claires et définie • La sécurité de la couche de transport (TLS) entre Kibana et Elasticsearch doit être activée.
Postcondition(s)	• Les résultats seront présentés dans la page Sécurité de l'interface utilisateur de Kibana

Scénarios	• **Scénario nominal**
	  ○ L'utilisateur accède à la page de sécurité afin de trouver des activités suspectes
	• **Scénario alternatif**
	  ○ L'utilisateur navigue vers la page de sécurité et ne trouve aucune alerte
	• **Scénario d'erreur**
	  ○ Obtient une erreur de socket TLS lors de la tentative d'accès à la page de sécurité

Tableau 14 : Description textuelle de cas d'utilisation « Voir les alertes ».

4 Diagramme de séquence

Les deux figures ci-dessous montrent un diagramme de séquence décrivant un scénario d'exécution de requêtes par l'administrateur pour afficher, filtrer les données dans l'interface utilisateur Kibana dans le cas de journaux non structurés analysés, indexées par logstash (figure 11) ou de journaux structurés envoyés directement à Elasticsearch (sous format JSON) (figure 12).

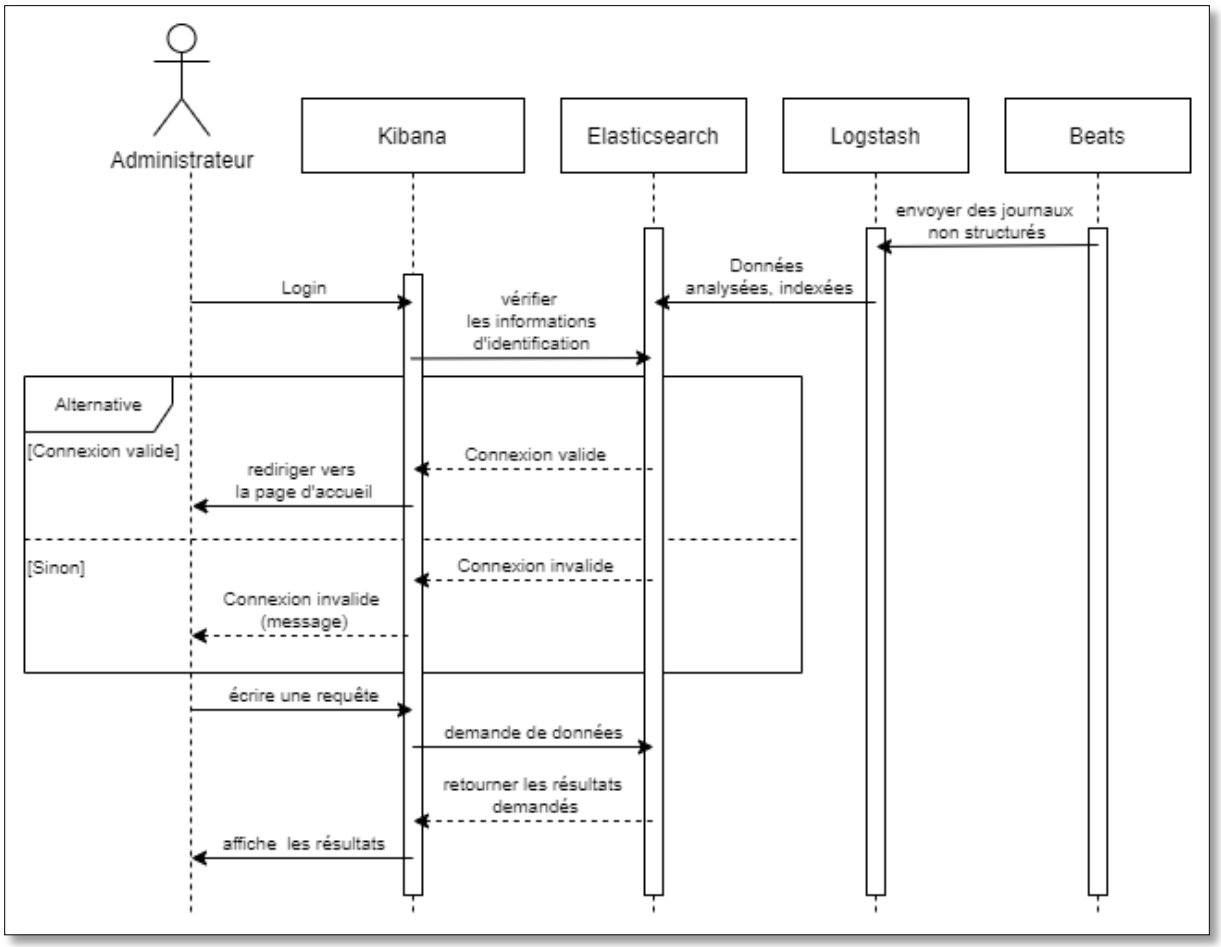


Figure 11 : Diagramme de séquence, exécution de requêtes par l'administrateur dans Kibana en cas de journaux non structurés.

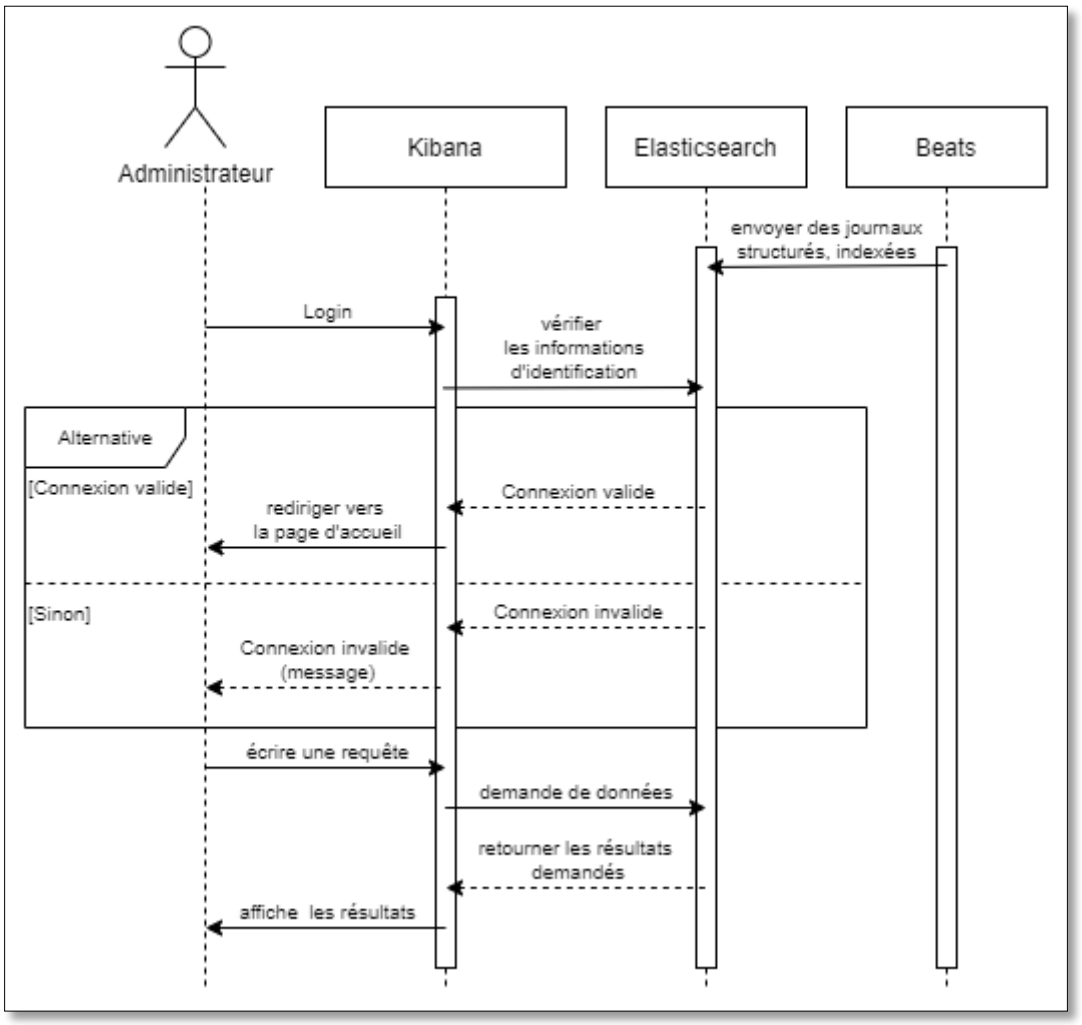


Figure 12 : Diagramme de séquence, exécution de requêtes par l'administrateur dans Kibana en cas de journaux structurés.

5 Conclusion

Ce chapitre a été consacré à la conception de la solution qui a été proposée pour répondre au besoin exprimé par l'organisme d'accueil en présentant les diagrammes de séquence, le diagramme de cas d'utilisation et en expliquant les différents cas d'utilisation. Ce qui nous a permis d'identifier l'architecture générale de notre solution.

Chapitre 4

Déploiement de la solution

1 Introduction

Après avoir mené des études conceptuelles dans le chapitre précédent, nous abordons dans cette partie l'implémentation de notre solution SIEM (Security Information and Event Management). Nous commencerons par définir les outils et les différentes technologies que nous avons utilisés. Par la suite la mise en œuvre de solution va être présentée d'une manière plus détaillée.

2 Environnement de travail

Dans cette section, nous allons définir l'environnement matériel et logiciel et l'architecture de la solution réalisée.

2.1 Environnement Matériel

Pour la réalisation du projet, nous avons utilisé un ordinateur de bureau ayant les caractéristiques suivantes :

Processeur	AMD Ryzen 5 PRO 4650G (6C/12T, 3.70 GHz, 8MB Cache, 65W)
Carte graphique	AMD RX Vega 7 2 GB up to 5 GB
Mémoire vive	16gb ddr4-2666mhz (2x8gb)
Système d'exploitation	Windows 10 Professional 64-bit
Disque dur	512 GB SSD

Tableau 15 : Environnement Matériel du projet.

2.2 Environnement Logiciel

2.2.1 VirtualBox

Oracle VM (Virtual Machine) VirtualBox est une application de virtualisation multiplateforme qui s'installe sur des ordinateurs Intel ou AMD, qu'ils exécutent n'importe quel type de système d'exploitation (Windows, Mac OS X, Linux ou Oracle Solaris), il étend les capacités de l'ordinateur afin qu'il puisse exécuter plusieurs systèmes d'exploitation, à l'intérieur de plusieurs machines virtuelles, en même temps. [42]

Oracle VM VirtualBox semble simple mais aussi très puissant. Il peut s'exécuter partout, des petits systèmes embarqués ou des machines de bureau jusqu'aux déploiements de centres de données et même aux environnements Cloud. [42]

Notre environnement virtuel est détaillé dans les tableaux ci-dessous :

Threads alloués	4 threads
Système d'exploitation	Ubuntu 18.04
Mémoire vive allouée	6 GB RAM
Espace disque	100 GB
Technologies et Outils	Elasticsearch 7.17, Kibana 7.17, Logstash 7.17, Elastic Agent 7.17, Filebeat 7.17, Metricbeat 7.17, Gnu nano

Tableau 16 : Caractéristiques du Serveur Principale (Machine SIEM).

Threads alloués	2 threads
Système d'exploitation	Ubuntu 18.04
Mémoire vive allouée	2 GB RAM
Espace disque	40 GB
Technologies et Outils	Filebeat 7.17, Auditbeat 7.17, Packetbeat7.17, Suricata, Gnu nano

Tableau 17 : Caractéristiques du Serveur de test 1.

Threads alloués	2 threads
Système d'exploitation	Ubuntu 18.04
Mémoire vive allouée	2 GB RAM
Espace disque	40 GB
Technologies et Outils	Hping3, Nmap, Gnu nano

Tableau 18 : Caractéristiques du Serveur de test 2 (Attaquant).

Threads alloués	2 threads
Système d'exploitation	Windows 7 Ultimate
Mémoire vive allouée	2 GB RAM
Espace disque	40 GB
Technologies et Outils	Elastic Agent 7.17, Eicar test file, Sublime Text

Tableau 19 : Caractéristiques du Serveur de test 3.

2.2.2 Java

Java est un langage de codage informatique qui sert à programmer des applications client-serveur. Java a été créé en 1995 par la société technologique américaine "Sun Microsystems", la syntaxe et le formatage java sont inspirés du langage C/C++. Les applications Java sont des programmes multiplateformes grâce à Java Runtime Environment (JRE) et Java Virtual Machine (JVM). La technologie Java est l'un des termes les plus utilisés en informatique et sur le Web. Il couvre diverses normes, communautés d'affaires et logiciels. [43]

2.2.3 Éditeur de fichiers .yml

a. GNU nano

Les environnements Unix et Linux s'appuient sur ***GNU nano*** comme éditeur de texte par défaut. Comme tout éditeur de texte, GNU nano possède la coloration syntaxique, de multiples tampons, la prise en charge de la recherche et du remplacement des expressions régulières, la vérification orthographique, l'encodage UTF-8, etc. [44]

b. Sublime Text

Sublime Text est un éditeur de code utilisé par les programmeurs du monde entier. Le programme possède de nombreuses fonctionnalités, telles que la reconnaissance du type de fichier, l'auto-indentation, la coloration syntaxique, les barres latérales, les macros, les plugins et les packages, pour aider les programmeurs à travailler avec des bases de code plus importantes. [45]

2.2.4 Pile Elastique (Elastic Stack)

La Pile Elastique est un ensemble d'outils qui peuvent de récupérer des données de diverses sources, structurées ou non structurées, et les examiner en détail et les afficher Immédiatement. Pour plus de détails, voir la **section 3.2.2 et section 4 du chapitre 2.**

3 Implémentation de la solution

3.1 Installation de Sysmon

Mark Russinovich et Thomas Garnier définissent ***Sysmon*** de cette manière : « *System Monitor (Sysmon) est un Windows service système et pilote d'appareil qui, une fois installé sur un système, reste résident dans les redémarrages du système pour surveiller et journaliser l'activité du système dans le journal des événements Windows. Il fournit des informations détaillées sur les créations de processus, les connexions réseau et les modifications apportées à l'heure de création de fichier* » [46]**.**

Sysmon est installé sur la machine principale "Windows 10" pour donner plus de détails aux événements qui seront envoyés ultérieurement à la machine SIEM (notamment Elasticsearch), pour plus d'information, voir **Annexe I.**

3.2 Installation d'Elasticsearch

Avant d'installer Elastic Stack (précisément Elasticsearch), Java doit être installé sur le système ou on va utiliser Elasticsearch, Ensuite, nous configurons la variable d'environnement **JAVA_HOME** dans le fichier **etc/environment** (**Annexe II**).

Une fois Java installé, nous procédons au processus d'installation et de configuration d'Elasticsearch, dans le fichier **/etc/elasticsearch/elasticsearch.yml**, nous spécifions l'adresse IP et le port du système pour qu'Elasticsearch écoute le trafic HTTP, et spécifions un cluster à nœud unique pour empêcher Elasticsearch de rejoindre d'autres nœuds ou d'autoriser d'autres nœuds à rejoindre le cluster.

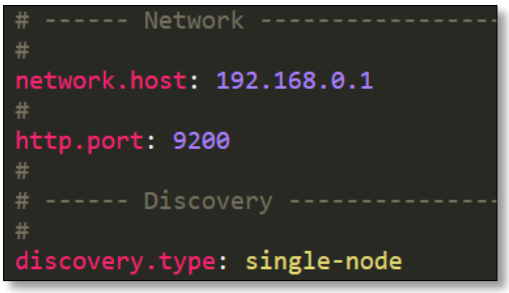


Figure 13 : Fichier de configuration d'Elasticsearch.

3.3 Installation de Kibana

Pour visualiser les données stockées dans Elasticsearch, nous devons installer et configuer Kibana, dans le fichier **/etc/kibana/kibana.yml**, nous spécifions l'adresse IP du système, un numéro de port et ajoutons l'URL de l'instance Elasticsearch.

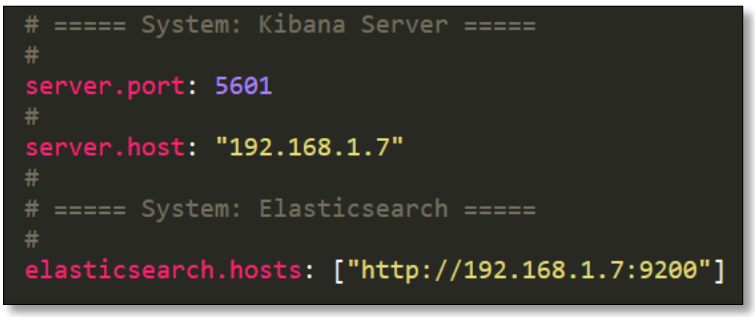


Figure 14 : Fichier de configuration de Kibana.

