





**UNIVERSITE GASTON BERGER DE SAINT-LOUIS**

**UFR DE SCIENCES APPLIQUEES ET DE TECHNOLOGIE**

**Centre de Formation Professionnelle Pluridisciplinaire**

**(CFPP)**

**MASTER PROFESSIONNELLE RESEAUX ET TELECOMMUNICATIONS**

**(MART)**

**Mémoire**

# Système intelligent de corrélation et automatisation de la réponse aux incidents

**Présentée et Soutenue par:**

Yarame Niang

**Encadré par :**

Pr. Cherif Diallo

**Maitre de stage :**

M. Mouhamadou M. Mbaye

M. Mouhamadou L. Diakhame

**Année Académique 2024/2025**

# Dédicaces

À ma brave mère, ce travail est le reflet de ton amour, de tes prières et de tes innombrables sacrifices. Merci pour ton soutien inconditionnel, ta patience et la force que tu m'as transmise dans chaque étape de ma vie. Sans toi, ce travail n'aurait jamais vu le jour. Je t'offre humblement ce mémoire comme témoignage de ma profonde gratitude et de mon amour éternel.

A mon papa Cheikh Abdoulaye Sow, pour ses nombreuses prières, sa présence et

tout l'amour et l'attention qu'il me porte.

A mon papa Matabara Niang que j'aime tant.

A mes tofos d'amour Ibrahima Khaliloulahi Diao et Mouhamed Niang.

A ma grande sœur, mon repère, Bana Diop, merci d'être toi.

A ma petite sœur chérie Adja Fatou Gueye.

A mon oncle Alioune Badara Gueye, merci pour tout ton soutien.

A mes amis, Yaye Arame Bakar Faye, Ndeye Rama Sarr, Cheikh Gningue, El Mokhtar Abdy, Abdallah Mahamat Bokhite.

A la personne Grande et Forte que je deviens.

A toutes les personnes qui m'accompagnent et me soutiennent.

# Remerciements

Gloire à Allah le tout puissant, lui qui éclaire chacun de mes pas et me donne la force de surmonter les épreuves. Je lui confie la suite de mon chemin et qu'il continue de me guider.

Je tiens à exprimer ma profonde gratitude à ma mère, Ndeye Khoudia Gueye la personne qui a le plus marqué mon parcours. Elle a porté mes rêves, soutenu mes ambitions et affronté les difficultés avec une force admirable. Chaque étape de mes études porte son empreinte et ce mémoire est autant le sien que le mien.

Je remercie vivement mes encadreurs, Professeur Chérif Diallo, Monsieur Mouhamadou Lamine Diakhame et Monsieur Mouhamadou Moustapha Mbaye pour leurs conseils avisés, leur disponibilité ainsi que leurs soutiens constants tout au long de ce travail. Leurs remarques pertinentes et exigences scientifiques ont grandement contribué à la qualité de ce mémoire.

Une sincère gratitude à l'ensemble des professeurs qui m'ont accompagné tout au long de mon parcours universitaire. Par leurs conseils et la qualité de leurs enseignements, ils ont contribué à renforcer mes connaissances, à développer mon esprit critique et à façonner la professionnelle que je deviens aujourd'hui.

J'exprime ma reconnaissance à Madame Ndeye Aissatou SY Sow dont la disponibilité et le professionnalisme ont été d'un soutien précieux. Sa capacité à accompagner et orienter avec douceur et efficacité a apporté un grand équilibre tout au long de ce processus.

Mes remerciements vont également à ma famille, qui m'a toujours entourée de son amour et de son soutien. Leur confiance en moi, leurs encouragements constants et leur présence, même silencieuse, ont été des piliers essentiels dans les moments de doutes comme dans les moments de joie.

## Résumé

La supervision de la sécurité d'un système d'information représente aujourd'hui un enjeu majeur pour les organisations, en raison de la complexité croissante des infrastructures informatiques et de la multiplication des menaces cybernétiques. Les équipes de sécurité doivent analyser en continu d'importants volumes de données issues des journaux d'événements et des systèmes de détection d'intrusion. Cette tâche, à la fois complexe et chronophage, rend la corrélation manuelle des alertes et la réponse rapide aux incidents difficiles à assurer. Face à cette problématique, ce mémoire présente le développement d'un système intelligent de corrélation et d'automatisation de la réponse aux incidents, conçu pour améliorer la détection et la gestion des menaces dans les environnements de cybersécurité modernes. Le document rappelle les fondements de la cybersécurité ainsi que l'apport de l'intelligence artificielle dans l'analyse avancée des données de sécurité. Il présente par la suite les concepts de SIEM, de corrélation d'événements et de SOAR, en montrant leur rôle complémentaire dans l'optimisation de la détection et de la réponse. La partie conception détaille l'architecture du mécanisme développé, intégrant un LLM (Large Language Model ou Modèle de Grand Langage) chargé d'analyser automatiquement les alertes, un SIEM ElasticStack (Elasticsearch, Kibana, Filebeat) pour la collecte des données, et un SOAR Cortex/TheHive pour l'orchestration des réponses. Le système est également connecté à un répondeur SentinelOne pour automatiser des actions comme l'isolement d'hôte. L'implémentation finale démontre que la combinaison de l'IA, du SIEM et du SOAR permet une corrélation intelligente, une réduction significative du bruit d'alertes et une réponse automatisée plus rapide et plus efficace face aux incidents de sécurité. Ce travail met en lumière le potentiel des approches intelligentes pour renforcer la posture de sécurité du Centre de Calcul Ousmane SECK (CCOS) et ouvre des perspectives pour des améliorations futures du système.



**Mots-clés :** Cybersécurité, SIEM, Corrélation, SOAR, LLM, Système d'information

## Abstract

The monitoring of an information system's security has become a major challenge for organizations today, due to the increasing complexity of IT infrastructures and the growing number of cyber threats. Security teams must continuously analyze large volumes of data generated by event logs and intrusion detection systems. This task, both complex and time-consuming, makes manual alert correlation and rapid incident response difficult to achieve. To address this issue, this dissertation presents the development of an intelligent system for correlating and automating incident response, designed to enhance threat detection and management in modern cybersecurity environments. The document outlines the fundamentals of cybersecurity as well as the contribution of artificial intelligence to advanced security data analysis. It then introduces the concepts of SIEM, event correlation, and SOAR, highlighting their complementary roles in optimizing detection and response. The design section details the architecture of the developed mechanism, integrating a Large Language Model (LLM) responsible for automatically analyzing alerts, an ElasticStack SIEM (Elasticsearch, Kibana, Filebeat) for data collection, and a Cortex/TheHive SOAR for orchestrating responses. The system is also connected to a SentinelOne responder to automate actions such as host isolation. The final implementation demonstrates that combining AI, SIEM, and SOAR enables intelligent correlation, a significant reduction in alert noise, and a faster and more efficient automated response to security incidents. This work highlights the potential of intelligent approaches to strengthen the security posture of the Ousmane SECK Computing Center (CCOS) and opens up perspectives for future improvements of the system.