3.4 Installation de Logstash

Pour collecter les journaux et les analyser, nous devons installer Logstash. Les journaux collectés et filtrés sont stockés dans Elasticsearch. **(Annexe III)**

Une fois Logstash installé avec succès, nous configurons les sections d'entrée, de filtre et de sortie afin que Logstash reçoive les journaux de différents Beats, les analyse et les transmette à Elasticsearch. **(Annexe III)**

3.4.1 Exemple

Pour mieux comprendre l'analyse des logs, voici un exemple de log Syslog (System Logging Protocol) :

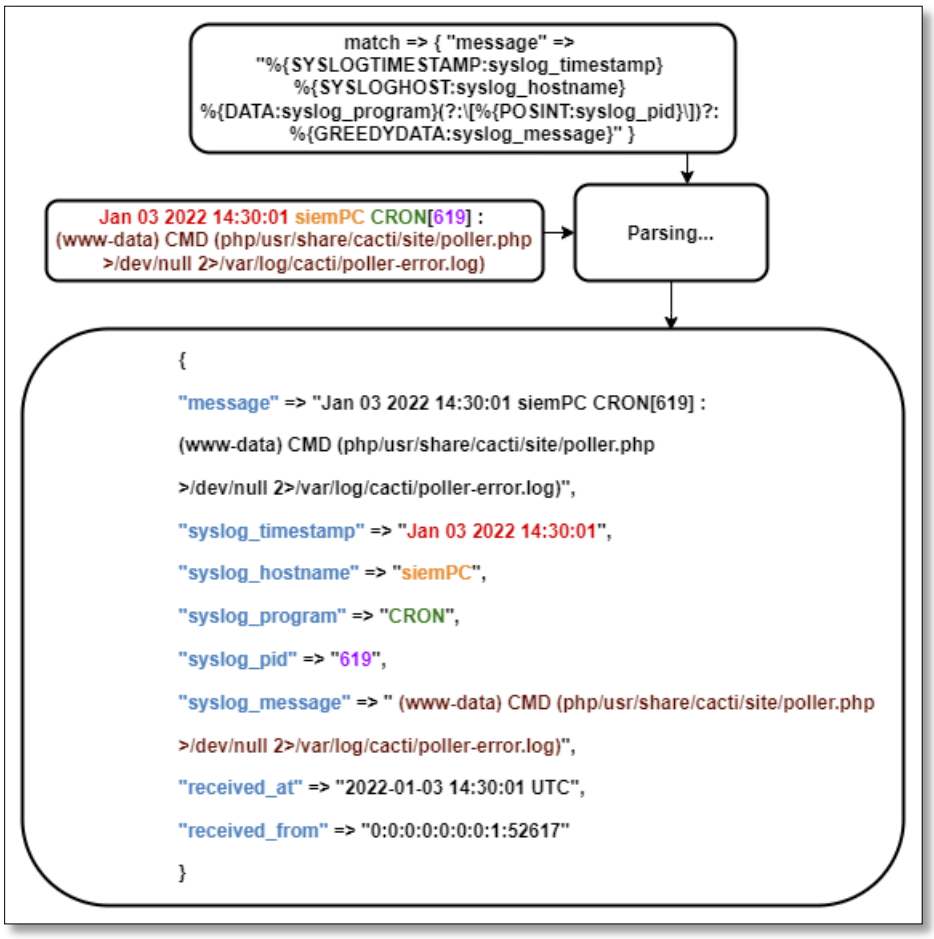


Figure 15 : Exemple d'analyse de journal.

3.5 Affectation des rôles aux utilisateurs

Les fonctionnalités de sécurité d'Elastic Stack ajoutent une autorisation, qui est le processus permettant de déterminer si un utilisateur qui génère une demande entrante est autorisé à exécuter la demande.

Ces fonctionnalités appliquent un rôle par défaut à tous les utilisateurs qui leur permet d'accéder au point de terminaison d'authentification, de modifier leurs propres mots de passe et d'obtenir des informations sur eux-mêmes. Pour en savoir plus sur la façon de créer un utilisateur et de lui affecter un rôle, voir **l'annexe IV**

3.5.1 Modèle RBAC

La figure suivante montre un mécanisme de contrôle d'accès basé sur les rôles (RBAC), qui nous permet d'autoriser les utilisateurs en attribuant des privilèges aux rôles et en attribuant des rôles aux utilisateurs ou aux groupes.

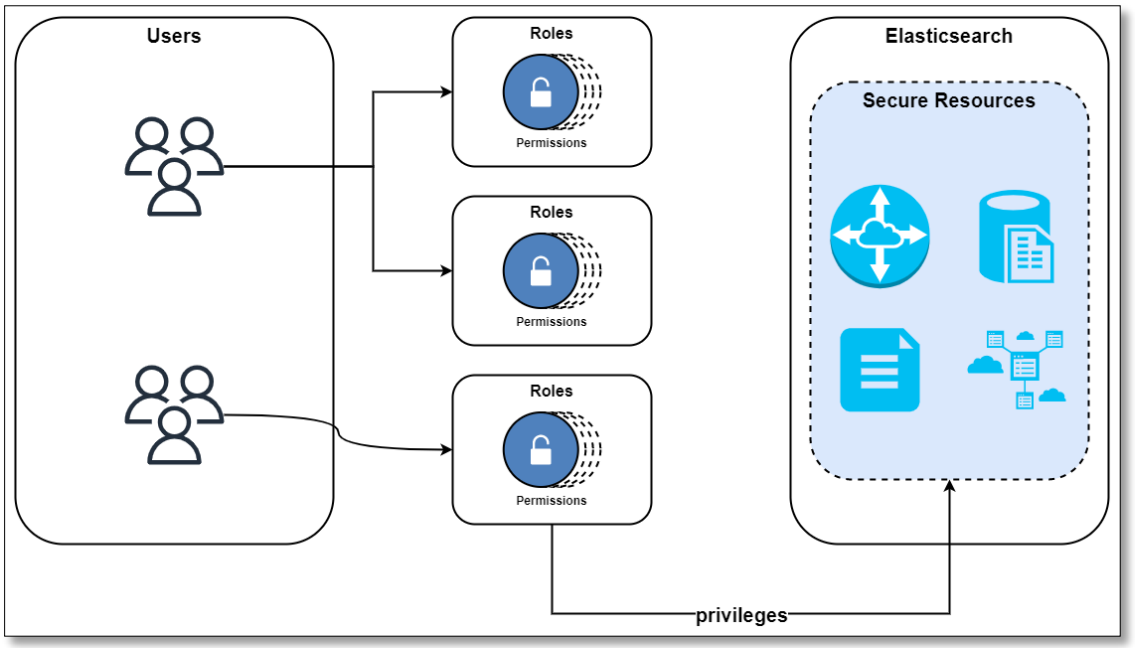


Figure 16 : Le RBAC dans Elasticsearch.

Dans la figure ci-dessus :

- **Users** représentent des utilisateurs authentifiés,
- **Roles** sont un ensemble d'autorisations,
- **Permissions** sont un ensemble d'un ou plusieurs privilèges,
- **Privileges** sont des actions qu'un utilisateur peut effectuer sur une ressource sécurisée,
- **Secure Ressources** sont des ressources dont l'accès est restreint. Les index, les alias, les documents, les champs, les utilisateurs et le cluster Elasticsearch lui-même sont tous des exemples d'objets sécurisés.

3.6 Installation de Beats

3.6.1 Auditbeat

Auditbeat est un expéditeur léger qui s'installe sur les serveurs pour auditer les activités des utilisateurs et des processus sur le système. Par exemple, on peut utiliser Auditbeat pour détecter les modifications apportées aux fichiers critiques, tels que les fichiers binaires et les fichiers de configuration, et identifier les violations potentielles des politiques de sécurité [39]**.** Nous avons choisi Auditbeat comme exemple de configuration de beats, ces étapes sont correctes pour chaque beat supplémentaire que nous voulons configurer.

Une fois installé, nous naviguons vers le fichier **etc/auditbeat/auditbeat.yml**, nous définissons l'hôte et le port où Auditbeat peut trouver l'installation d'Elasticsearch, et définissons le nom d'utilisateur et le mot de passe d'un utilisateur autorisé à utiliser Auditbeat, comme expliqué dans la figure ci-dessous.

```
# ---------------------------- Elasticsearch Output ----------------------------
output.elasticsearch:
  # Array of hosts to connect to.
  hosts: ["http://192.168.1.7:9200"]

  # Protocol - either `http` (default) or `https`.
  protocol: "http"

  # Authentication credentials - either API key or username/password.
  #api_key: "id:api_key"
  username: "admin"
  password: "m2SSI2022"

# ------------------------------ Logstash Output -------------------------------
```

Figure 17 : Fichier de configuration Auditbeat -Elasticsearch Output.

Ensuite, nous devons autoriser Auditbeat à se connecter au point de terminaison Kibana afin d'utiliser ses tableaux de bord préconstruits, en fournissant l'adresse IP et le numéro de port utilisé lors de l'installation de Kibana (**section 3.3 du Chapitre 4**), les informations d'identification d'un utilisateur qui a des privilèges pour écrire dans Kibana.

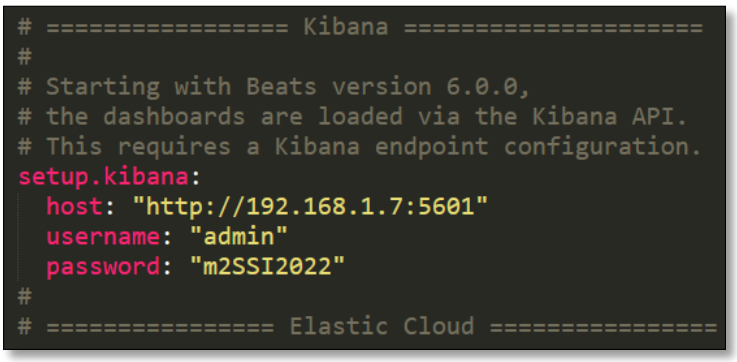


Figure 18 : Fichier de configuration Auditbeat -Kibana Output.

La dernière étape de l'installation d'Auditbeat consiste à configurer les modules de collecte de données. La figure suivante montre le module d'intégrité des fichiers configuré pour générer des événements une fois qu'un changement se produit dans l'un des chemins spécifiés (créer, supprimer, mettre à jour le fichier).

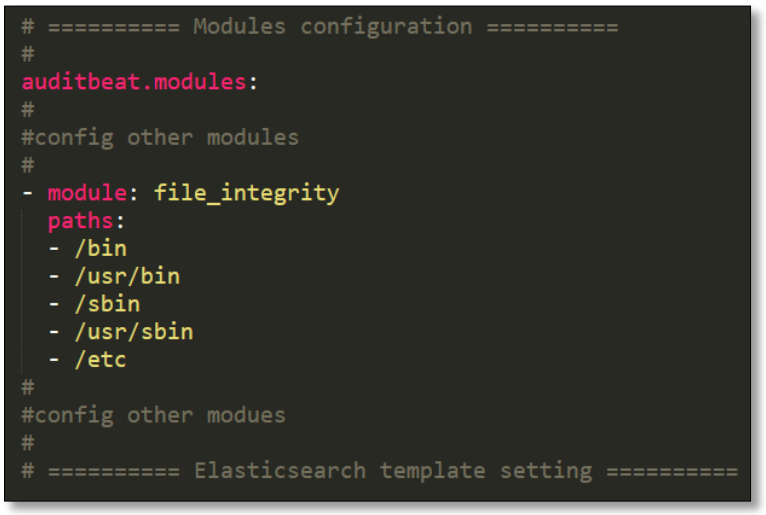


Figure 19 : Fichier de configuration Auditbeat -Module d'intégrité.

Pour en savoir plus sur le processus d'installation d'Auditbeat et les commandes utilisées, consulter **l'Annexe V.**

3.6.2 Filebeat

Filebeat est un collecteur de données de journal léger qui peut être installé en tant qu'agent sur des serveurs. Il surveille les fichiers journaux, collecte les données d'événement et les indexe dans Elasticsearch ou dans Logstash [39]. Pour plus d'informations sur le l'installation et la configuration de Filebeat, voir **l'Annexe VI**.

3.6.3 Metricbeat

Metricbeat est un petit programme qui peut être installé sur un serveur pour collecter des métriques sur le système d'exploitation et d'autres services. Il envoie les statistiques collectées à la sortie spécifiée, comme Logstash ou Elasticsearch. [39]. Pour savoir comment installer Metricbeat et configurer les sorties Elasticsearch et Kibana, consulter **l'Annexe VII.**

3.6.4 Packetbeat

Packetbeat est un examinateur de paquets réseau en temps réel qui fonctionne avec Elasticsearch afin de produire un système d'observation des applications et d vérification des performances. Packetbeat fonctionne en capturant le trafic réseau entre les serveurs d'applications, en décodant les protocoles de la couche application (HTTP, MySQL, Redis, etc.), en corrélant les requêtes avec les réponses et en enregistrant les champs intéressants pour chaque transaction [39]. Pour en savoir plus sur le processus d'installation de Packetbeat et les commandes utilisées, consulter l'**Annexe VIII**.

3.6.5 Winlogbeat

Winlogbeat est Elastic Beat qui envoie les journaux d'événements Windows à Elasticsearch ou Logstash. Winlogbeat utilise les API Windows pour lire un ou plusieurs journaux d'événements, filtrer les événements en fonction de critères configurés par l'utilisateur et envoyer des données d'événement à une sortie configurée (Elasticsearch ou Logstash). Winlogbeat surveille les journaux d'événements sur l'ordinateur, de sorte que s'il y a de nouveaux événements, ils sont récupérés rapidement. La position de

chaque journal est enregistrée sur le disque dur, de sorte que Winlogbeat peut continuer là où il s'est arrêté si l'ordinateur s'éteint ou redémarre [39]. Pour connaitre les étapes d'installations de Winlogbeat sur notre machine Windows 10, voir **Annexe IX.**

3.7 Politique et gestion du cycle de vie des index (ILM)

Index lifecycle management est un gestionnaire des index qui crée, gère et supprime automatiquement les index Elasticsearch.

Pouvoir automatiser la création d'un nouvel index ou optimiser l'index pour économiser les ressources du cluster une fois qu'un index n'est plus en cours d'écriture est très utile. Il existe plusieurs méthodes pour optimiser les index, notamment :

- Forcer la fusion de l'index pour optimiser l'espace utilisé par l'index sur le disque.
- Réduire l'index pour réduire le nombre de fragments.
- Allocation de la partition à du matériel moins performant/optimisé pour le stockage.

Lorsqu'un index approche de la fin de sa durée de vie (atteignant une certaine durée ou taille prédéterminée par l'analyste), nous pouvons prendre un instantané de l'index ou le supprimer complètement en fonction des besoins de l'entreprise.

3.7.1 Définir la politique de gestion du cycle de vie

Une politique de gestion du cycle de vie des index peut être appliquée à n'importe quel nombre d'index. La politique se compose d'un certain nombre de phases définies et indique les mesures à prendre lors du passage d'une phase à la suivante.

Toutes les politiques doivent commencer par une **phase HOT**. Les autres phases sont facultatives. Le passage d'une phase à la suivante est défini par le temps écoulé depuis la création de l'indice ou par l'atteinte d'une taille maximale. Nous avons défini les différentes phases dans notre système comme suit :

a. Hot Phase

Les index actifs reçoivent activement des données dans l'index et traitent les requêtes, à cette phase, les index sont activement mis à jour et sont utilisés pour l'ingestion de données

Le passage de cette phase à la suivante est défini par les politiques suivantes :

- Déclenche le roulement (rollover) après 15 jours écoulés depuis la création de l'index, ou lorsque la taille de l'index atteint 1 Go.
- Un nouvel index sera créé lorsque l'index existant satisfait une ou plusieurs des politiques de roulement.

b. Cold Phase

L'index de cette phase n'est plus mis à jour et reçoit rarement des requêtes. Les informations doivent toujours être consultables, mais ce n'est pas grave si ces requêtes sont plus lentes. Le passage de cette phase à la suivante est défini par une politique qui se déclenche au bout de 15 jours depuis le roulement de la phase précédente.

c. Delete Phase

L'index n'est plus nécessaire et peut être supprimé en toute sécurité.

Les figures suivantes montrent comment nous avons défini les politiques de cycle de vie des index pour chaque phase

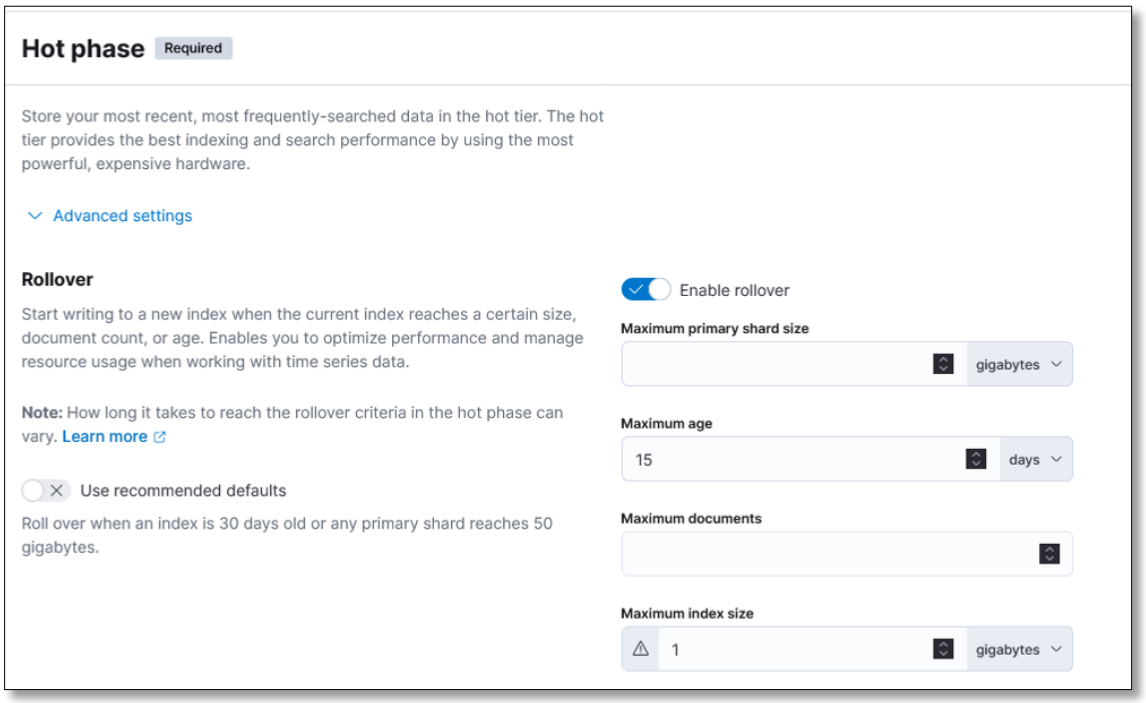


Figure 20 : politique de cycle de vie d'index -Hot phase.

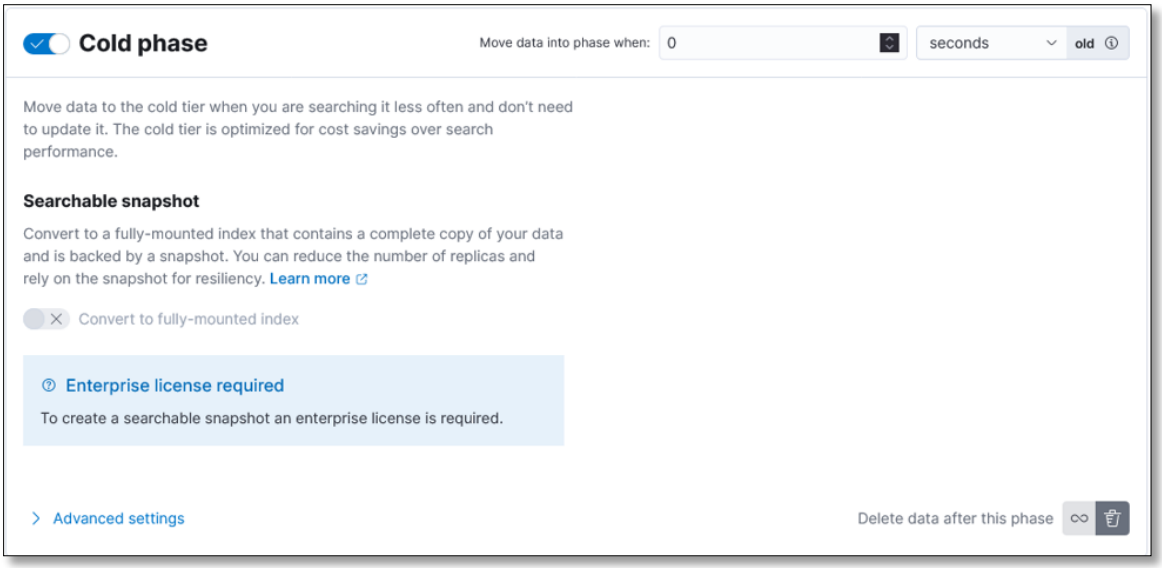


Figure 21 : politique de cycle de vie d'index -Cold phase

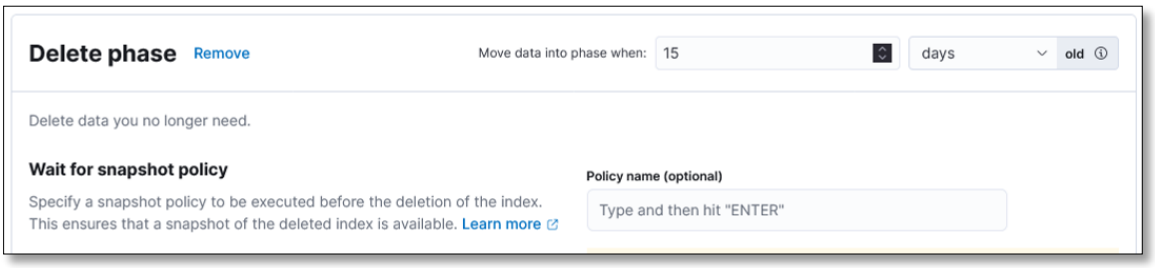


Figure 22 : politique de cycle de vie d'index -Delete phase.

3.8 Configurer la sécurité minimale pour Elasticsearch

Les fonctions de sécurité d'Elasticsearch sont désactivées par défaut. L'activation des fonctionnalités de sécurité Elasticsearch active l'authentification de base afin d'avoir la possibilité d'exécuter un cluster local avec une authentification par nom d'utilisateur et mot de passe. La configuration se fait sur le fichier **/etc/elasticsearch/elasticsearch.yml**.

3.8.1 Créer des mots de passe pour les utilisateurs intégrés

Il est nécessaire de configurer un nom d'utilisateur et mot de passe pour les utilisateurs intégrés pour communiquer avec le cluster, toutes les demandes qui n'incluent pas un nom d'utilisateur et un mot de passe sont rejetées.

Apres l'activation des fonctions de sécurité d'Elasticsearch, Nous définissons les mots de passe des utilisateurs intégrés en exécutant l'utilitaire **elasticsearch-setup-passwords auto.**

3.8.2 Configurer Kibana pour s'authentifier à Elasticsearch

Lorsque les fonctionnalités de sécurité sont activées, Kibana ne peut pas accéder aux données stockées dans Elasticsearch, nous devons accorder l'accès à Kibana en configurant des rôles pour les utilisateurs de Kibana afin de contrôler les données auxquelles ces utilisateurs peuvent accéder.

Dans le fichier de configuration de Kibana, nous ajoutons le nom d'un utilisateur ayant accès à Elasticsearch, nous créons le **keystore Kibana** et ajoutons le mot de passe de l'utilisateur **kibana_system**

```
elasticsearch.username: "kibana_system"
```

Figure 23 : Fichier de configuration Kibana -nom d'utilisateur.

```
./bin/kibana-keystore add elasticsearch.password
```

Figure 24 : Ajout d'un mot de passe au keystore de kibana.

3.9 Configurer la sécurité de base pour Elastic Stack

Après avoir ajouté la protection par mot de passe dans **la section précédente**, nous devrons configurer Transport Layer Security (TLS). La couche de transport gère toutes les communications internes entre les nœuds de notre cluster.

La couche de transport s'appuie sur **Mutual TLS** pour le chiffrement et l'authentification des nœuds. L'application correcte de TLS garantit qu'un nœud malveillant ne peut pas rejoindre le cluster et échanger des données avec d'autres nœuds. Bien que la mise en œuvre de l'authentification au niveau de la couche HTTP soit utile pour sécuriser un cluster local, la sécurité de la communication entre les nœuds nécessite TLS.

3.9.1 Générer l'autorité de certification

Dans un cluster sécurisé, les nœuds Elasticsearch utilisent des certificats pour s'identifier lorsqu'ils communiquent avec d'autres nœuds.

Le cluster doit valider l'authenticité de ces certificats. L'approche recommandée consiste à faire confiance à une autorité de certification (CA) spécifique. Lorsque des

nœuds sont ajoutés au cluster, ils doivent utiliser un certificat signé par la même autorité de certification.

```
./bin/elasticsearch-certutil ca
```

Figure 25 : génération d'autorité de certification.

3.9.2 Chiffrer les communications inter-nœuds avec TLS

La couche réseau de transport est utilisée pour la communication interne entre les nœuds d'un cluster. Lorsque les fonctions de sécurité sont activées, nous devons utiliser TLS (Transport Layer Security) pour assurer que la communication entre les nœuds est chiffrée.

Afin de joindre le même cluster, tous les nœuds doivent partager le même nom de cluster, Dans le fichier **/etc/elasticsearch/elasticsearch.yml**, nous ajoutons les paramètres suivants pour activer la communication entre les nœuds et donner accès au certificat du nœud.

```
xpack.security.transport.ssl.enabled: true
xpack.security.transport.ssl.verification_mode: certificate 1
xpack.security.transport.ssl.client_authentication: required
xpack.security.transport.ssl.keystore.path: elastic-certificates.p12
xpack.security.transport.ssl.truststore.path: elastic-certificates.p12
```

Figure 26 : Fichier de configuration Elasticsearch -configuration ssl.

3.10 Sécurisation du trafic https

Après avoir complété la configuration basique de sécurité et afin de déverrouiller plus de fonctionnalités d'Elasticsearch, il est nécessaire d'activer TLS sur la couche HTTP. Cette couche de sécurité supplémentaire garantit que toutes les communications vers et depuis notre cluster sont sécurisées.

3.10.1 Chiffrer les communications du client HTTP pour Elasticsearch

Nous exécutons l'outil de certificat HTTP Elasticsearch pour générer une demande de signature de certificat (CSR), cette commande génère un fichier compressé qui contient des certificats et des clés à utiliser avec Elasticsearch et Kibana.

```
./bin/elasticsearch-certutil http
```

Figure 27 : génération de demande de signature de certificat (CSR).

```
/elasticsearch
|_ README.txt
|_ http.p12
|_ sample-elasticsearch.yml
```

```
/kibana
|_ README.txt
|_ elasticsearch-ca.pem
|_ sample-kibana.yml
```

Figure 28 : Contenu du fichier elasticsearch-ssl-http.zip.

3.11 Crypter le trafic entre le navigateur et Kibana

Pour chiffrer le trafic entre le navigateur et Kibana, il est nécessaire de générer un certificat de serveur et une clé privée pour Kibana. Kibana utilise ce certificat de serveur et la clé privée correspondante lors de la réception de connexions à partir du navigateurs Web.

```
/kibana-server
|_ kibana-server.crt
|_ kibana-server.key
```

Figure 29 : Contenu du certificat de serveur.

3.12 Configurer la sécurité Beats

Les Beats sont des expéditeurs de données open source que nous installons en tant qu'agents sur nos serveurs pour envoyer des données opérationnelles à Elasticsearch, pour plus de détails, voir la **section 4.1.3 du chapitre 2.**

Nous avons choisi **metricbeat** comme exemple de configuration de sécurité des beats, ces étapes sont correctes pour chaque beat supplémentaire pour lequel nous voulons configurer la sécurité.

Il est nécessaire d'ajouter les champs suivants dans le fichier de configuration **metricbeat.yml** pour configurer la connexion avec Elasticsearch et Kibana.

```
output.elasticsearch:
 hosts: ["192.168.1.7:9200"]
 protocol: "https"
 username: "metricbeat"
 password: "m2ssi2022"
 ssl:
   certificate_authorities: ["elasticsearch-ca.pem"]
   verification_mode: "certificate"
```

Figure 30 : Fichier de configuration Metricbeat -Elasticsearch Output ssl configuration.

```
setup.kibana
 host: "https://192.168.1.7:5601"
 ssl.enabled: true
 username: "metricbeat"
 password: m2ssi2022
```

Figure 31 : Fichier de configuration Metricbeat -Kibana Output ssl configuration.

Configurer la sécurité des Beats ne suffira pas à envoyer les journaux à Elasticsearch en toute sécurité, nous devons configurer la sécurité des modules afin que n'importe quel module de n'importe quel Beat puisse se connecter à Elasticsearch en toute sécurité.

Tout d'abord, dans le fichier de configuration **/modules.d/elasticsearch-xpack.yml**, il faut modifier l'URL Elasticsearch de **http:// à https**://, afin que les données échangées entre Elasticsearch et elasticsearch-xpack soient sécurisées via Transport Layer Security (TLS), puis on configure l'authentification à Elasticsearch et on indiquer le chemin d'autorités de certification.

```
- module: elasticsearch
  xpack.enabled: true
  period: 10s
  hosts: ["https://192.168.1.7:9200"]
  username: "admin"
  password: "m2ssi2022"
  ssl:
    enabled: true
    certificate_authorities: ["elasticsearch-ca.pem"]
    verification_mode: "certificate"
```

Figure 32 : Fichier de configuration Elasticsearch-xpack -activer la sécurité SSL

3.13 Déploiement de Fleet Server

Une fois Elasticsearch et Kibana sécurisés, nous pouvons commencer le processus de déploiement de Fleet Server. Tout d'abord, nous devons installer Elastic Agent et l'inscrire dans une politique d'agent contenant l'intégration Fleet Server, en accédant à la page de téléchargement d'Elastic Agent (https://www.elastic.co/downloads/elastic-agent) et suivons les étapes d'installation.

3.13.1 Ajouter un Fleet Server

Nous devons nous connecter à l'interface Kibana et aller dans **Management > Fleet**, Une fois la page chargée, nous suivons les étapes d'intégration de Fleet Server. Après une installation réussie, nous pouvons voir **Fleet Server Agent** dans l'onglet **Agents de Fleet.**

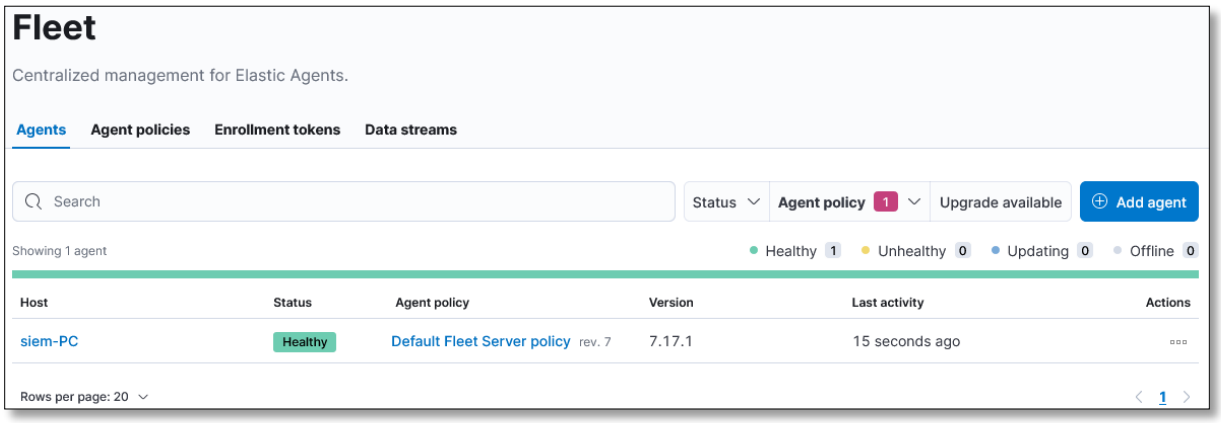


Figure 33 : Agent de serveur Fleet.

3.14 Créer une politique d'agent

Une fois le serveur Fleet bien configuré, nous passons maintenant à la création des politiques d'agent Elastique pour envoyer les données à Elasticsearch. La figure suivante montre la politique d'agent Elastique installé sur la machine Windows pour récolter les évènements de système et Windows.

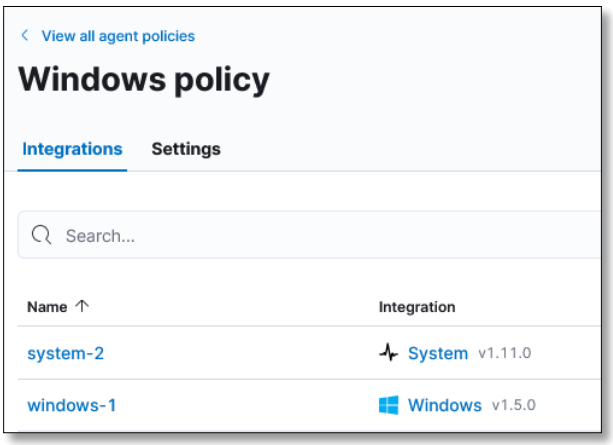


Figure 34 : Intégrations de la politique "Windows policy".

3.15 Suricata

Développé par L'Open Information Security Foundation (OSIF), le moteur de détection open source Suricata peut être utilisé comme IDS et IPS, Le système utilise un ensemble de règles et un langage de signature pour détecter et prévenir les menaces. Suricata peut fonctionner sous Windows, Mac, Unix et Linux [47].

Une fois installé avec succès, nous définissons les règles sur lesquelles Suricata s'appuie pour détecter les activités suspectes et les alertes lorsqu'une telle activité est découverte, dans l'exemple ci-dessous, nous avons créé une règle de détection DOS (Denial Of Service) pour alerter l'attaque par inondation SYN (Synchronize), l'alerte sera générée s'il y a plus de 5000 paquets TCP (Transmission Control Protocol) entrants dans les 5 prochaines secondes.

```
alert tcp $EXTERNAL_NET any -> $HOME_NET any (msg:"LOCAL DOS SYN packet flood inbound, Potential DOS";
flow:to_server; flags: S,12; threshold: type both, track by_dst, count 5000, seconds 5;
classtype:misc-activity; sid:5;)
```

Figure 35 : Règle Suricata.

Les paramètres communs de la règle sont les actions :

- **alert** qui alerte lorsque les conditions de la règle sont satisfaites.
- **tcp** le protocole sur lequel se concentre la règle.
- **$HOME_NET** & **$EXTERNAL_NET** est un ensemble d'adresses IP spécifiées dans le fichier **/etc/suricata/suricata.yaml**.
- **msg** le Message à afficher lors de l'alerte.
- **flags** contient S pour le paquet SYN.
- **threshold** pour définir un seuil minimum pour une règle avant qu'elle ne génère des alertes.

4 Conclusion

Au cours de ce chapitre, nous avons présenté le déploiement de notre solution SIEM (Security Information and Event Management), en commençant par la présentation de l'environnement de travail sur lequel nous avons implémenté notre projet suivi d'une présentation détaillée des différents outils et technologies utilisés dans la mise en place de ce projet

Chapitre 5
Tests et évaluations

1 Introduction

Après avoir implémenté notre solution, nous passons à la présentation de plusieurs cas de test dans ce chapitre pour nous assurer que notre système est efficace et fonctionne comme prévu. La dernière étape consiste alors à étudier les différentes données et graphes produits par Kibana et à évaluer l'efficacité de la solution à travers différents scénarios et tests.

2 Tableaux de bord et visualisations

La meilleure façon de comprendre nos données est de les visualiser avec les tableaux de bord, nous pouvons transformer nos données d'un ou plusieurs modèles d'index en une collection de panneaux qui apportent de la clarté à nos données et nous permettent de concentrer uniquement sur les données qui sont importantes.

Kibana propose diverses visualisations telles que : des cartes thermiques, des graphiques à barres, des graphiques en aires et de nombreux autres types de visualisation. Nous avons choisi le tableau de bord Suricata comme exemple de tableaux de bord créés, pour voir plus de Dashboards (tableaux de bord), consulter l'**Annexe X**.

Le tableau de bord suricata nous donne une vue détaillée des attaques et des menaces présentes sur le réseau en temps réel, ainsi qu'une carte des emplacements IP des attaques (source et récepteur).

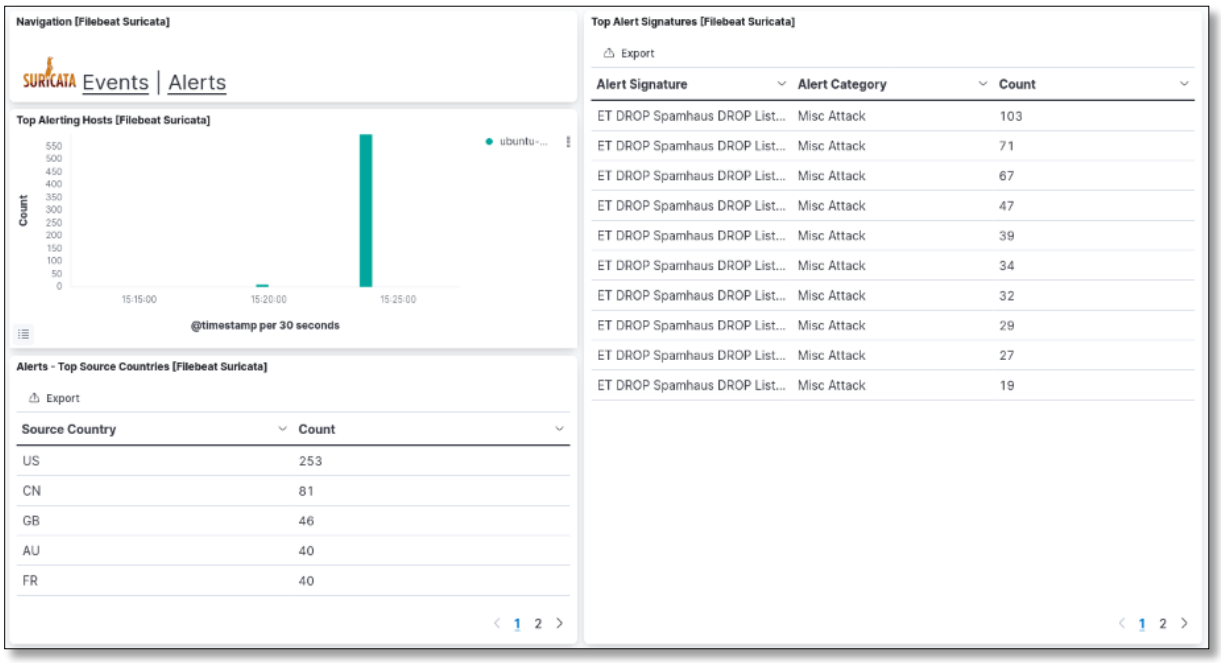


Figure 36 : Dashboard_Suricata Part 1.