**Keywords:** Cybersecurity, SIEM, Correlation, SOAR, LLM, Information system

# Table des matières

Dédicaces ..... 2
Remerciements ..... 3
Résumé ..... 4
Abstract ..... 5
Introduction Générale ..... 12
CHAPITRE 1 : FONDEMENTS DE LA SECURITE RESEAU ET APPORTS DE L'INTELLIGENCE ARTIFICIELLE ..... 14
Introduction : ..... 14
I. Définition de la sécurité informatique ..... 14
II. Principales menaces et typologies des attaques informatiques ..... 15
II.1. Les attaques par déni de service (DOS/DDOS)…. ..... 16
II.2. Les attaques par hameçonnage(phishing) ..... 18
II.3. Les logiciels malveillants ..... 19
II.4. Les attaques de l'homme du milieu ..... 20
II.5. Attaque par injection SQL (Structured Query Language) ..... 23
II.6. Menaces Internes ..... 24
III. Approches classiques de sécurisation et limites ..... 26
III.1. Approches classiques de sécurisation ..... 26
III.2. Limites des approches traditionnelles ..... 32
IV. Introduction à l'intelligence artificielle ..... 32
IV.1. Le Machine Learning ..... 34
IV.2. Le Deep Learning ..... 35
IV.3. Le traitement du langage naturel ..... 35
IV.4. La vision par ordinateur ..... 36
V. Application de l'intelligence artificielle en cybersécurité ..... 37
V.1. Protection et authentification par mot de passe ..... 37
V.2. Détection du phishing et contrôle de la prévention ..... 38
V.3. Gestion des vulnérabilités ..... 38
V.4. Sécurité réseau ..... 39
V.5. Analyse comportementale ..... 39
Conclusion ..... 40
CHAPITRE 2 : NOTIONS DE SIEM, CORRELATION et SOAR ..... 41
Introduction ..... 41

I. Notions de SIEM (Security Information and Event Management) ..... 41
I.1 Définition et rôle ..... 41
I.2.Architecture et fonctionnement du SIEM ..... 42
I.3. Avantages et limites du SIEM ..... 43
II. Notions de Corrélation des événements de sécurité ..... 45
III. Automatisation et Orchestration de la réponse aux incidents ..... 48
III.3. Apports et bénéfices du SOAR ..... 48
III.4. Limites et défis du SOAR ..... 49
Conclusion ..... 51
CHAPITRE 3 : CONCEPTION DU SYSTEME INTELLIGENT ..... 52
Introduction ..... 52
I. Mise en place du SIEM ..... 52
II. Mise en place du SOAR(Thehive+Cortex) et connexion au SIEM ..... 58
II.1. Mise en place du SOAR (Thehive+Cortex) ..... 58
II.2. Connexion du SOAR au SIEM ..... 63
III. Développement du système intelligent de corrélation ..... 64
IV. Système complet de corrélation intelligente et d'automatisation de la réponse 69
Conclusion ..... 74
Conclusion Générale ..... 76
Bibliographie ..... 78

## Liste des Abréviations

**API** : Application Programming Interface

**CCTV**: Closed-Circuit Television

**CPU**: Central Processing Unit

**DDoS**: Distributed Denial of Service

**DNS**: Domain Name System

**DoS**: Deni of Service

**HTTP**: HyperText Transfer Protocol

**IA** : Intelligence Artificielle

**IP** : Internet Protocol

**EDR**: Endpoint Detection and Response

**LLM:** Large Language Model

**OSI** : Open System Interconnection

**RGPD** : Règlement Général sur la Protection des Données

**SIEM**: Security Information and Event Management

**SOAR:** Security Orchestration, Automation and Response

**SOC** : Security Operations Center

**SSI** : Sécurité des Systèmes d'Information

**SQL**: Structured Query Language

**SSL:** Secure Sockets Layer

**STIX**: Structured Threat Information eXpression

**TCP**: Transmission Control Protocol

**TLS**: Transport Layer Security

**TAXII**: Trusted Automated eXchange of Intelligence Information

**URL**: Uniform Resource Locator

**VPN**: Virtual Private Network

# Liste des figures

Figure 1: Attaque DDoS ..... 18
Figure 2: Attaque par hameçonnage ..... 19
Figure 3: Attaque de l'homme du milieu ..... 20
Figure 4: Injection SQL ..... 24
Figure 5: Les types de menaces internes ..... 26
Figure 6: Fonctionnement du pare-feu ..... 27
Figure 7: IDS/IPS ..... 28
Figure 8: Gestion des identités et des accès ..... 29
Figure 9 : Réseau Privé Virtuel ..... 30
Figure 10: Les types d'intelligence artificielle ..... 33
Figure 11: Architecture d'un SIEM ..... 43
Figure 12: Définition de l'interface de capture ..... 53
Figure 13: Définition du réseau surveillé ..... 53
Figure 14: Configuration de l'interface de sortie ..... 54
Figure 15: Chargement des règles IDS ..... 54
Figure 16: Définition du fichier de log à surveiller ..... 55
Figure 17: Activation des modules ..... 55
Figure 18: Envoie des logs vers Elasticsearch ..... 55
Figure 19: Paramètres réseaux ..... 56
Figure 20: Sécurisation d'elasticsearch ..... 56
Figure 21: Activation de la xpack ..... 57
Figure 22: Tableau de bord Kibana ..... 57
Figure 23: Architecture physique SIEM ..... 58
Figure 24: Configuration de thehive ..... 59
Figure 25: plateforme thehive ..... 60
Figure 26: Configuration du conteneur cortex ..... 61
Figure 27: plateforme Cortex ..... 61
Figure 28 : Configuration du conteneur elasticsearch dédiée à cortex ..... 62
Figure 29: Connexion thehive et cortex ..... 62
Figure 30: Création de l'utilisateur API ..... 63
Figure 31: Création du connecteur webhook ..... 64
Figure 32: Obtention de la clé d'accès sur Hugging Face ..... 65
Figure 33: Chargement du modèle TinyLlama ..... 66
Figure 34: Définition du prompt de corrélation ..... 66
Figure 35: Module de corrélation pour 2 alertes ..... 67
Figure 36: Module de corrélation pour un lots d'alertes ..... 67
Figure 37: Structure de l'API ..... 68
Figure 38: Interface FastAPI ..... 69
Figure 39: Configuration Thehive+Cortex ..... 70