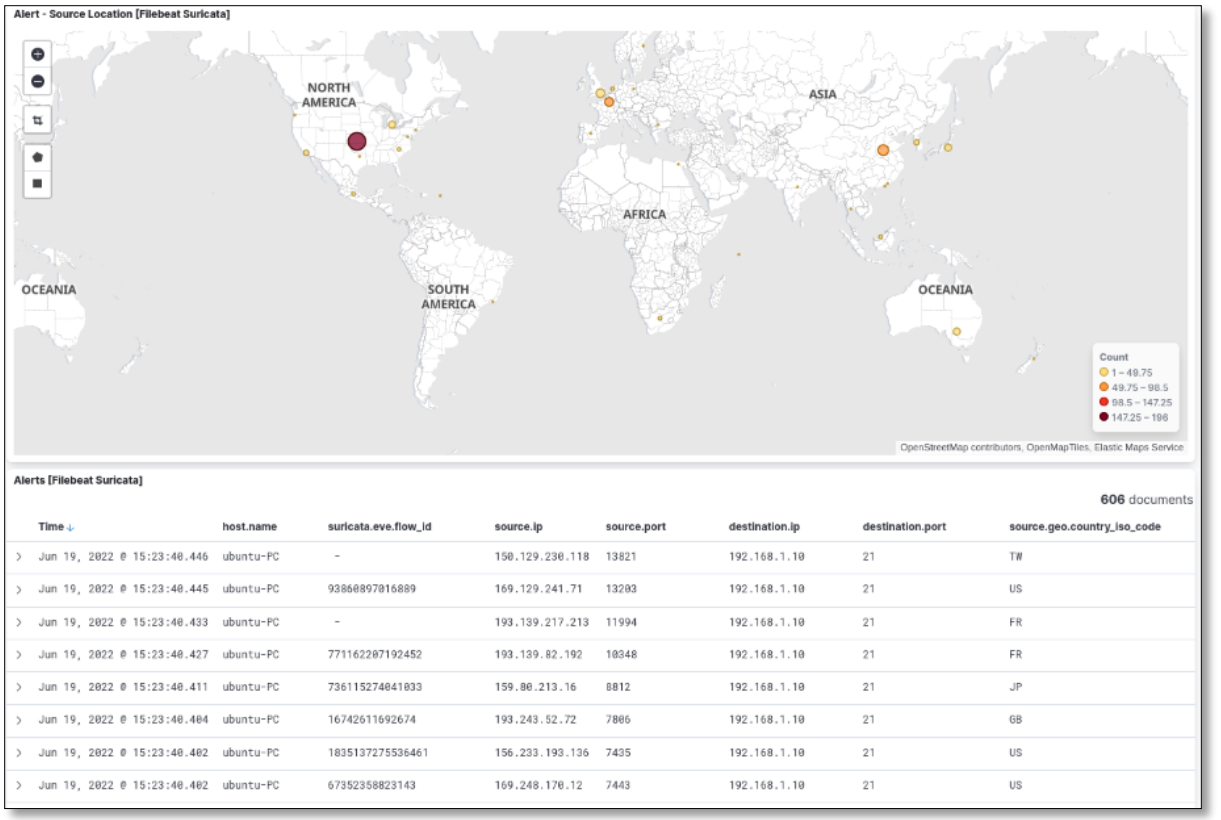
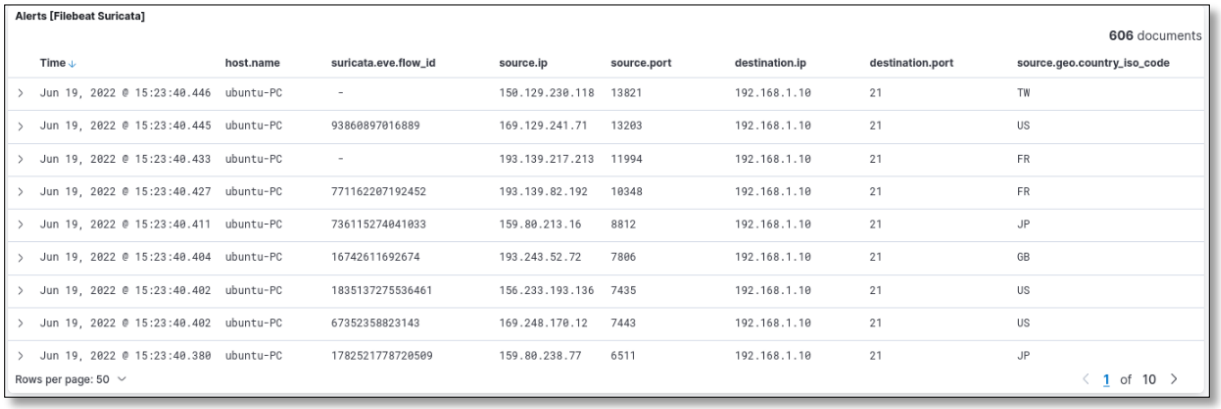


Figure 37 : Dashboard_Suricata Part 2.

Le tableau présenté dans la figure suivante nous montre des détails sur les événements Filebeat collectés à partir de Suricata, dans ces détails on peut retrouver : l'adresse IP source de l'attaquant, le port source, la destination qui est notre machine Ubuntu, le port destinataire qui nous permet de voir sur quel protocole a été lancer l'attaque. Ces informations peuvent nous permettre de bloquer ces adresses ou d'arrêter le service du protocole servant à faire l'attaque.



Alerts [Filebeat Suricata]

606 documents

Time	host.name	suricata.eve.flow_id	source.ip	source.port	destination.ip	destination.port	source.geo.country_iso_code
Jun 19, 2022 @ 15:23:40.446	ubuntu-PC	-	150.129.230.118	13821	192.168.1.10	21	TW
Jun 19, 2022 @ 15:23:40.445	ubuntu-PC	93860897016889	169.129.241.71	13203	192.168.1.10	21	US
Jun 19, 2022 @ 15:23:40.433	ubuntu-PC	-	193.139.217.213	11994	192.168.1.10	21	FR
Jun 19, 2022 @ 15:23:40.427	ubuntu-PC	771162207192452	193.139.82.192	10348	192.168.1.10	21	FR
Jun 19, 2022 @ 15:23:40.411	ubuntu-PC	736115274041033	159.80.213.16	8812	192.168.1.10	21	JP
Jun 19, 2022 @ 15:23:40.404	ubuntu-PC	16742611692674	193.243.52.72	7806	192.168.1.10	21	GB
Jun 19, 2022 @ 15:23:40.402	ubuntu-PC	1835137275536461	156.233.193.136	7435	192.168.1.10	21	US
Jun 19, 2022 @ 15:23:40.402	ubuntu-PC	67352358823143	169.248.170.12	7443	192.168.1.10	21	US
Jun 19, 2022 @ 15:23:40.380	ubuntu-PC	1782521778720509	159.80.238.77	6511	192.168.1.10	21	JP

Rows per page: 50 1 of 10

Figure 38 : Dashboard_Suricata Filebeat Suricata logs

Le Dashboard va même loin en nous montrant sur un maps la localisation de l'IP de l'attaquant.

Figure 39 : Dashboard_Suricata Sources d'attaques.

Le nombre des attaques par pays ainsi que le **top alerting host** (dans notre cas c'est Ubuntu machine) sont représentées dans la figure suivante :

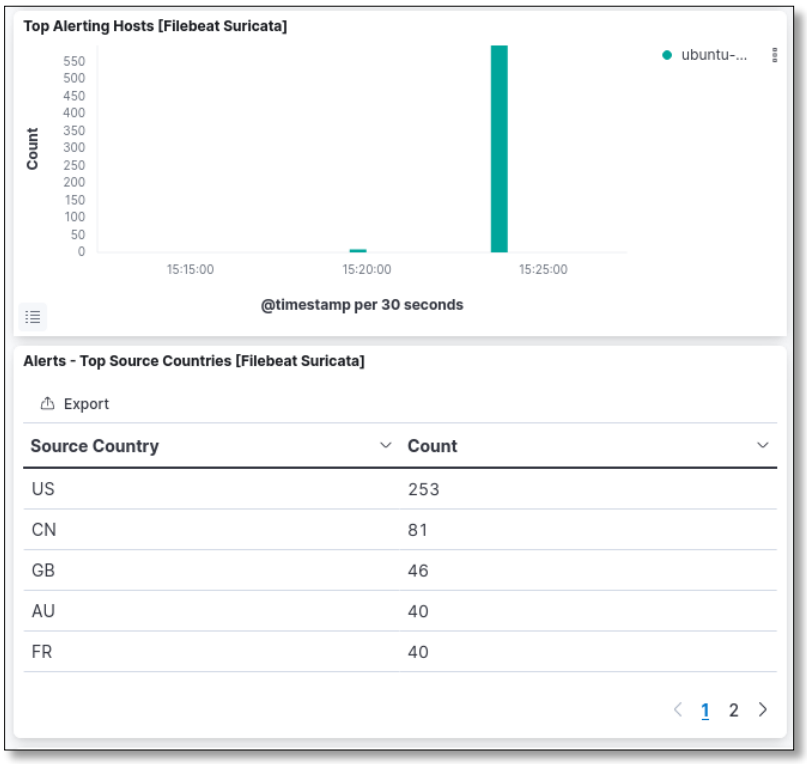


Figure 40 : Dashboard_Suricata Top alerts (Hosts / Countries).

Le tableau de bord Suricata fournit également plus de détails sur les attaques qui se sont produites en présentant la signature d'alerte et sa catégorie et le nombre d'occurrences de chaque attaque.

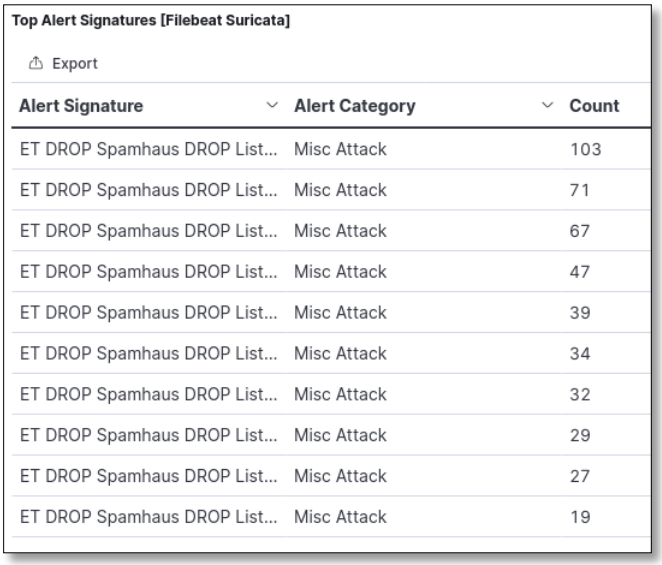


Figure 41 : Dashboard_Suricata Top Alerts Signatures.

3 Cas d'utilisation

Afin de tester notre solution, cette section sera consacrée à une série de scénarios d'attaques et de menaces informatiques pour observer le comportement de notre SIEM (Security Information and Event Management).

3.1 Intégrité des fichiers

La surveillance de l'intégrité des fichiers aide les organisations à détecter les modifications inappropriées apportées aux fichiers critiques de leurs systèmes, réduisant ainsi le risque de vol ou de compromission de données pouvant entraîner une perte de temps et d'argent, une perte de revenus, une atteinte à la réputation et des sanctions juridiques et de conformité.

3.1.1 Intégrité des fichiers « FILE_CREATED »

Une fois un nouveau répertoire ou fichier créé dans les chemins supervisés par le module Auditbeat **File_Integrity** (voir **section 3.6.1 du chapitre 4**) une alerte se déclenche.

- Pour faire cela, nous naviguons vers **Security > Detect > Rules > Create new rule**.
- Nous choisissons **custom query rule** pour créer une règle basée sur une requête KQL (Kibana Query Language).
- Nous définissons les indices **auditbeat-*** dans le champ **index patterns**
- Nous entrons notre requête dans le champ **custom query.**
- Dans la page **About rule**, nous remplissons les champs suivants :
  - **Name** : le nom de la règle.
  - **Description** : une description de ce que fait la règle.
  - **Default severity** : le niveau de gravité des alertes créées par la règle :
    - **Low** : Une alerte intéressante qui n'est généralement pas considérée comme un événement de sécurité. Parfois, un ensemble d'alertes de bas niveau peut signaler une action inhabituelle.

        - **Medium** : Alertes nécessitant une enquête.
        - **High** : Alertes nécessitant une enquête immédiate.
        - **Critical** : Alertes qui indiquent une forte probabilité d'incident de sécurité.
    - **Default risk score** : une valeur numérique comprise entre 0 et 100 qui correspond au niveau de gravité. Les directives générales sont :
        - **0 - 21** représente une gravité faible.
        - **22 - 47** représente une sévérité moyenne.
        - **48 - 73** représente une sévérité élevée.
        - **74 - 100** représente la gravité critique.
- Dans la page **Schedule rule**, nous sélectionnons la fréquence d'exécution de la règle (temps de vérification de l'alerte)
- Nous créons la règle en cliquant sur **Create & activate rule**

Cela recherchera les indices **auditbeat-*** pour les exécutions **event.action** avec l'argument **"created",** qui est utilisé pour créer un fichier.

Note : Nous notons que les étapes mentionnées ci-dessus sont les mêmes pour chaque création d'alerte.

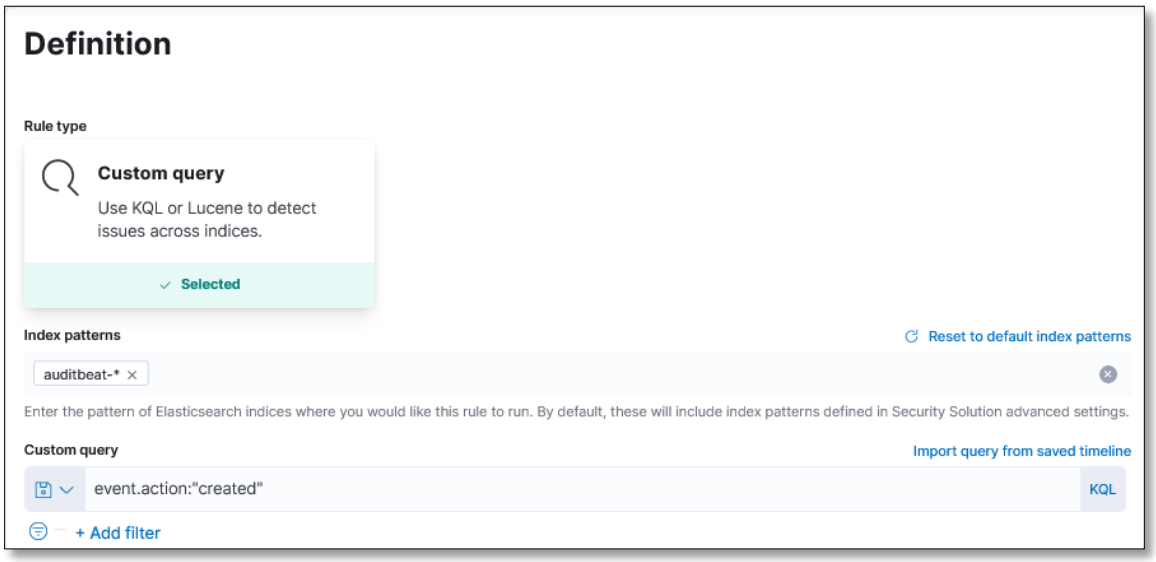


Figure 42 : Intégrité des fichiers « FILE_CREATED » Définition page.

Figure 43 : Intégrité des fichiers « FILE_CREATED » About page.

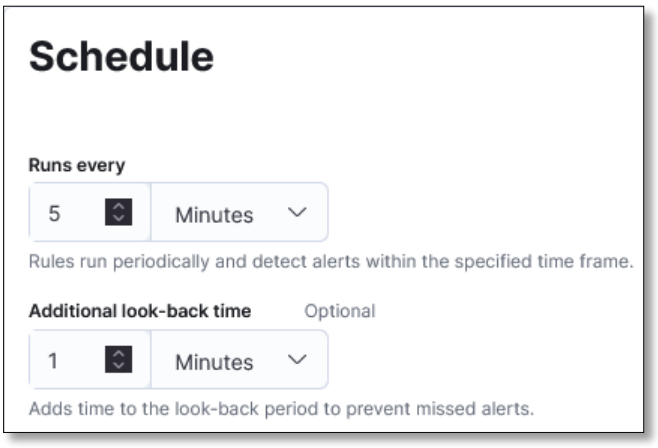


Figure 44 : Intégrité des fichiers « FILE_CREATED » Schedule page.

3.1.2 Intégrité des fichiers « FILE_UPDATED »

Dès qu'un répertoire ou un fichier est mis à jour dans les chemins supervisés par le module Auditbeat **File_Integrity** (voir **section 3.6.1 du chapitre 4**) une alerte se déclenche.

3.1.3 Intégrité des fichiers « FILE_DELETED »

Lors de la suppression d'un répertoire ou d'un fichier dans les chemins supervisés par le module Auditbeat **File_Integrity** (voir **section 3.6.1 du chapitre 4**) une alerte se déclenche.

3.1.4 Test et Evaluation

Les privilèges d'administrateur peuvent être attribués à des utilisateurs ou à des groupes via le fichier **/etc/sudoers**. L'addition des lignes avec les bonnes règles peut donner à l'utilisateur le droit de modifier les autorisations dans notre système. Une fois qu'un attaquant a accès à notre système, il peut modifier ce fichier sans que l'administrateur système ne s'en aperçoive.

La figure suivante montre un scénario d'attaque où l'attaquant a modifié le fichier **sudoers** et ajouté une règle (**sudo ALL=NOPASSWD: ALL**) permettant aux commandes

sudo de ne pas avoir besoin d'un mot de passe pour l'exécuter, et cela lui accordera les privilèges **root** qui ont l'autorisation complète de lire, écrire et exécuter tout fichier.

```
# User privilege specification
root    ALL=(ALL:ALL) ALL

# Members of the admin group may gain root privileges
%admin ALL=(ALL) ALL

# Allow members of group sudo to execute any command
#%sudo  ALL=(ALL:ALL) ALL
%sudo ALL=NOPASSWD: ALL

# See sudoers(5) for more information on "#include" directives:
```

Figure 45 : modification de fichier sudoers.

On peut voir qu'après **5 minutes** de modification de fichier (temps périodique de vérification de l'alerte) une alerte se déclenche indiquant qu'un fichier a été mis à jour (Figure 46). Nous pouvons avoir le même résultat si nous consultons le tableau de bord FILE_INTEGRITY comme le montre les Figures 47,48.

Figure 46 : Alerte Kibana Intégrité des fichiers.

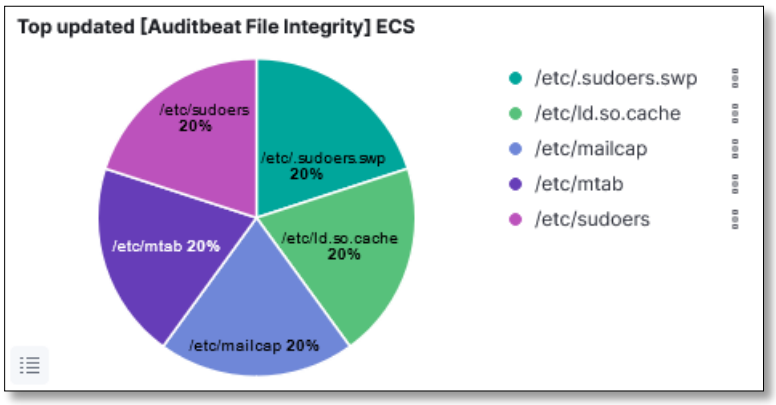


Figure 47 : visualisation graphique circulaire Dashboard File Integrity.

File Integrity Events [Auditbeat File Integrity] ECS

Time	file.path	event.action
Jun 24, 2022 @ 17:05:09.307	/etc/.sudoers.swp	updated, attributes_modified
Jun 24, 2022 @ 17:05:05.211	/etc/.sudoers.swp	created
Jun 24, 2022 @ 17:03:29.883	/etc/sudoers	updated, attributes_modified

Figure 48 : visualisation graphique tableau Dashboard File Integrity.

3.2 Attaque par déni de service (DOS)

CISA ou l'Agence de cybersécurité et de sécurité des infrastructures est une agence fédérale des États-Unis, son activité principale consiste à défendre l'infrastructure d'Internet et à améliorer sa résilience et sa sécurité en identifiant et en évaluant les menaces à l'infrastructure et en fournissant des outils de cybersécurité, une analyse des menaces et une réponse aux incidents sur tous les sites Web **.gov**. [48].

Selon CISA « *Une attaque par déni de service (DOS) se produit lorsque des utilisateurs légitimes ne peuvent pas accéder aux systèmes d'information, aux appareils ou à d'autres ressources réseau en raison des actions d'un acteur de la cybermenace malveillante. Les services concernés peuvent inclure les e-mails, les sites Web, les comptes en ligne (par exemple, les services bancaires) ou d'autres services qui dépendent de l'ordinateur ou du réseau concerné. Une condition de déni de service est accomplie en inondant l'hôte ou le réseau ciblé de trafic jusqu'à ce que la cible ne puisse pas répondre ou tombe simplement en panne,*

empêchant l'accès aux utilisateurs légitimes. Les attaques DOS peuvent coûter du temps et de l'argent à une organisation alors que ses ressources et ses services sont inaccessibles. » [49].

3.2.1 Test et Evaluation

Maintenant, l'attaque DOS (Denial Of Service) est menée à partir de la machine de l'attaquant à l'aide de l'utilitaire réseau hping3 pour générer et inonder le paquet TCP SYN (Synchronize) vers l'adresse IP cible.

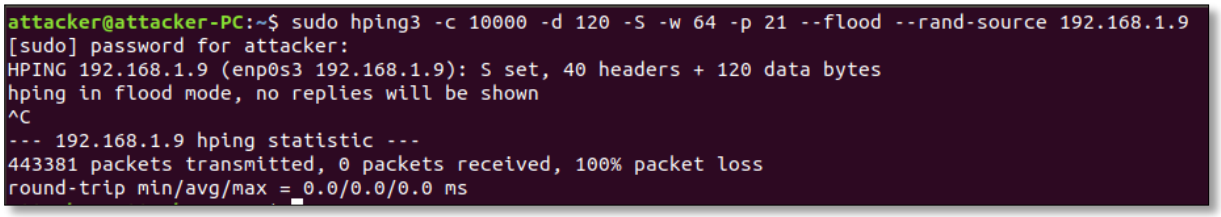


Figure 49 : Commande d'attaque DOS.

- Les paramètres communs de la commande sont les actions :
  - **sudo hping3** donne les privilèges nécessaires pour exécuter hping3.
  - **-c** Nombre de paquets à envoyer.
  - **-d** Pour définir la taille du paquet
  - **-S** Spécifie les paquets SYN (connexion de paquet TCP légitime).
  - **-w** Définir la taille de la fenêtre TCP.
  - **-p** Spécifier le port ciblé.
  - **–flood :** les réponses seront ignorées et les paquets seront envoyés le plus rapidement possible.
  - **--rand-source** Randomiser l'adresse IP source.
  - **192.168.1.9** Adresse IP de destination.

Nous pouvons arrêter l'attaque après quelques minutes de fonctionnement et noter qu'une nouvelle alerte est présente informant qu'il y a une possible attaque DOS.

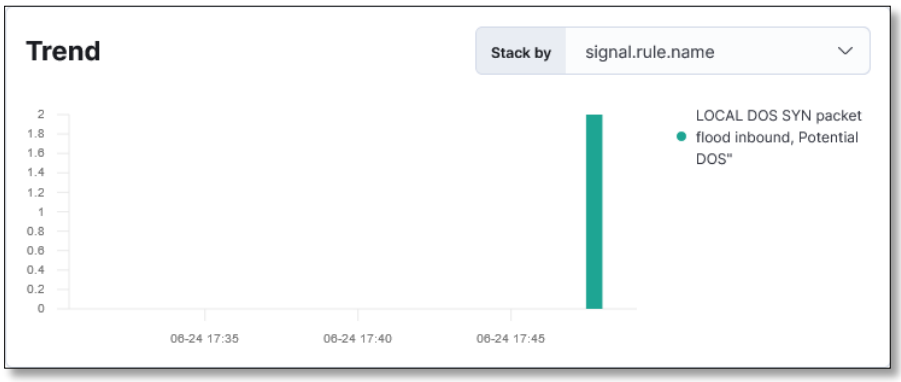


Figure 50 : Alerte Kibana attaque par déni de service.

Nous pouvons voir le même résultat dans le tableau de bord prédéfini de Suricata, le tableau de bord va même loin en nous indiquant sur une carte la localisation de l'IP de l'attaquant.

Figure 51 : visualisation graphique carte de localisation Dashboard Suricata.

3.3 Scan de port réseau avec Nmap

Selon K. Scarfone, T. Grance, et K. Masone, une analyse de port consiste à « *utiliser un programme pour déterminer à distance quels ports d'un système sont ouverts* » [50].

Nmap (Network Mapper) est un programme gratuit qui détecte le système d'exploitation utilisé par un périphérique et les services exécutés sur les ports ouverts du réseau en envoyant divers paquets aux ports hôtes ciblés du réseau [51].

Notre solution SIEM détecte les scans Nmap grâce au moteur de détection open source Suricata, de manière plus spécifique à partir des règles Suricata créées précédemment (voir **section 3.15 du chapitre 4**).

3.3.1 Test et Evaluation

Une fois que nous avons fini d'établir notre règle de détection, nous passons à la machine de l'attaquant et effectuons une analyse Nmap pour tester notre règle. Nous exécutons la commande suivante pour analyser les ports ouverts de la machine victime.

Figure 52 : Commande de scan nmap.

Cette commande lance une analyse des ports **TCP** (Transmission Control Protocol) des 1000 ports les plus populaires répertoriés dans **nmap-services**.

Une fois notre analyse nmap terminée, nous recevons une alerte de l'interface Kibana Security nous informant qu'une analyse Nmap a été détectée.

Figure 53 : Alerte Kibana Scan de port réseau avec Nmap.

3.4 Attaque de logiciel malveillant (malware)

Un logiciel malveillant est un logiciel informatique conçu pour perturber le système/réseau. Il peut facilement accéder aux informations confidentielles des entreprises et des systèmes personnels [52].

Les logiciels malveillants constituent une menace sérieuse pour toutes les entreprises. Cela pose plusieurs problèmes de sécurité aux entreprises, tels que les interruptions et la désactivation des services, la rupture de l'infrastructure de sécurité d'un réseau/système, le contrôle de toutes les applications exécutées sur les appareils des organisations, l'accès aux informations sensibles, etc [52]. Pour recevoir des logs de ce type (détection de malware), nous avons besoin de l'aide de **Winlogbeat** que nous avons préalablement installée.

3.4.1 Test et Evaluation

Lors de nos tests, nous avons téléchargé le fichier de test de virus **EICAR** compressé dans un fichier ZIP qui est un programme DOS (Denial Of Service) créé par l'Institut européen de recherche sur les antivirus informatiques, qui affiche uniquement le message "EICAR-STANDARD-ANTIVIRUS-TEST-FILE" à l'écran, puis se termine lui-même.

Une fois le virus extrait, Windows Defender détecte le fichier malveillant et instantanément nous recevons une alerte nous informant qu'un Malware a été détecté, Cela nous confirme que la règle de détection fonctionne parfaitement.

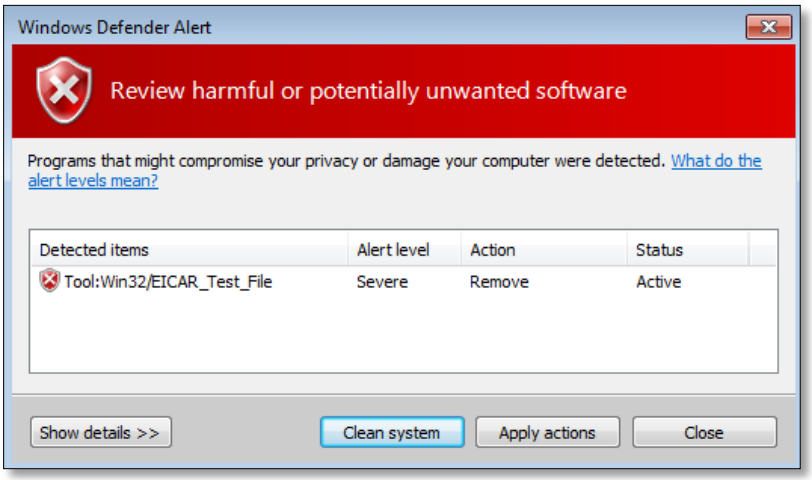


Figure 54 : Windows Defende alert sur la machine windows 7.

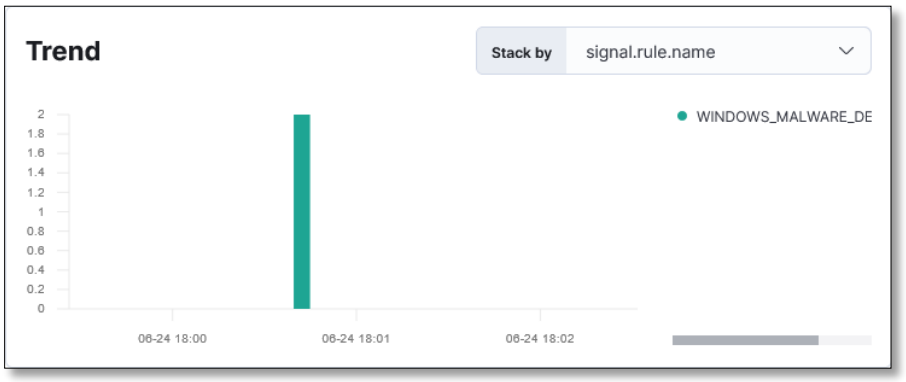


Figure 55 : Alerte Kibana Attaque de logiciel malveillant.

3.5 Surveillance des tentatives de connexion échouées

L'une des choses que les administrateurs font souvent est de vérifier les multiples tentatives de connexion infructueuses. Leur surveillance nous permet d'évaluer si des tentatives illégales ou indésirables sur notre réseau ont eu lieu.

3.5.1 Test et Evaluation

Nous avons testé cette règle sur le compte local de notre machine Windows 10, après avoir tapé 3 fois de suite de faux mots de passe, trois événements Windows d'échec d'authentification sont envoyés à Kibana et une alerte est immédiatement générée notifiant les échecs successifs des tentatives de connexion.

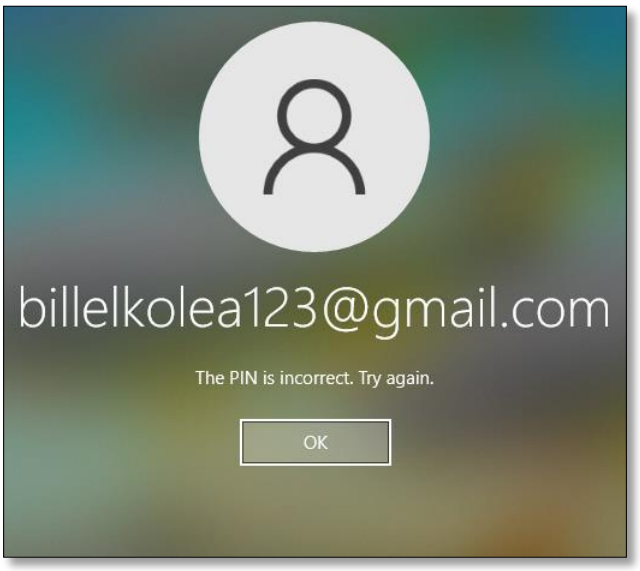


Figure 56 : Echec de tentative de connexion windows.

Figure 57 : Alerte Kibana Tentative de connexion échouées.

4 Conclusion

Dans ce chapitre, nous nous sommes familiarisés avec la pile Elastic, puis l'expérimenter pour mettre en évidence ses capacités en créant des visualisations et des tableaux de bord qui nous permettent de superviser nos systèmes et d'obtenir une vue globale et des informations plus détaillées sur la sécurité et la santé de notre système. Nous avons pu effectuer une série d'attaques sur notre solution et nous sommes capables de détecter ces attaques grâce aux règles de corrélations qui nous permettent de réagir pour les prévenir ou les éviter.

Conclusion Générale

Ce projet de fin d'études porte sur la mise en place d'une solution de gestion des informations et des événements de sécurité au sein de la société **NAFTAL** afin de déceler les failles de sécurité du système informatique et des systèmes de communications de l'entreprise.

Au terme de ce projet, nous avons pu atteindre les objectifs que nous nous étions fixés nous avons conçu un système qui permet de rassembler, analyser, stocker et indexer les fichiers journaux de divers appareils, services, serveurs et systèmes, corréler les événements de plusieurs ressources et surveiller l'accès aux ressources critiques en générant des alertes dès qu'une attaque est effectuée ou tentée, ces alertes seront ensuite traitées par l'équipe SOC.

Les différents tests effectués sur le système ont tous bien fonctionné en détectant les tentatives d'attaques testées.

Ce travail a été développé à l'aide d'Elastic Stack qui effectue une variété de tâches avec un haut niveau de réussite et offre la possibilité d'utiliser de puissants tableaux de bord pour analyser tous les types de données.

En termes de perspectives, nous notons que notre solution pourrait être améliorée et optimisée par :

- Intégrer l'apprentissage machine Elasticsearch pour disposer d'un système capable d'automatiser le processus de dépannage.
- Un système d'envoi d'alertes par email en cas de détection d'attaque.
- Optimisation des règles de corrélation pour réduire le taux de faux positifs.
- Enrichir le système en intégrant de nouveaux modules de collecte de données pour mieux corréler les événements et assurer la sécurité du système d'information.
- Créer de nouvelles règles de détection d'incidents et se tenir au courant des nouvelles cyberattaques pour ne pas mettre le système d'information en danger.

Enfin, ce stage a été une très bonne expérience pour nous en tant qu'étudiants en Sécurité des Systèmes d'Information car il nous a permis d'appliquer nos connaissances théoriques et d'acquérir de nouvelles connaissances et d'améliorer notre capacité à communiquer, collaborer et nous adapter avec le milieu professionnel en réalisant un projet dans une grande entreprise nationale.

Références

[1] BORIS CIZELJ, « Information and cyber security », *meer*, sept. 05, 2021. https://www.meer.com/en/66841-information-and-cyber-security

[2] R. K. Pandey et M. Misra, « Cyber security threats — Smart grid infrastructure », in *2016 National Power Systems Conference (NPSC)*, déc. 2016, p. 1-6. doi: 10.1109/NPSC.2016.7858950.

[3] R. G. Yende, « SUPPORT DE COURS DE SÉCURITÉ INFORMATIQUE ET CRYPTO. », oct. 2018, Consulté le: août 12, 2022. [En ligne]. Available: https://hal.archives-ouvertes.fr/cel-01965300

[4] CNSS, « Committee on National Security Systems », *Cnss.gov*, juin 05, 2016. https://www.cnss.gov/cnss/

[5] « Committee on National Security Systems Committee on National Security Systems (CNSS) Glossary », 2022, Consulté le: août 12, 2022. [En ligne]. Available: https://www.cnss.gov.

[6] B. Jefferson, « The 15 Most Common Types of Cyber Attacks », *Lepide Home | Data Security & Compliance Blog*, août 02, 2020. https://www.lepide.com/blog/the-15-most-common-types-of-cyber-attacks/

[7] W. Duo, M. Zhou, et A. Abusorrah, « A Survey of Cyber Attacks on Cyber Physical Systems: Recent Advances and Challenges », *IEEE/CAA Journal of Automatica Sinica*, vol. 9, nº 5, p. 784-800, mai 2022, doi: 10.1109/JAS.2022.105548.