Figure 40: Configuration d'elasticsearch ..... 70
Figure 41: Ajout du champs action ..... 71
Figure 42: Exécution du répondeur de cortex ..... 71
Figure 43 : Création d'un case thehive ..... 72
Figure 44: Récupération des alertes depuis Elasticsearch ..... 72
Figure 45 : Définition de la récupération périodique ..... 73
Figure 46: Système intelligent de corrélation et d'automatisation de la réponse ..... 74

## Introduction Générale

À l'ère de la transformation numérique, les systèmes d'information occupent une place essentielle dans le fonctionnement des entreprises, des institutions et des services publics. Cette dépendance accrue s'accompagne d'une multiplication des attaques informatiques, notamment les intrusions, les vols de données ou encore les attaques par déni de service. Les organisations doivent donc assurer une surveillance continue de leurs réseaux afin de détecter, d'analyser et de traiter efficacement les incidents de sécurité.

Traditionnellement, cette surveillance repose sur la collecte et l'analyse des logs et des alertes générées par différents systèmes de sécurité comme les IDS, les IPS, les pares-feux, les antivirus ou les SIEM. Cependant, ces outils produisent une quantité très importante de données souvent difficile à interpréter sans un mécanisme de corrélation performant.

Ce mémoire s'inscrit dans cette problématique et vise à développer une solution intelligente combinant la corrélation automatique des alertes et une réponse adaptative. L'objectif est de renforcer la posture de cybersécurité du Centre de Calcul Ousmane SECK en améliorant la capacité à comprendre et à traiter les incidents de manière rapide et pertinente.

Face à la complexité croissante des attaques et à la volumétrie importante de données issues des différents outils de sécurité, les centres de supervision de la sécurité (SOC) sont confrontés à un volume croissant d'alertes et de journaux d'événements. Dès lors, une question existentielle se pose ; Comment corréler efficacement les logs et alertes provenant de multiples sources pour identifier des incidents pertinents et automatiser la réponse à ces incidents pour réduire le temps de réaction et limiter les impacts sur le système d'information ?

L'objectif de ce travail est donc de développer un système intelligent capable d'automatiser la corrélation des alertes afin d'améliorer la détection et la réponse aux incidents de sécurité en concevant une architecture intégrant les sources de données (logs et alertes), mettant en œuvre des algorithmes de corrélation basés sur des techniques de Machine Learning et de règles de détection, automatiser certaines actions de réponse aux incidents (blocage d'adresse IP, isolement d'hôte, notification, …),

concevoir une interface de visualisation et de pilotage du système et évaluer les performances du système sur des jeux de données réels.

La méthodologie adoptée s'appuie sur une approche à la fois analytique et expérimentale, il s'agira donc d'étudier des outils SIEM existants, modèles de corrélation et solutions d'automatisation (SOAR), de définir une architecture technique, des flux de données et des composants (collecte, corrélation, réponse), d'implémenter des modules d'analyse, d'apprentissage automatique et d'automatisation.

Ce présent mémoire propose une implémentation d'un système intelligent de corrélation et d'automatisation de la réponse. Son apport réside dans la proposition d'une solution innovante pour améliorer la détection et la gestion des incidents, contribuant ainsi à renforcer la posture de cybersécurité du Centre de Calcul Ousmane SECK (CCOS), ainsi que l'exploration de l'intégration de l'intelligence artificielle dans la corrélation d'alertes, un domaine encore en pleine évolution.

Ce travail est structuré comme suit :

- Le premier chapitre se concentre sur les fondements de la cybersécurité et de l'apport de l'intelligence artificielle.
- Le deuxième chapitre s'intéresse aux notions de SIEM, corrélation et d'automatisation de la réponse aux incidents (SOAR).
- Le troisième chapitre présente la conception du système intelligent.

# CHAPITRE 1 : FONDEMENTS DE LA SECURITE RESEAU ET APPORTS DE L'INTELLIGENCE ARTIFICIELLE

**Introduction :**

Dans un contexte où le monde est de plus en plus connecté, les composants techniques qui assurent la communication et l'échange de données occupent une place essentielle au sein des systèmes d'information. Ils assurent la transmission, le partage et le traitement des données essentielles aux activités personnelles, professionnelles et institutionnelles. Cependant, cette interconnexion croissante s'accompagne d'une augmentation significative des risques de sécurité. Les cyberattaques, de plus en plus fréquentes et sophistiquées, menacent la confidentialité, l'intégrité et la disponibilité des ressources numériques. Face à ces défis, les approches classiques de défense telles que les pares-feux, VPN ou systèmes de détection d'intrusion bien qu'efficaces dans certains cas, montrent aujourd'hui leurs limites. C'est dans cette dynamique que s'inscrit l'émergence de l'intelligence artificielle (IA) comme levier innovant en matière de cybersécurité. Grâce à ses capacités d'apprentissage automatique, de reconnaissance de schémas complexes et d'automatisation des processus décisionnels, l'IA permet d'envisager de nouvelles approches de sécurité plus proactives, intelligentes et autonomes.

## I. Définition de la sécurité informatique

La sécurité informatique regroupe l'ensemble des moyens techniques, organisationnels, juridiques et humains destinés à protéger l'intégrité et la confidentialité des informations stockées dans un système informatique. La SSI, Sécurité des Systèmes d'Information, désigne l'ensemble des mesures nécessaires pour prévenir l'utilisation non autorisée, le mauvais usage, la modification ou le détournement d'un système d'information.