[8] F. David Janos et N. Huu Phuoc Dai, « Security Concerns Towards Security Operations Centers », in *2018 IEEE 12th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, mai 2018, p. 000273-000278. doi: 10.1109/SACI.2018.8440963.

[9] M. BARBEZAT, « Les Security Operations Centers SOC et leur rôle pour la cybersecurité », oct. 17, 2017. https://www.ledecodeur.ch/2017/10/17/security-operations-centers-soc-role-cybersecurite/

[10] L. Gorber et J. Thorpe, « Les Nouvelles Installations De Sécurité De CGI Au Royaume-Uni Permettront D'offrir Une Surveillance De Protection Et Des Services D'analyse De Cybermenaces », *CGI Inc, Centre des médias*, janv. 25, 2017. https://www.cgi.com/fr/CGI-Royaume-Uni-instatllations-securite-analyse-cybermenaces

[11] S. Al-Fedaghi et F. Mahdi, « Events Classification in Log Audit », *International journal of Network Security & Its Applications*, vol. 2, nº 2, p. 58-73, avr. 2010, doi: 10.5121/ijnsa.2010.2205.

[12] M. Souppaya et K. Scarfone, *Guide to Computer Security Log Management*, vol. 92. 2006.

[13] SIX B, « Log Formats – a (Mostly) Complete Guide », *Graylog*, janv. 15, 2020. https://www.graylog.org/post/log-formats-a-complete-guide (consulté le août 12, 2022).

[14] E. Kostrecová et H. Bínová, « Security Information and Event Management », *INDIAN JOURNAL OF RESEARCH*, vol. 4, nº 2, p. 119-120, févr. 2015.

[15] M. di Mauro et C. di Sarno, « Improving SIEM capabilities through an enhanced probe for encrypted Skype traffic detection », *Journal of Information Security and Applications*, vol. 38, p. 85-95, févr. 2018, doi: 10.1016/j.jisa.2017.12.001.

[16] N. Miloslavskaya, « Analysis of SIEM Systems and Their Usage in Security Operations and Security Intelligence Centers », in *Biologically Inspired Cognitive Architectures (BICA) for Young Scientists*, 2018, p. 282-288.

[17] Cybersecurity FOREVER, *Day-146: SIM vs. SEM vs. SIEM--What's The Difference?,* (déc. 13, 2020). [En ligne Video]. Available: https://www.youtube.com/watch?v=TFZXbRvXBzw

[18] T. L. Wiant, « Information security policy's impact on reporting security incidents », *Comput Secur*, vol. 24, nº 6, p. 448-459, sept. 2005, doi: 10.1016/J.COSE.2005.03.008.

[19] R. Werlinger, K. Muldner, K. Hawkey, et K. Beznosov, « Preparation, detection, and analysis: the diagnostic work of IT security incident response », *Information Management & Computer Security*, vol. 18, nº 1, p. 26-42, janv. 2010, doi: 10.1108/09685221011035241.

[20] S. M. M. Hossain, R. Couturier, J. Rusk, et K. B. Kent, « Automatic Event Categorizer for SIEM », in *Proceedings of the 31st Annual International Conference on Computer Science and Software Engineering*, 2021, p. 104-112.

[21] R. Gabriel, T. Hoppe, A. Pastwa, et S. Sowa, « Analyzing Malware Log Data to Support Security Information and Event Management: Some Research Results », in *2009 First International Confernce on Advances in Databases, Knowledge, and Data Applications*, 2009, p. 108-113. doi: 10.1109/DBKDA.2009.26.

[22] S. Kliger, S. Yemini, Y. Yemini, D. Ohsie, et S. Stolfo, « A Coding Approach to Event Correlation », in *Integrated Network Management IV: Proceedings of the fourth international symposium on integrated network management, 1995*, A. S. Sethi, Y. Raynaud, et F. Faure-Vincent, Éd. Boston, MA: Springer US, 1995, p. 266-277. doi: 10.1007/978-0-387-34890-2_24.

[23] Cable News Network, « IT - Gartner Inc Company Profile », *CNN*, juill. 01, 2018. https://money.cnn.com/quote/profile/profile.html?symb=IT (consulté le août 12, 2022).

[24] K. Kanavagh, T. Bussa, et J. Collins, « 2021 Gartner Magic Quadrant for SIEM », juin 2021. [En ligne]. Available: https://www.rapid7.com/c/2021-siem-mq/

[25] IBM Corporation, « QRadar overview - IBM Documentation », févr. 16, 2022. https://www.ibm.com/docs/en/qsip/7.5?topic=started-qradar-overview

[26] M. Hristov, M. Nenova, G. Iliev, et D. Avresky, « Integration of Splunk Enterprise SIEM for DDoS Attack Detection in IoT », in *2021 IEEE 20th International Symposium on Network Computing and Applications (NCA)*, nov. 2021, p. 1-5. doi: 10.1109/NCA53618.2021.9685977.

[27] LogRhythm Inc, « Ready to Defend », févr. 2022. [En ligne]. Available: https://gallery.logrhythm.com/brochures/ready-to-defend-overview-brochure.pdf

[28] McAfee, « SIEM Solutions from McAfee Data Sheet », mai 2019, [En ligne]. Available: https://www.mcafee.com/enterprise/en-us/assets/data-sheets/ds-siem-solutions-from-mcafee.pdf

[29] Mozilla Corporation, « Mozilla Enterprise Defense Platform documentation », juin 21, 2021. https://mozdef.readthedocs.io/en/latest/index.html

[30] Elastic NV, « What is the ELK Stack? », *Elastic.co*, 2022. https://www.elastic.co/what-is/elk-stack

[31] Elastic NV, « Beats: Data Shippers for Elasticsearch », *Elastic.co*, 2022. https://www.elastic.co/beats/

[32] I. Sklavidis, C. Angelidis, R. Babagiannou, et A. Liapis, « Enhancing SIEM Technology for protecting Electrical Power and Energy Sector », in *2021 IEEE International Conference on Cyber Security and Resilience (CSR)*, juill. 2021, p. 473-478. doi: 10.1109/CSR51186.2021.9527944.

[33] SIEMonster, « SIEMonster | Affordable Security Monitoring Software Solution », déc. 20, 2015. https://siemonster.com/

[34] Elastic NV, « Elastic Stack: Elasticsearch, Kibana, Beats & Logstash », *Elastic.co*, 2022. https://www.elastic.co/elastic-stack/

[35] Dell Technologies Info Hub, « Architecture overview | Elastic Stack on Dell EMC VxRail », nov. 23, 2020. https://infohub.delltechnologies.com/l/elastic-stack-on-dell-emc-vxrail/architecture-overview-75

[36] Elastic NV, « What is Elasticsearch », *Elastic.co*, juill. 17, 2019. https://www.elastic.co/what-is/elasticsearch

[37] A. Goliva, « What is Elastic Stack and Where to Use it », *Tecnología*, févr. 06, 2020. https://medium.com/tecnolog%C3%ADa/what-is-elastic-stack-and-where-to-use-it-614cdccc0b03

[38] Elastic NV, « Amazon Web Services Inc », déc. 19, 2018. https://aws.amazon.com/opensearch-service/the-elk-stack/logstash/

[39] Elastic NV, « What are Beats ? », *Elastic.co*, oct. 22, 2015. https://www.elastic.co/guide/en/beats/libbeat/current/beats-reference.html

[40] Elastic NV, « Qu'est-ce que Kibana ? », *Elastic.co*, mars 21, 2020. https://www.elastic.co/fr/what-is/kibana

[41] A. Aleryani, « Comparative Study between Data Flow Diagram and Use Case Diagram », *International Journal of Scientific and Research Publications*, vol. 6, p. 124-2250, avr. 2016.

[42] Oracle Corporation, « Oracle VM VirtualBox User Manual Version 6.1.34 », *Virtualbox.org*, juill. 19, 2022. https://www.virtualbox.org/manual/ch01.html

[43] Z. Sikora, « Introduction », in *Java*, Elsevier, 2003, p. 1-6. doi: 10.1016/B978-155860909-9/50001-6.

[44] C. Allegretta, « GNU nano a small and friendly text editor version 6.4 ». juin 14, 2016. [En ligne]. Available: https://www.nano-editor.org/dist/latest/nano.pdf

[45] K. Kinder, « Sublime text: one editor to rule them all? », *Linux Journal*, vol. 2013, août 2013.

[46] M. Russinovich et T. Garnier, « Sysmon - Windows Sysinternals », *docs.microsoft.com*, mai 25, 2017. https://docs.microsoft.com/fr-fr/sysinternals/downloads/sysmon

[47] D. J. Day et B. M. Burns, « A performance analysis of Snort and Suricata Network Intrusion Detection and Prevention Engines », in *ICDS 2011 : The Fifth International Conference on Digital Society*, 2011, p. 187-192.

[48] Cybersecurity & Infrastructure Security Agency, « ABOUT CISA », *Library of Congress, Washington, D.C., 20540 USA*, sept. 10, 2019. https://www.cisa.gov/about-cisa

[49] CISA, « Understanding Denial-of-Service Attacks », *www.cisa.gov*, nov. 04, 2009. https://www.cisa.gov/uscert/ncas/tips/ST04-015

[50] K. A. Scarfone, T. Grance, et K. Masone, « Computer security incident handling guide », Gaithersburg, MD, 2008. doi: 10.6028/NIST.SP.800-61r1.

[51] K. Chhillar et S. Shrivastava, « University Computer Network Vulnerability Management using Nmap and Nexpose », *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 10, n° 6, 2021, doi: 10.30534/ijatcse/2021/021062021.

[52] M. van Eeten et J. Bauer, « Economics of Malware », *OECD Science, Technology and Industry Working Papers*, mai 2008, doi: 10.1787/241440230621.

[53] C. Liu, Y. Sheng, Z. Wei, et Y.-Q. Yang, « Research of Text Classification Based on Improved TF-IDF Algorithm », in *2018 IEEE International Conference of Intelligent Robotic and Control Engineering (IRCE)*, août 2018, p. 218-222. doi: 10.1109/IRCE.2018.8492945.

Annexe

I. Installation de Sysmon

```
E:\SYSMON>sysmon -accepteula -i c:\windows\config.xml
```

Annexe 1 : Commande d'installation Sysmon.

- **-accepteula :** pour accepter automatiquement le Contrat de Licence d'Utilisateur Final lors de l'installation.
- **-i :** pour spécifier un fichier de configuration.
- **c:\windows\config.xml :** Il s'agit d'un fichier de configuration Microsoft Sysmon avec un suivi des événements de haute qualité fourni par SwiftOnSecurity.

II. Installation de JAVA

La commande pour installer Java OpenJDK (Open Java Development Kit) 11 est :

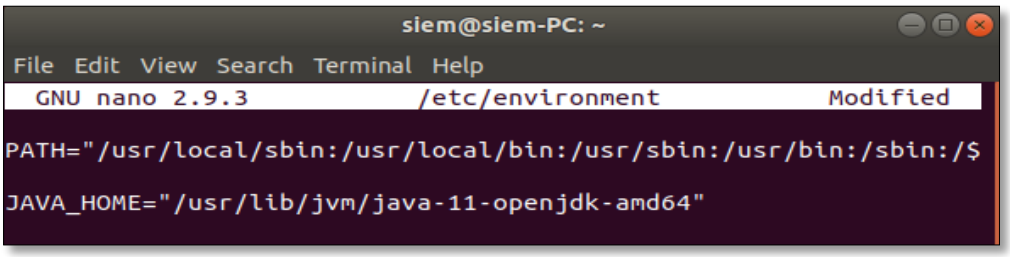
```
siem@siem-PC:~$ sudo apt-get install openjdk-11-jdk
```

Annexe 2 : Commande d'installation de Java.

Pour s'assurer que Java est correctement installé, nous testons une commande qui renvoie la version java. Ensuite, et une fois que java est correctement installé, nous configurons la variable d'environnement **JAVA_HOME** dans le fichier **etc/environment**.

```
siem@siem-PC:~$ sudo java -version
[sudo] password for siem:
openjdk version "11.0.15" 2022-04-19
OpenJDK Runtime Environment (build 11.0.15+10-Ubuntu-0ubuntu0.18.04.1)
OpenJDK 64-Bit Server VM (build 11.0.15+10-Ubuntu-0ubuntu0.18.04.1, mixed mode,
sharing)
```

Annexe 3 : Commande d'affichage de la version de Java.



Annexe 4 : Variable d'environnement JAVA_HOME.

III. Installation de Logstash

```
siem@siem-PC:~$ sudo apt-get install logstash
```

Annexe 5 : Commande d'installation de Logstash.

Tout d'abord, la section d'entrée (input), nous ajoutons le port qui fait référence au port Beats que Logstash écoutera pour le trafic entrant.

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

Annexe 6 : Fichier de configuration Logstash input.

Ensuite, la section de filtre (Parsing), nous configurons l'analyse des journaux Syslog (System Logging Protocol) avec Grok (Grok est un outil simple qui permet de transformer facilement les données non structurées des journaux et des événements en données structurées JSON).

```
filter {
  if [type] == "syslog" {
    grok {
      match => { "message" => "%{SYSLOGTIMESTAMP:syslog_timestamp} %{SYSLOGHOST:syslog_hostname}
        %{DATA:syslog_program}(?:\[%{POSINT:syslog_pid}\])?: %{GREEDYDATA:syslog_message}" }
      add_field => [ "received_at", "%{@timestamp}" ]
      add_field => [ "received_from", "%{host}" ]
    }
    date {
      match => [ "syslog_timestamp", "MMM  d HH:mm:ss", "MMM dd HH:mm:ss" ]
    }
  }
}
```

Annexe 7 : Fichier de configuration Logstash filter.

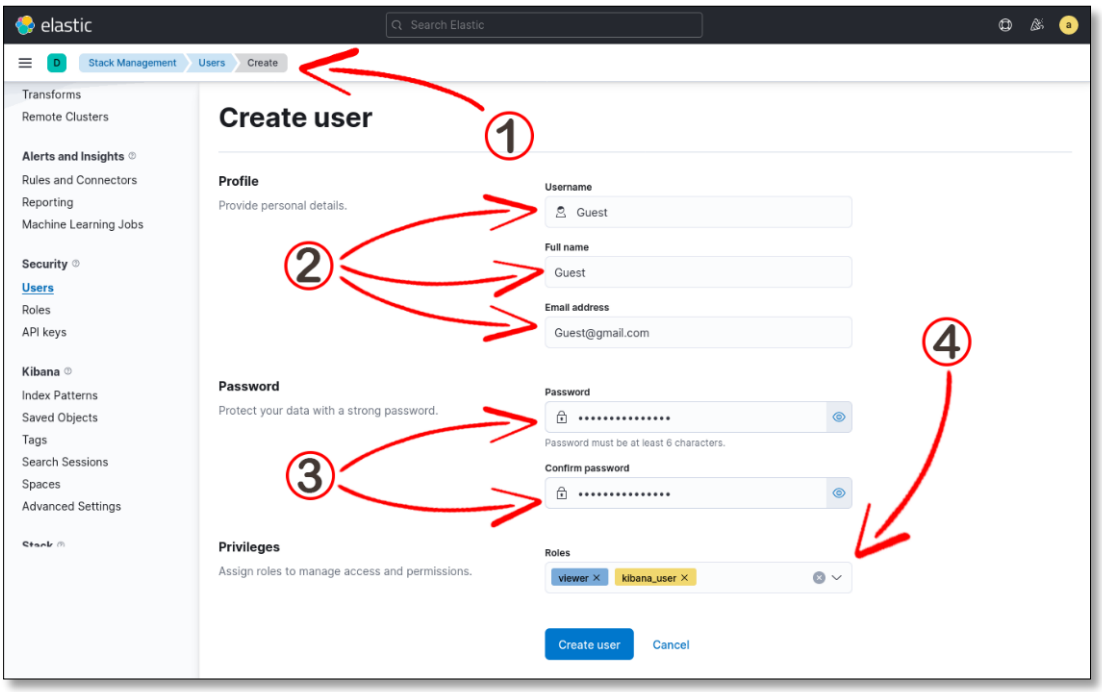
La dernière étape est de configurer la section de sortie (output), nous spécifions l'URL de l'instance Elasticsearch et l'index où nous stockerons les journaux dans Elasticsearch.

```
output {

  elasticsearch {

    hosts => ["192.168.1.7:9200"]

    index => "%{[@metadata][beat]}-%{[@metadata][version]}-%{+YYYY.MM.dd}"

  }

}
```

Annexe 8 : Fichier de configuration Logstash output.

IV. Création des utilisateurs

Il existe également un ensemble de rôles intégrés que nous pouvons attribuer aux utilisateurs, ces rôles ont un ensemble fixe de privilèges et ne peuvent pas être mis à jour, par exemple **Kibana_admin** accorde l'accès à toutes les fonctionnalités de Kibana, **Logstash_admin** accorde l'accès aux index Logstash pour la gestion de la configuration et accorde l'accès nécessaire aux API spécifiques à Logstash exposées par le plug-in **logstash x-pack.** La figure suivante montre les étapes de création d'un utilisateur en attribuant différents rôles.

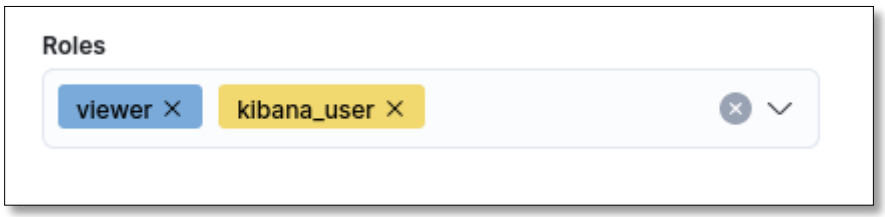


Annexe 9 : étapes de création d'un utilisateur.

1. Représente le chemin où se trouve cette page

Annexe 10 : chemin d'accès au processus de création de l'utilisateur.

2. Les informations personnelles de l'utilisateur telles que : **Username, Full name, Email address**.
3. Champ de mot de passe + confirmation du mot de passe
4. Champ où attribuer des rôles à l'utilisateur



Annexe 11 : Champ des rôles.

V. Auditbeat

Pour installer Auditbeat :

```
ubuntu@ubuntu-PC:~$ sudo apt-get install auditbeat
```

Annexe 12 : Commande d'installation d'Auditbeat.