Selon le rapport annuel Data Breach Investigations Report publié par Verizon en 2023, la majorité des incidents de sécurité identifiés dans les organisations impliquent des atteintes à la confidentialité, à l'intégrité ou à la disponibilité des données, souvent en lien avec des erreurs humaines ou des compromissions d'identifiants. C'est dans cette perspective que la sécurité des systèmes d'informations vise les objectifs suivants :

- **La confidentialité :** fait référence au droit de contrôle sur la façon dont les informations et données à caractère personnel sont collectées, conservées et utilisées. Les politiques de confidentialité sont mises en place pour protéger les informations sensibles et l'identité des utilisateurs. Le respect de la confidentialité va de pair avec une utilisation responsable des informations sensibles. Les organisations sont tenues d'être transparentes quant aux types de données qu'elles ont l'intention de collecter, aux finalités de la collecte, au contexte et aux tiers avec lesquels ces informations seront partagées. L'utilisateur doit accepter au préalable ces conditions et a donc le droit de contrôler les informations qu'il partage. Seules les personnes autorisées peuvent avoir accès aux informations qui leurs sont destinées. Tout accès indésirable doit être empêché.
- **L'authenticité :** Permet d'assurer l'identité d'un utilisateur grâce à l'usage d'un code d'accès. Ici, l'objectif de la sécurité informatique est de garantir que chaque utilisateur est bien celui qu'il prétend être. Le contrôle d'accès, par un mot de passe par exemple, permet de limiter la consultation ou l'utilisation de certaines ressources seulement aux personnes autorisées et de maintenir la confiance dans les relations d'échange.
- **L'intégrité :** Consiste à garantir que les données soient celles que l'on attend, et ne soient pas altérées de façon fortuite, illicite ou malveillante. En clair, les éléments considérés doivent être exacts et complets. Cet objectif utilise généralement des méthodes de calcul de checksum ou de hachage.
- **La Disponibilité :** Elle définit l'accessibilité aux ressources du système d'information de manière permanente et sans faille durant les plages d'utilisation prévues. Les services et ressources doivent être accessibles rapidement et régulièrement.

## II. Principales menaces et typologies des attaques informatiques

Afin d'atteindre l'objectif qui est de voler des informations personnelles tels que les identifiants de connexion, les détails de compte et les numéros de carte crédit pour les utiliser à des fins multiples notamment l'usurpation d'identité, les transferts de fonds non autorisés ou la modification illicite de mots de passe, les systèmes informatiques

sont exposés à plusieurs types de menaces. Ces attaques sont multiples et chacune a sa méthodologie d'agression.

### II.1. Les attaques par déni de service (DOS/DDOS)

Une attaque par déni de service comme représenté dans la figure 1[3] est une forme d'attaque informatique ayant pour but de rendre indisponible un service, d'empêcher les utilisateurs légitimes d'un service de l'utiliser en saturant ce dernier de requêtes. Elle agit sur le CPU, la mémoire ou la bande passante. Une attaque DDoS est essentiellement l'utilisation légitime d'un service en ligne poussée à l'extrême. Par exemple, un site web peut être capable de traiter un certain nombre de requêtes par minute. Si ce nombre est dépassé, les performances du site web sont dégradées, ce qui peut rendre le système complètement inaccessible. Cette surcharge peut être due à une attaque ou même à une utilisation légitime, comme un site de commerce électronique submergé le jour du "Black Friday" ou une plateforme de vente de billets qui tombe en panne au moment de l'ouverture des ventes pour un événement populaire. Les attaques DDoS sont capables de submerger une cible à différents niveaux. Par exemple, une application Web peut avoir un nombre maximum de requêtes qu'elle peut traiter. Il se peut également que le serveur sur lequel il fonctionne ait une limite quant au nombre de connexions simultanées qu'il peut gérer. Un réseau d'entreprise est probablement soumis à des restrictions de bande passante qui pourraient être dépassées par un pirate. Le dépassement de l'un de ces seuils peut signifier une attaque DoS ou DDoS. Les attaques DoS & DDoS peuvent prendre de nombreuses formes et être utilisées à des fins diverses :

- **Attaques par amplification :** Certains services, comme le DNS, ont des réponses beaucoup plus volumineuses que la demande correspondante. Dans les attaques par amplification DDoS, les attaquants envoient une requête à un serveur DNS avec leur adresse IP usurpée à l'adresse IP de la cible, ce qui fait que la cible reçoit un grand nombre de réponses non sollicitées qui consomment des ressources.
- **Saturation de la bande passante :** Tous les réseaux ont une bande passante et un débit maximum qu'ils peuvent maintenir. Les attaques par saturation de la bande passante tentent de consommer cette bande passante avec du trafic illégitime.

- **Attaque au niveau de l'application :** Une attaque au niveau de l'application est un type d'attaque DDoS qui cible la couche 7 du modèle OSI. Un exemple de ce type d'attaque est Slowloris, dans laquelle l'attaquant envoie des requêtes HyperText Transfer Protocol (HTTP) partielles, mais ne les termine pas. Des en-têtes HTTP sont envoyés périodiquement pour chaque requête, ce qui entraîne l'immobilisation des ressources réseau. L'attaquant continue l'assaut jusqu'à ce qu'aucune nouvelle connexion ne puisse être établie par le serveur. Ce type d'attaque est très difficile à détecter, car plutôt que d'envoyer des paquets corrompus, il envoie des paquets partiels et utilise un volume faible de la bande passante.
- **Attaque de protocole :** Une attaque de protocole est un type d'attaque DDoS qui exploite les faiblesses des couches 3 et 4 du modèle OSI. Par exemple, l'attaquant peut exploiter la séquence de connexion TCP en envoyant des requêtes, mais en ne répondant pas comme prévu ou en répondant avec une autre requête avec utilisation d'une adresse IP source usurpée. Les requêtes sans réponse utilisent les ressources du réseau jusqu'à ce que ce dernier devienne indisponible.
- **Dégradation de service :** Les cybercriminels se livrent de plus en plus à ces types d'attaques qui consistent à envoyer à une cible un faible volume de trafic qui diminue l'accessibilité du service sans pour autant le mettre complètement indisponible. Ces attaques sont devenues populaires parce qu'elles sont plus difficiles à détecter pour une organisation qu'une attaque DDoS de grande envergure.