Avant de démarrer Auditbeat, nous devons charger les actifs prédéfinis pour l'analyse, l'indexation et la visualisation des données, puis nous pouvons lancer Auditbeat pour commencer à collecter les journaux et les envoyer à Elasticsearch.

```
ubuntu@ubuntu-PC:~$ sudo auditbeat setup
```

Annexe 13 : Commande de chargement des actifs prédéfinis d'Auditbeat.

```
ubuntu@ubuntu-PC:~$ sudo systemctl start auditbeat
```

Annexe 14 : Commande de démarrage d'Auditbeat.

VI. Filebeat

Une fois que Logstash est configuré pour recevoir les logs de beats, nous procédons à l'installation de Filebeat en saisissant la commande suivante :

```
siem@siem-PC:~$ sudo apt-get install filebeat
```

Annexe 15 : Commande d'installation de Filebeat.

Ensuite, nous configurons Filebeat en accédant au fichier **/etc/filebeat/ filebeat.yml,** et spécifions l'adresse IP du système **192.168.1.7** avec le port **5044** pour que Filebeat puisse se connecte à Logstash.

```
# ------------------------------ Logstash Output -------------------------------
output.logstash:
  # The Logstash hosts
  hosts: ["192.168.1.7:5044"]

  # Optional SSL. By default is off.
  # List of root certificates for HTTPS server verifications
  #ssl.certificate_authorities: ["/etc/pki/root/ca.pem"]

  # Certificate for SSL client authentication
  #ssl.certificate: "/etc/pki/client/cert.pem"
```

Annexe 16 : Fichier de configuration Filebeat -Logstash Output.

VII. Metricbeat

Le processus d'installation de Metricbeat et de configuration des sorties Elasticsearch et Kibana est le même que pour toute autre installation/configuration de beats.

```
siem@siem-PC:~$ sudo apt-get install metricbeat
```

Annexe 17 : Commande d'installation de Metricbeat.

```
# ---------------------------- Elasticsearch Output ----------------------------
output.elasticsearch:
  # Array of hosts to connect to.
  hosts: ["http://192.168.1.7:9200"]
  # Protocol - either `http` (default) or `https`.
  protocol: "http"

  # Authentication credentials - either API key or username/password.
  #api_key: "id:api_key"
  username: "admin"
  password: "m2SSI2022"

# ------------------------------ Logstash Output -------------------------------
```

Annexe 18 : Fichier de configuration Metricbeat -Elasticsearch Output.

```
# ================================== Kibana ===================================

# Starting with Beats version 6.0.0, the dashboards are loaded via the Kibana API.
# This requires a Kibana endpoint configuration.
setup.kibana:
  host: "http://192.168.1.7:5601"
  username: "admin"
  password: "m2SSI2022"
  # Kibana Host
  # Scheme and port can be left out and will be set to the default (http and 5601)
  # In case you specify and additional path, the scheme is required: http://localhost:5601/path
  # IPv6 addresses should always be defined as: https://[2001:db8::1]:5601
  #host: "localhost:5601"

# ============================= Elastic Cloud ==============================
```

Annexe 19 : Fichier de configuration Metricbeat -Kibana Output.

Metricbeat utilise des modules pour collecter des métriques. Chaque module définit la logique de base pour collecter les données d'un service spécifique. Un module se compose d'ensembles de métriques qui récupèrent et structurent les données, dans notre solution, nous avons choisi le **module système** et le **module Elasticsearch xpack**. Pour voir la liste des modules disponibles :

```
siem@siem-PC:~$ sudo metricbeat modules list
```

Annexe 20 : Commande d'affichage liste des modules Metricbeat.

La commande suivante active un module Metricbeat :

```
siem@siem-PC:~$ sudo metricbeat modules enable {nom_module}
```

Annexe 21 : Commande d'activation d'un module Metricbeat.

Une fois que tout est configuré, nous chargeons les actifs prédéfinis pour l'analyse, l'indexation et la visualisation des données, puis nous lançons Metricbeat

```
siem@siem-PC:~$ sudo metricbeat setup
```

Annexe 22 : Commande de chargement des actifs prédéfinis de Metricbeat.

```
siem@siem-PC:~$ sudo systemctl start metricbeat
```

Annexe 23 : Commande de démarrage de Metricbeat.

VIII. Packetbeat

Sur la plupart des plates-formes, Packetbeat requiert la bibliothèque de capture de paquets **libpcap (Promiscuous Capture Library)**. Pour installer libpcap :

```
ubuntu@ubuntu-PC:~$ sudo apt-get install libpcap0.8
```

Annexe 24 : Commande d'installation de la bibliothèque libpcap.

Comme toute autre installation de beats, nous installons Packetbeat et configurons les sorties Elasticsearch et Kibana

```
ubuntu@ubuntu-PC:~$ sudo apt-get install packetbeat
```

Annexe 25 : Commande d'installation de Packetbeat.

```
# ---------------------------- Elasticsearch Output ----------------------------
output.elasticsearch:
  # Array of hosts to connect to.
  hosts: ["http://192.168.1.7:9200"]

  # Protocol - either `http` (default) or `https`.
  protocol: "http"

  # Authentication credentials - either API key or username/password.
  #api_key: "id:api_key"
  username: "admin"
  password: "m2SSI2022"

# ------------------------------ Logstash Output -------------------------------
```

Annexe 26 : Fichier de configuration Packetbeat -Elasticsearch Output.

```
# ================================== Kibana ===================================

# Starting with Beats version 6.0.0, the dashboards are loaded via the Kibana AP
# This requires a Kibana endpoint configuration.
setup.kibana:

  # Kibana Host
  # Scheme and port can be left out and will be set to the default (http and 560
  # In case you specify and additional path, the scheme is required: http://loca
  # IPv6 addresses should always be defined as: https://[2001:db8::1]:5601
  host: "http://192.168.1.7:5601"
  username: "admin"
  password: "m2SSI2022"

# ============================== Elastic Cloud ================================
```

Annexe 27 : Fichier de configuration Packetbeat -Kibana Output.

Après avoir terminé la configuration de la connexion de Packetbeat à Elasticsearch et Kibana, nous configurons les paramètres de sniffing, nous devons configurer les périphériques réseau et les protocoles afin de capturer le trafic réseau.

Tout d'abord, nous définissons le type de renifleur (sniffer) sur **« af_packet »** pour utiliser le reniflement mappé en mémoire et spécifions les interfaces réseau à **« any »** pour capturer tous les messages envoyés ou reçus par le serveur sur lequel Packetbeat est installé.

```
# ============================ Network device ================================

# Select the network interface to sniff the data. On Linux, you can use the
# "any" keyword to sniff on all connected interfaces.
packetbeat.interfaces.type: af_packet
packetbeat.interfaces.device: any

# ================================== Flows ====================================
```

Annexe 28 : Fichier de configuration Packetbeat -Interfaces réseau.

Dans la section des protocoles, nous configurons les ports où Packetbeat peut trouver chaque protocole

```
# ========================== Transaction protocols ============================
packetbeat.protocols:
- type: icmp
  # Enable ICMPv4 and ICMPv6 monitoring.
  enabled: true

- type: dhcpv4
  # Configure the DHCP for IPv4 ports.
  ports: [67, 68]

- type: dns
  # Configure the ports where to listen for DNS traffic.
  ports: [53]

- type: http
  # Configure the ports where to listen for HTTP traffic.
  ports: [80, 8080, 8000, 5000, 8002]

- type: mysql
  # Configure the ports where to listen for MySQL traffic.
  ports: [3306,3307]

- type: pgsql
  # Configure the ports where to listen for Pgsql traffic.
  ports: [5432]

- type: redis
  # Configure the ports where to listen for Redis traffic.
  ports: [6379]

- type: mongodb
  # Configure the ports where to listen for MongoDB traffic.
  ports: [27017]

- type: nfs
  # Configure the ports where to listen for Network File System traffic.
  ports: [2049]

- type: tls
  # Configure the ports where to listen for TLS traffic.
  ports:
    - 443   # HTTPS
# ======================= Elasticsearch template setting ======================
```

Annexe 29 : Fichier de configuration Packetbeat -Protocoles de transactions.

Une fois la phase de configuration terminée, nous chargeons les actifs prédéfinis pour l'analyse, l'indexation et la visualisation des données. Puis on démarre Packetbeat

```
ubuntu@ubuntu-PC:~$ sudo packetbeat setup
```

Annexe 30 : Commande de chargement des actifs prédéfinis de Packetbeat.

```
ubuntu@ubuntu-PC:~$ sudo systemctl start packetbeat
```

Annexe 31 : Commande de démarrage de Packetbeat.

IX. Winlogbeat

Pour installer Winlogbeat sur notre machine Windows 10, nous suivons ces étapes:

- Nous téléchargeons le fichier Winlogbeat compressé à partir de la page de téléchargement https://www.elastic.co/downloads/beats/winlogbeat et nous l'extrayons dans le chemin **C:\Program Files\Winlogbeat**
- Nous ouvrons une invite PowerShell en tant qu'administrateur, nous accédons au chemin du répertoire Winlogbeat et exécutons la commande d'installation comme le montre la figure ci-dessous :

```
PS C:\Program Files\Winlogbeat> .\install-service-winlogbeat.ps1
```

Annexe 32 : Commande d'installation de winlogbeat.

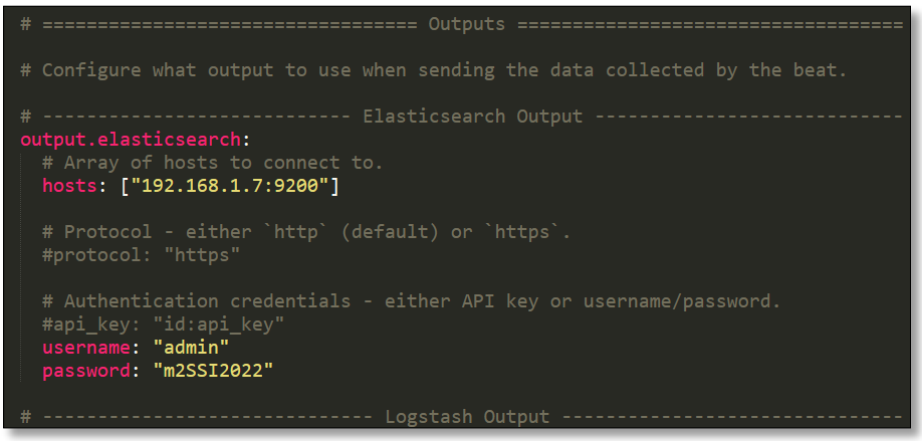
- Un avertissement de sécurité apparaîtra dans l'invite PowerShell nous demandant d'autoriser l'installation des scripts Winlogbeat, nous autorisons le processus d'installation en appuyant sur "R"

```
Security warning
Run only scripts that you trust. While scripts from the internet can be useful, this script can potent
 the Unblock-File cmdlet to allow the script to run without this warning message. Do you want to run C
Files\Winlogbeat\install-service-winlogbeat.ps1?
[D] Do not run  [R] Run once  [S] Suspend  [?] Help (default is "D"): R
```

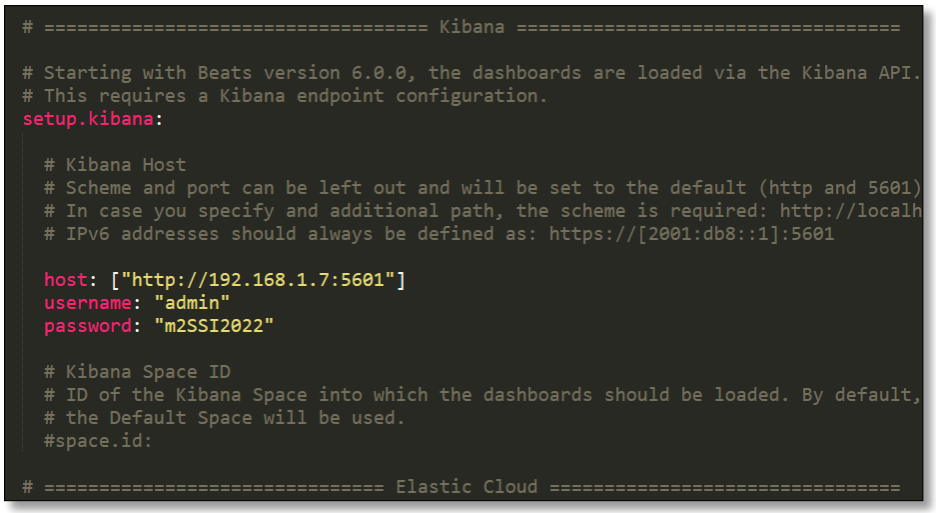
Annexe 33 : avertissement de sécurité PowerShell.

- Une fois le processus d'installation terminé, Winlogbeat fonctionnera en tant que service Windows.

Ensuite, nous configurons les sorties Elasticsearch et Kibana afin que Winlogbeat puisse se connecter à la fois à Elasticsearch pour envoyer des journaux et à Kibana pour charger les ressources prédéfinies.



Annexe 34 : Fichier de configuration Winlogbeat -Elasticsearch Output.



Annexe 35 : Fichier de configuration Winlogbeat -Kibana Output.

Après avoir terminé la configuration de la connexion de Winlogbeat à la Pile Elastic, nous passons à la configuration des journaux d'événements que Winlogbeat surveillera, dans le fichier **winlgobeat.yml** et sous **winlogbeat.event_logs**, nous spécifions une liste de journaux d'événements à surveiller.

```
# ========================== Winlogbeat specific options ==========================

# event_logs specifies a list of event logs to monitor as well as any
# accompanying options. The YAML data type of event_logs is a list of
# dictionaries.
#
# Please visit the documentation for the complete details of each option.
# https://go.es.io/WinlogbeatConfig

winlogbeat.event_logs:
  - name: Application
    ignore_older: 72h

  - name: System

  - name: Security
    processors:
      - script:
          lang: javascript
          id: security
          file: ${path.home}/module/security/config/winlogbeat-security.js

  - name: Microsoft-Windows-Sysmon/Operational
    processors:
      - script:
          lang: javascript
          id: sysmon
          file: ${path.home}/module/sysmon/config/winlogbeat-sysmon.js

  #configuration .....

# ====================== Elasticsearch template settings ======================
```

Annexe 36 : Fichier de configuration Winlogbeat -Winlogbeat specific options.

Winlogbeat est fourni avec des ressources prédéfinies pour analyser, indexer et visualiser les données. Pour charger ces éléments :

```
PS C:\Program Files\Winlogbeat> .\winlogbeat.exe setup -e
```

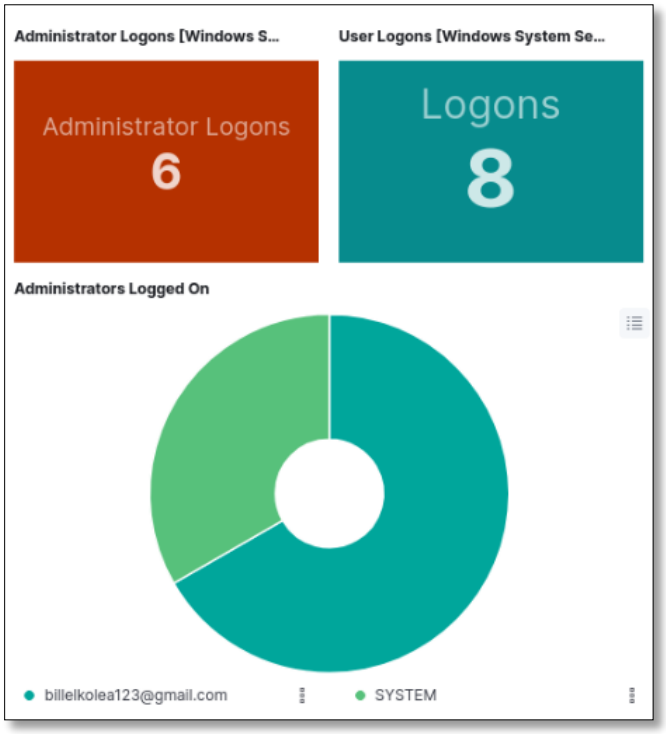
Annexe 37 : Commande de chargement des actifs prédéfinis de Winlogbeat.

Une fois que nous avons tout configuré, nous lançons Winlogbeat avec la commande suivante :

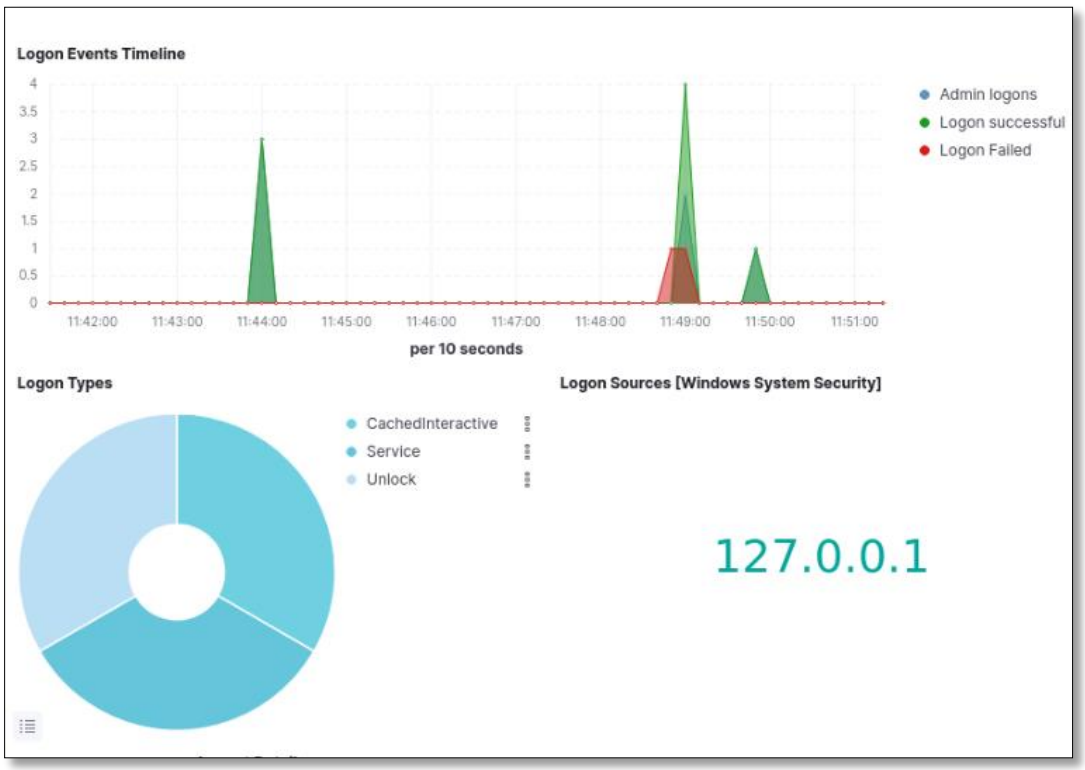
```
PS C:\Program Files\Winlogbeat> Start-Service winlogbeat
```

Annexe 38 : Commande de démarrage de Winlogbeat.