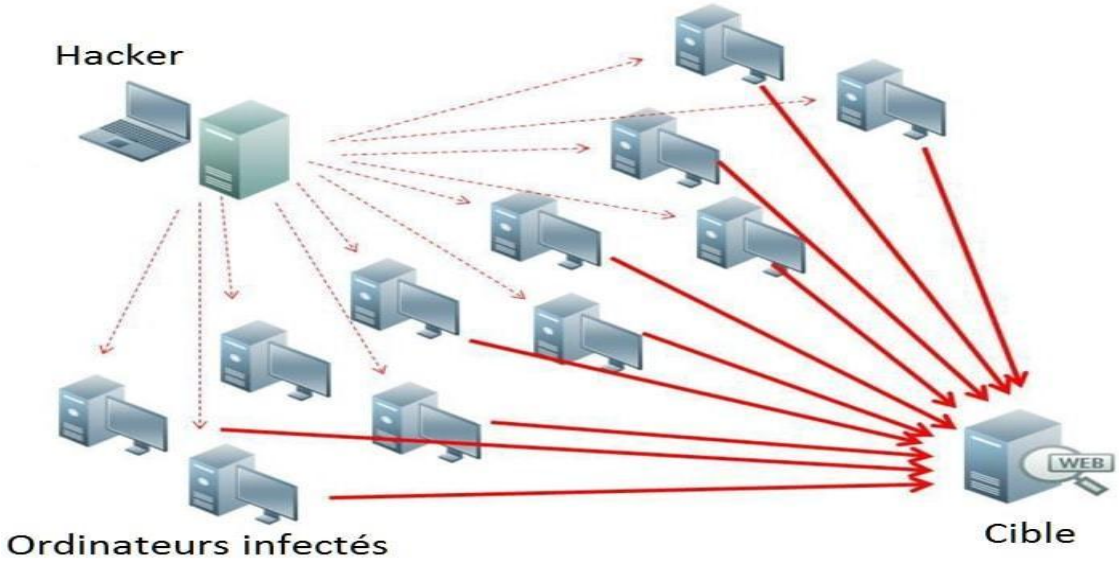



Figure 1: Attaque DDoS

**II.2. Les attaques par hameçonnage(phishing)**

Le phishing, ou hameçonnage en français, désigne un ensemble de techniques d'ingénierie sociale utilisé par les cybercriminels pour inciter les victimes à révéler volontairement des informations confidentielles. Il repose sur la duperie : l'attaquant se fait passer pour un tiers de confiance, généralement par le biais d'un email frauduleux, d'un site web contrefait, ou d'un message SMS ou WhatsApp, afin d'exploiter la crédulité ou l'urgence perçue par la cible. Typiquement, la victime reçoit un message l'invitant à cliquer sur un lien (vers un faux site web) ou à télécharger une pièce jointe infectée. Le faux site imite parfaitement un site légitime (comme une banque ou une plateforme de messagerie), incitant la victime à entrer ses identifiants. Ces données sont alors automatiquement transmises à l'attaquant, qui peut les utiliser pour accéder aux comptes de la victime, effectuer des transactions frauduleuses ou propager d'autres attaques. Plusieurs paramètres peuvent permettre de détecter l'attaque : adresse email de l'expéditeur suspecte ou légèrement modifiée, fautes d'orthographe ou de grammaire dans le message, lien hypertexte étrange (ne correspond pas au site officiel), demande urgente ou menaçante, pièce jointe inattendue ou douteuse. La figure 2 [4] ci-dessous représente un scenario de phishing.

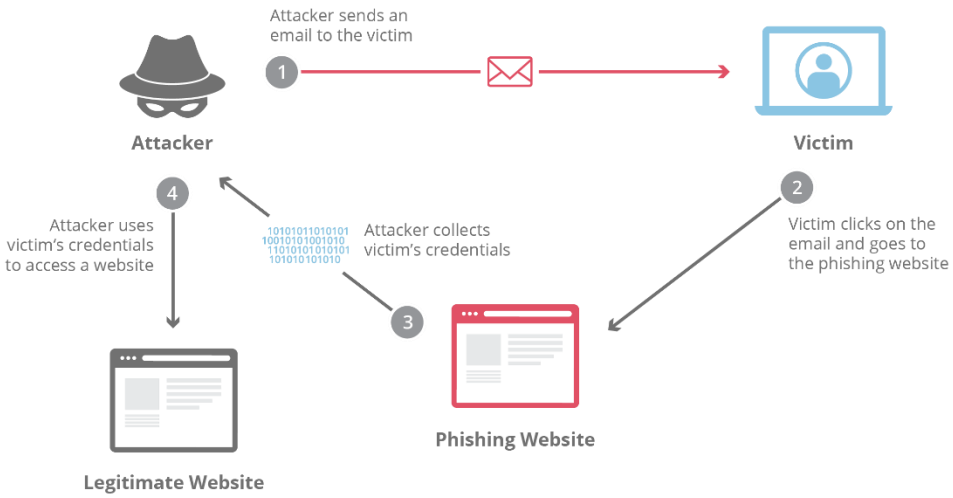



Figure 2: Attaque par hameçonnage

### II.3. Les logiciels malveillants

Les logiciels malveillants, ou malwares, regroupent un ensemble de programmes conçus pour infiltrer, endommager ou perturber un système informatique sans le consentement de son utilisateur. Leur mode d'action sur le SI varie en fonction de leur typologie (virus, vers, chevaux de Troie, ransomwares, spywares, etc.), mais repose sur des mécanismes communs d'intrusion et de propagation dont :

- **La phase d'infection :**

  Le malware pénètre dans un SI de différentes manières tels que des pièces jointes malveillants inclus dans des courriels, téléchargements piégés depuis des sites compromis, supports amovibles infectés (clés USB, disques durs externes), exploitation de failles dans les services réseau ou applications non mises à jour.

- **La phase d'exécution :**

  Une fois introduit, le malware s'exécute généralement de manière furtive. En effet, il modifie des fichiers systèmes ou du registre pour assurer une persistance après redémarrage, ouvre des ports non autorisés ou modifie des règles du pare-feu interne avant de créer des connexions sortantes vers un serveur de commande et de contrôle afin de recevoir des instructions.

- **Phase de propagation et d'attaque :**

  Selon le but de son attaque, le malware peut se propager latéralement au sein du SI en exploitant les partages de fichiers ou les vulnérabilités des protocoles internes (SMB, RDP, etc.), exfiltrer des données sensibles vers un serveur distant, chiffrer ou détruire des fichiers (cas des ransomwares), détourner les

ressources pour lancer d'autres attaques, telles que des campagnes de spam ou des attaques DDoS.

### II.4. Les attaques de l'homme du milieu

Une attaque de type man-in-the-middle (MITM) est une forme de cyberattaque dans laquelle les criminels exploitant des protocoles faibles s'insèrent entre les entités d'un canal de communication pour voler des données. En effet, aucune des parties envoyant des e-mails, des SMS ou des chats lors d'un appel vidéo ne sait qu'un attaquant a inséré sa présence dans la conversation et que l'attaquant vole ses données. Alors que la plupart des cyberattaques sont silencieuses et menées à l'insu des victimes, certaines attaques MITM sont le contraire. Il peut s'agir d'un bot générant des messages texte crédibles, de se faire passer pour la voix d'une personne lors d'un appel ou d'usurper un système de communication entier pour extraire des données que l'attaquant pense importantes des appareils des participants. La figure 3 [5] ici présente est une illustration d'une attaque de l'homme du milieu.