X. Dashboards



Annexe 39 : Dashboard_User Logons Windows part1.



Annexe 40 : Dashboard_User Logons Windows part 2.

Logon Details

8 documents

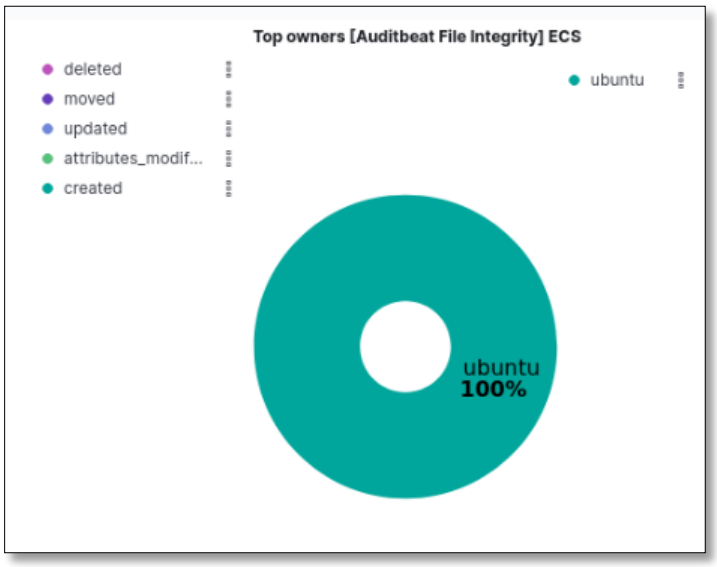
Time	user.name	winlog.logon.type	source.domain
Jun 16, 2022 @ 11:49:54.935	SYSTEM	Service	-
Jun 16, 2022 @ 11:49:05.603	billelkolea123	Unlock	DESKTOP-G8P3JB8
Jun 16, 2022 @ 11:49:05.603	billelkolea123	Unlock	DESKTOP-G8P3JB8
Jun 16, 2022 @ 11:49:05.553	billelkolea123	CachedInteractive	DESKTOP-G8P3JB8
Jun 16, 2022 @ 11:49:05.553	billelkolea123	CachedInteractive	DESKTOP-G8P3JB8
Jun 16, 2022 @ 11:44:03.426	SYSTEM	Service	-
Jun 16, 2022 @ 11:44:03.419	SYSTEM	Service	-
Jun 16, 2022 @ 11:44:01.709	SYSTEM	Service	-

Logout Details

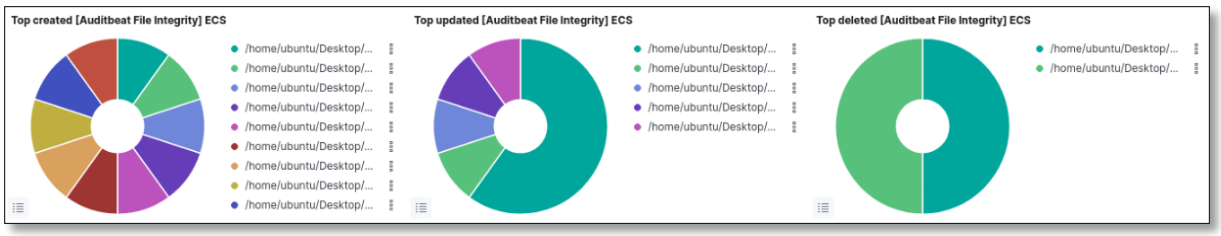
Time	user.name	user.domain	winlog.logon.id
Jun 16, 2022 @ 11:49:00.858	-	-	0x3e7
Jun 16, 2022 @ 11:48:59.802	-	-	0x3e7

Annexe 41 : Dashboard_User Logons Windows part 3.

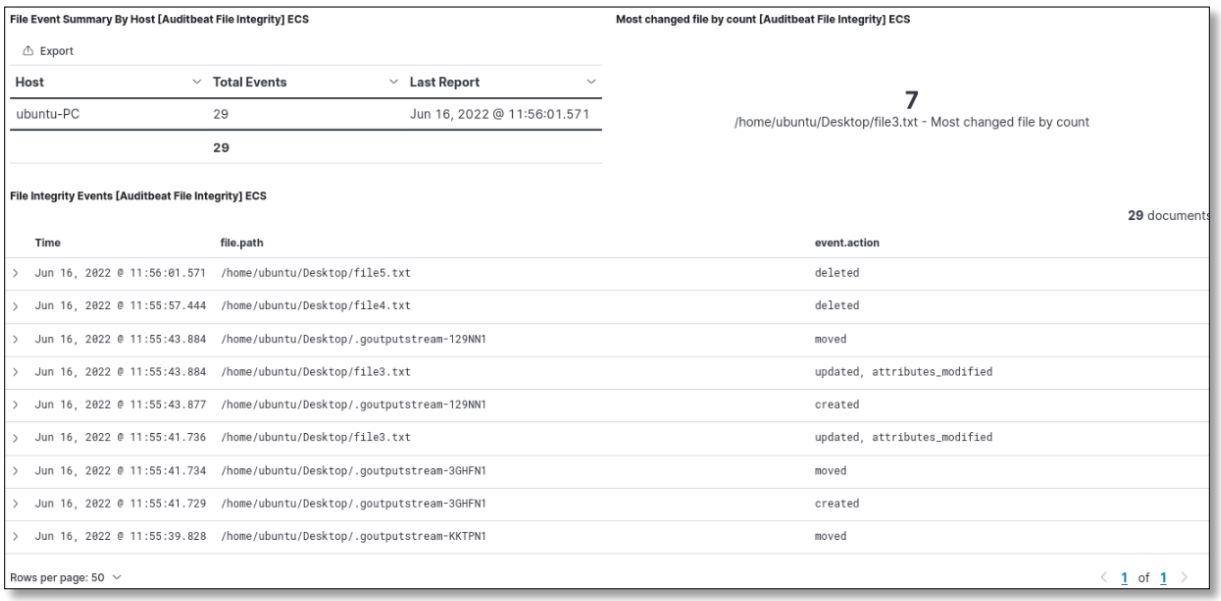
Annexe 42 : Dashboard_Auditbeat File Integrity Overview part1.



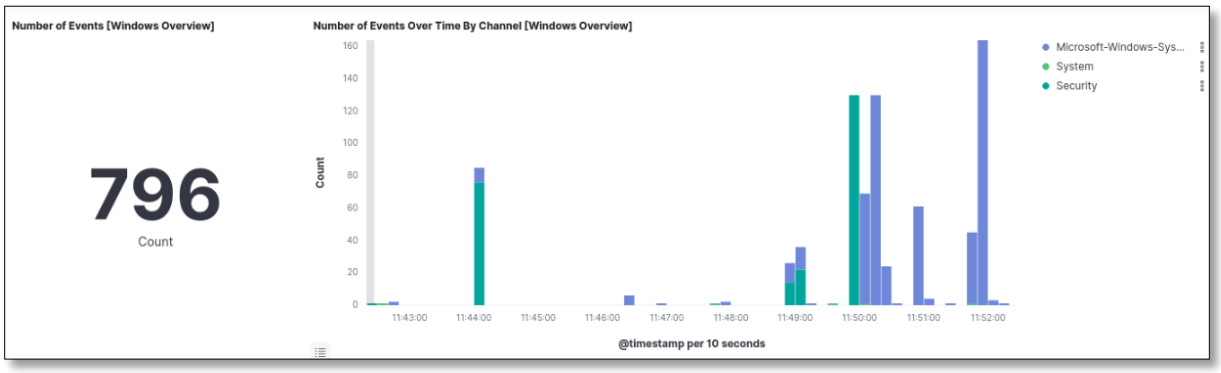
Annexe 43 : Dashboard_Auditbeat File Integrity Overview part 2.



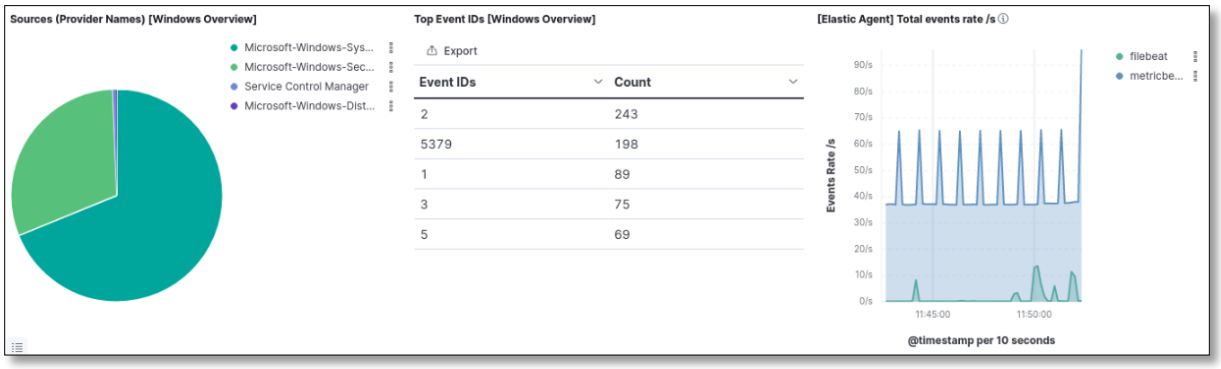
Annexe 44 : Dashboard_Auditbeat File Integrity Overview part 3.



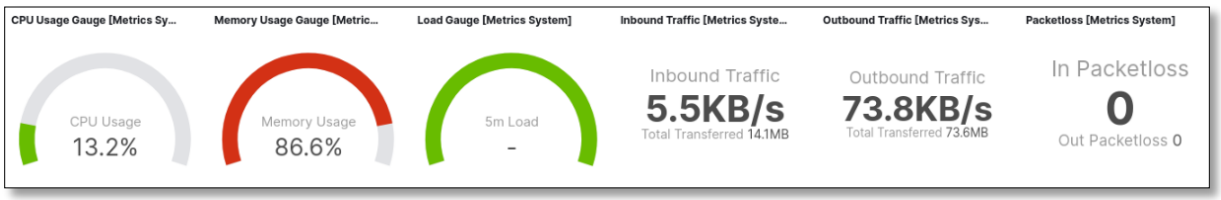
Annexe 45 : Dashboard_Auditbeat File Integrity Overview part 4.



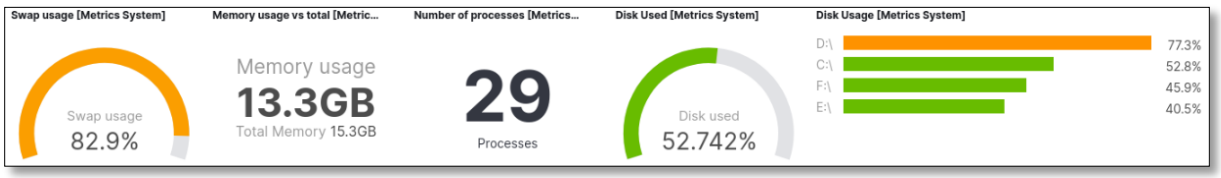
Annexe 46 : Dashboard_Windows Service Overview part 1.



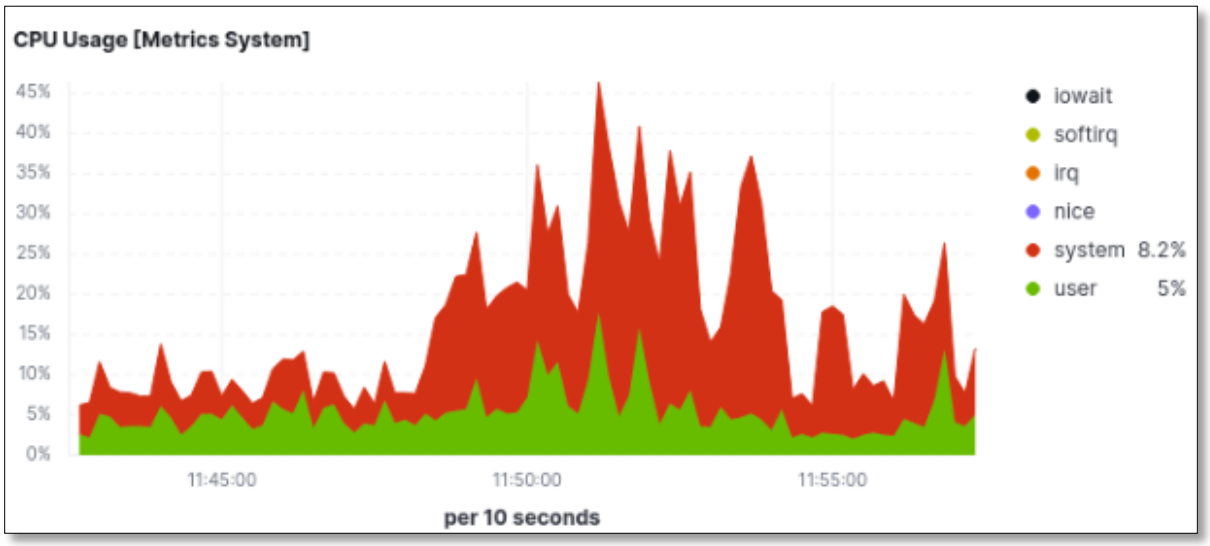
Annexe 47 : Dashboard_Windows Service Overview part 2.



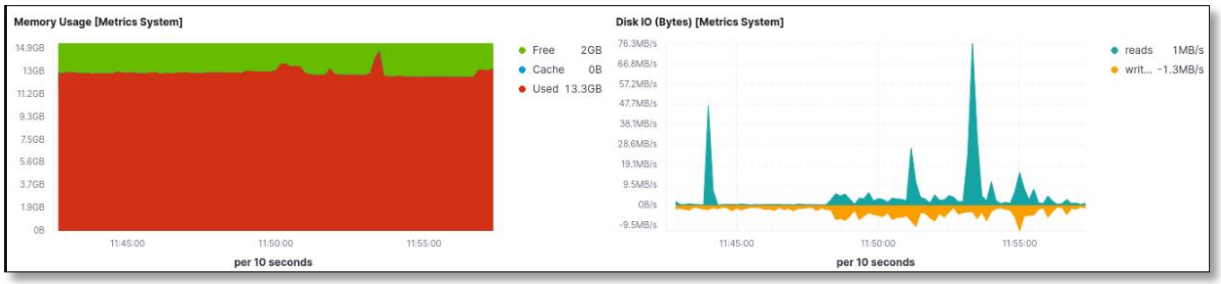
Annexe 48 : Dashboard Host Overview part 1.



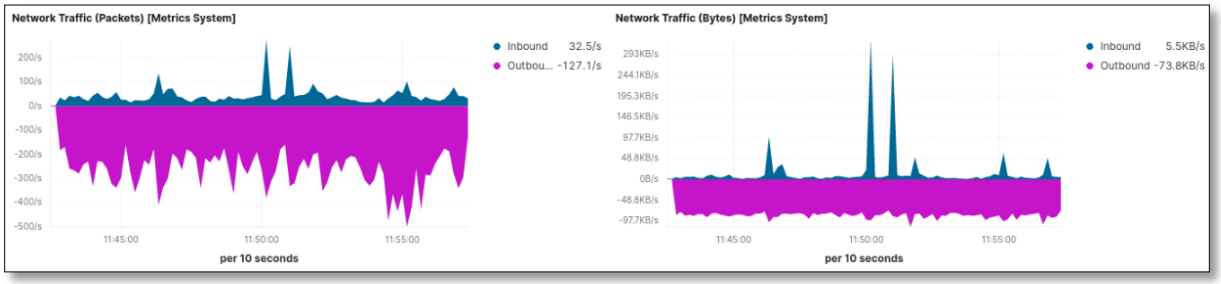
Annexe 49 : Dashboard Host Overview part 2.



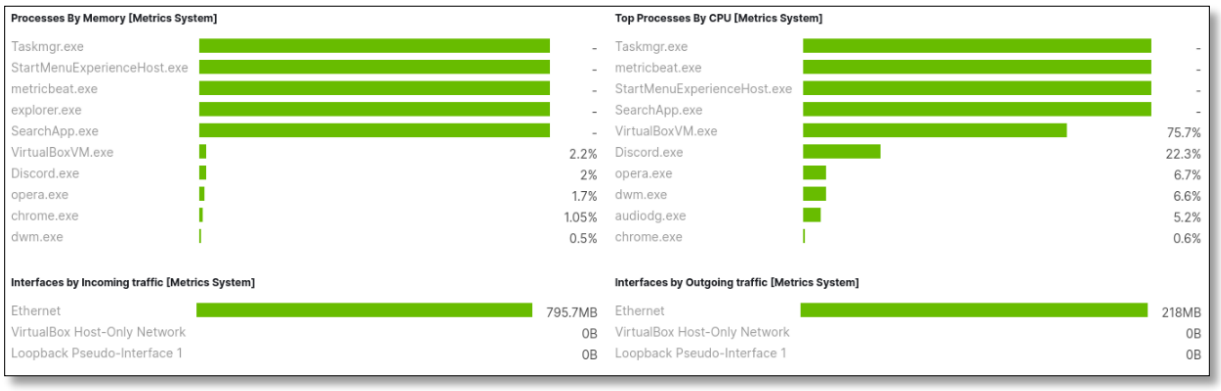
Annexe 50 : Dashboard Host Overview part 3.



Annexe 51 : Dashboard Host Overview part 4.



Annexe 52 : Dashboard Host Overview part 5.



Annexe 53 : Dashboard Host Overview part 6.

XI. Algorithme TF-IDF

« *L'algorithme TF-IDF est une méthode statistique utilisée pour évaluer l'importance d'un mot pour un document dans un ensemble de fichiers. L'idée principale est que si un mot ou une phrase apparaît fréquemment dans un article et qu'il est rarement trouvé dans d'autres articles, on considère que le mot ou la phrase a une bonne capacité de distinction de classe et convient à la classification.* » [53].