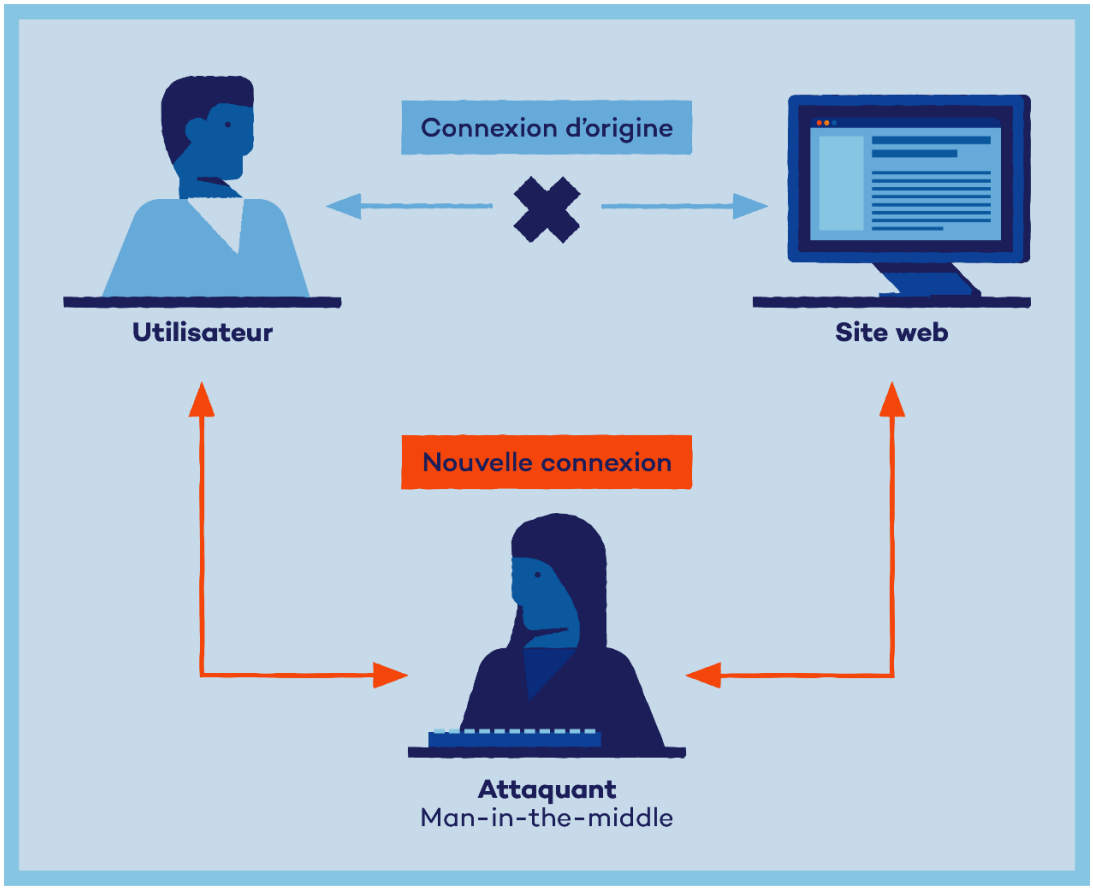



Figure 3: Attaque de l'homme du milieu

Il existe différents procédés d'attaques comme :

- **Détournement d'E-mails :** Comme son nom l'indique, dans ce type d'attaque, les cybercriminels prennent le contrôle des comptes de messagerie des banques, des institutions financières ou d'autres entreprises de confiance qui ont accès à des données sensibles et à de l'argent. Une fois à l'intérieur, les attaquants peuvent surveiller les transactions et la correspondance entre la banque et ses clients. Dans des scénarios plus malveillants, les assaillants usurpent ou falsifient l'adresse e-mail de la banque et envoient aux clients des e-mails leur demandant de renvoyer leurs identifiants, ou pire, d'envoyer de l'argent, à un compte contrôlé par les assaillants.
- **Ecoute WI-FI :** Dans l'écoute du Wi-Fi, les cybercriminels incitent les victimes à se connecter à un réseau sans fil voisin avec un nom légitime. Mais en réalité, le réseau est configuré pour s'engager dans une activité malveillante. Le réseau sans fil peut sembler appartenir à une entreprise voisine que l'utilisateur fréquente ou avoir un nom générique, apparemment inoffensif, tel que « Réseau Wi-Fi public gratuit ». Dans certains cas, l'utilisateur n'a même pas besoin de saisir un mot de passe pour se connecter. Une fois que les victimes sont connectées au Wi-Fi malveillant, l'assaillant dispose d'options : surveiller l'activité en ligne de l'utilisateur ou récupérer les identifiants de connexion, les informations de carte de crédit ou de paiement et d'autres données sensibles.
- **Usurpation DNS :** L'usurpation de système de noms de domaine ou l'empoisonnement du cache DNS se produit lorsque des enregistrements DNS manipulés sont utilisés pour détourner le trafic en ligne légitime vers un faux site Web ou un site Web usurpé conçu pour ressembler à un site Web que l'utilisateur connaîtrait et en qui il aurait probablement confiance. Comme pour toutes les techniques d'usurpation d'identité, les attaquants invitent les utilisateurs à se connecter involontairement au faux site Web et à les convaincre qu'ils doivent prendre une mesure spécifique, telle que payer des frais ou transférer de l'argent sur un compte spécifique. Les assaillants volent autant de données que possible aux victimes dans le processus.
- **Détournement de session :** Le détournement de session est un type d'attaque MITM dans laquelle l'assaillant attend qu'une victime se connecte à une application, par exemple pour la banque ou l'e-mail, puis vole le cookie de session. L'assaillant utilise ensuite le cookie pour se connecter au même compte

que celui de la victime, mais à la place à partir du navigateur de l'assaillant. Une session est un élément de données qui identifie un échange d'informations temporaire entre deux appareils ou entre un ordinateur et un utilisateur. Les attaquants exploitent les sessions parce qu'elles sont utilisées pour identifier un utilisateur qui s'est connecté à un site Web. Cependant, les attaquants doivent travailler rapidement à mesure que les sessions expirent après un certain temps, qui peut être aussi court que quelques minutes.

- **Détournement Secure Sockets Layer (SSL) :** Aujourd'hui, la plupart des sites Web affichent qu'ils utilisent un serveur sécurisé. Ils ont l'abréviation « HTTPS », c'est-à-dire HyperText Transfer Protocol Secure, au lieu de « HTTP » ou HyperText Transfer Protocol dans la première partie de l'URL (Uniform Resource Locator) qui apparaît dans la barre d'adresse du navigateur. Même lorsque les utilisateurs saisissent HTTP, ou pas du tout HTTP, le HTTPS ou la version sécurisée sera rendu dans la fenêtre du navigateur. Il s'agit d'un protocole de sécurité standard, et toutes les données partagées avec ce serveur sécurisé sont protégées. SSL et sa couche de sécurité de transport (TLS) apparue ultérieurement sont des protocoles pour établir la sécurité entre les ordinateurs en réseau. Dans un détournement SSL, l'attaquant intercepte toutes les données passant entre un serveur et l'ordinateur de l'utilisateur. Cela est possible, car le protocole SSL est un protocole de sécurité plus ancien et vulnérable qui a nécessité son remplacement, la version 3.0 a été obsolète en juin 2015, avec le protocole TLS plus fort.
- **Intoxication du cache ARP :** Le protocole de résolution d'adresse (ARP) est un protocole de communication utilisé pour découvrir l'adresse de la couche de liaison, telle qu'une adresse de contrôle d'accès multimédia (MAC), associée à une adresse de couche Internet donnée. L'ARP est important car il traduit l'adresse de la couche de liaison en adresse IP (Internet Protocol) sur le réseau local. Dans ce schéma d'attaque, l'ordinateur de la victime est induit en erreur par de fausses informations du cybercriminel pour penser que l'ordinateur du fraudeur est la passerelle réseau. Ainsi, l'ordinateur de la victime, une fois connecté au réseau, envoie essentiellement tout son trafic réseau à l'acteur malveillant plutôt que via la passerelle réseau réelle. L'assaillant utilise ensuite ce trafic détourné pour analyser et voler toutes les informations dont il a besoin,

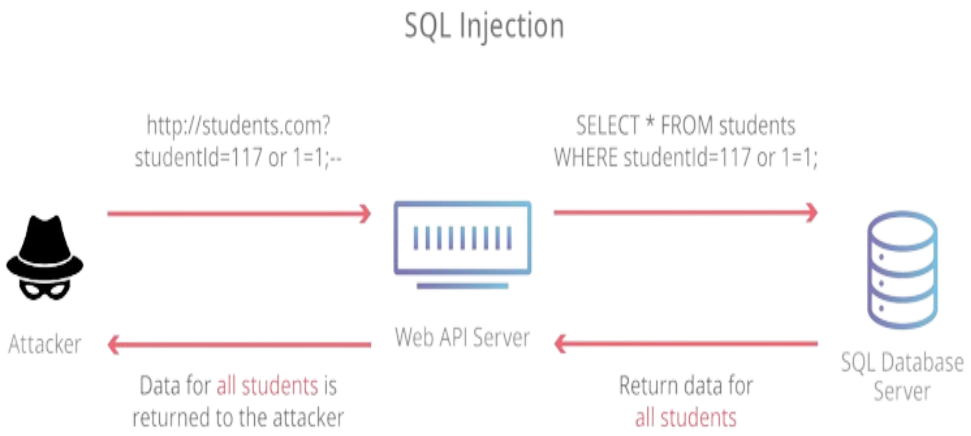
telles que les informations personnellement identifiables (PII) stockées dans le navigateur.

- **Usurpation d'adresse IP :** L'usurpation d'adresse IP est similaire à l'usurpation DNS, car l'attaquant redirige le trafic destiné à un site légitime vers un site frauduleux. Au lieu d'usurper l'enregistrement DNS du site Web, l'assaillant modifie l'adresse IP du site malveillant pour qu'elle apparaisse comme s'il s'agissait de l'adresse IP des utilisateurs légitimes du site Web destinés à la visite.
- **Vols des cookies de navigateur :** En informatique, un cookie est un petit élément d'information stocké. Un cookie de navigateur, également connu sous le nom de cookie HTTP, est constitué de données collectées par un navigateur Web et stockées localement sur l'ordinateur d'un utilisateur. Le cookie de navigateur aide les sites Web à mémoriser des informations pour améliorer l'expérience de navigation de l'utilisateur. Par exemple, lorsque les cookies sont activés, un utilisateur n'a pas besoin de continuer à remplir les mêmes éléments sur un formulaire, tels que le prénom et le nom de famille. Le vol de cookies de navigateur doit être associé à une autre technique d'attaque MITM, telle que l'écoute Wi-Fi ou le détournement de session, pour être effectué. Les cybercriminels peuvent accéder à l'appareil d'un utilisateur à l'aide de l'une des autres techniques MITM pour voler des cookies de navigateur et exploiter tout le potentiel d'une attaque MITM. Avec l'accès aux cookies de navigateur, les attaquants peuvent accéder aux mots de passe, aux numéros de carte de crédit et à d'autres informations sensibles que les utilisateurs stockent régulièrement dans leurs navigateurs.

### II.5. Attaque par injection SQL (Structured Query Language)

L'injection SQL comme le montre la figure 4 [6] est une méthode courante utilisée pour attaquer les sites Web qui utilisent des bases de données. Une attaque de type SQL Injection utilise une requête SQL envoyée par le client au serveur. Cette dernière est insérée, ou « injectée », dans un plan de données à la place d'autre chose qui y est normalement insérée, tel qu'un mot de passe ou un identifiant. Le serveur qui héberge la base de données exécute ensuite la commande et le système est pénétré. Si une injection SQL réussit, plusieurs choses peuvent se produire, y compris la divulgation de données sensibles, la modification ou la suppression de données importantes. De

plus, un assaillant peut exécuter des opérations d'administrateur comme une commande d'arrêt, qui peut interrompre la fonction de la base de données.




[1]

Figure 4: Injection SQL

### II.6. Menaces Internes

Une menace interne est un type de cyberattaque provenant d'une personne qui travaille pour une organisation ou qui a un accès autorisé à ses réseaux ou systèmes. Une menace interne peut être un employé, un consultant, un membre du conseil d'administration ou un partenaire commercial actuel ou ancien et peut être intentionnelle, involontaire ou malveillante. Généralement, une menace interne en cybersécurité fait référence à une personne qui utilise son accès autorisé aux données et ressources d'une organisation pour nuire aux équipements, aux informations, aux réseaux et aux systèmes de l'entreprise. Cela inclut la corruption, l'espionnage, la dégradation des ressources, le sabotage, le terrorisme et la divulgation d'informations non autorisées. Il peut également s'agir d'un point de départ pour les cybercriminels pour lancer des logiciels malveillants ou des attaques de ransomware. Différents types de menaces internes (figure 5 [7]), peuvent conduire une organisation à subir une perte de données ou d'autres exploits de sécurité. Celles-ci incluent les éléments suivants :

- **Intentionnelle :** Une menace interne intentionnelle se produit lorsqu'une personne cherche à causer intentionnellement des dommages à une organisation. De nombreuses menaces internes intentionnelles visent à atteindre

[1]

même une entreprise en raison d'un manque de reconnaissance ou d'un manquement à répondre aux attentes, par exemple en ne recevant pas la prime ou la promotion souhaitée.

- **Non intentionnel :** Une menace interne involontaire implique la perte ou le vol de données à la suite d'une erreur ou d'une négligence de l'employé. Des menaces internes accidentelles et involontaires se produisent en raison d'erreurs humaines et d'individus qui commettent une erreur entraînant une fuite de données, une attaque de sécurité ou le vol d'identifiants. Les fuites accidentelles de données comprennent l'envoi d'informations commerciales à la mauvaise adresse e-mail, le fait de cliquer par erreur sur des hyperliens malveillants ou d'ouvrir des pièces jointes malveillantes en envoyant des e-mails par phishing, ou le fait de ne pas supprimer ou éliminer efficacement des informations sensibles. Ces menaces peuvent souvent être évitées en suivant les meilleures pratiques de sécurité. Ce genre d'incident se produit par une négligence qui conduit à exposer une organisation à une menace. Par exemple, ignorer les politiques de sécurité informatique, égarer les dispositifs de stockage portables, utiliser des mots de passe faibles et ignorer les mises à jour logicielles ou les correctifs de sécurité peut rendre les organisations vulnérables à une cyberattaque.
- **Menaces tierces :** Une menace tierce est généralement un partenaire commercial ou un sous-traitant qui compromet la sécurité d'une organisation. Les menaces tierces peuvent être le résultat d'une activité par négligence ou malveillance.
- **Menaces malveillantes :** C'est une forme de menace interne intentionnelle qui a l'intention de causer des dommages, que ce soit à des fins personnelles ou en tant qu'acte de vengeance. Les menaces internes malveillantes visent à divulguer des données sensibles, à harceler les directeurs de l'entreprise, à saboter les équipements et systèmes de l'entreprise ou à voler des données pour essayer de faire progresser leur carrière. Nombre de ces menaces malveillantes sont motivées financièrement, car les employés volent des données d'entreprise pour les vendre à des hackers, des organisations tierces ou des entreprises concurrentes.

- **Menaces collusoires :** Il s'agit d'un type d'initié malveillant, dans lequel une ou plusieurs personnes malveillantes travaillent avec un partenaire externe pour compromettre leur organisation. Les menaces internes collusoires impliquent souvent qu'un cybercriminel recrute un employé pour voler la propriété intellectuelle en son nom à des fins financières.

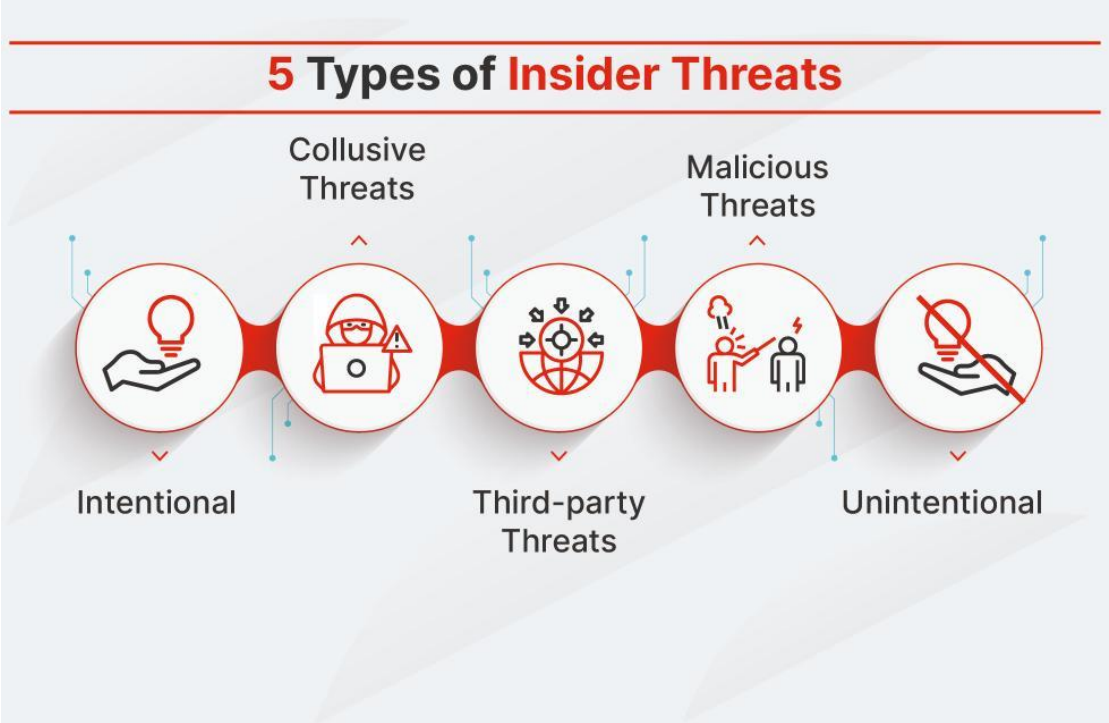



Figure 5: Les types de menaces internes

## III. Approches classiques de sécurisation et limites

### III.1. Approches classiques de sécurisation

#### III.1.1. Utilisations des pares-feux

Un pare-feu est un dispositif de sécurité réseau conçu pour bloquer les accès non autorisés, autoriser les communications légitimes et filtrer le trafic réseau entrant et sortant en fonction de règles de sécurité préméditées. L'objectif principal d'un pare-feu est d'établir une barrière entre un réseau interne de confiance et un réseau externe non fiable (figure 6 [8]) notamment internet. En effet, le fonctionnement des pares-feux dépend d'une multitude de facteurs :

- **Le filtrage par adresse IP :** le pare-feu analyse les adresses IP source et destination, s'ils correspondent à une règle autorisée, le trafic peut passer.
- **Le filtrage par port :** certains services utilisent des ports spécifiques. Par exemple, le port 80 pour le HTTP. Le pare-feu peut bloquer ou autoriser le trafic sur le port.

- **Le filtrage par contenu :** Le pare-feu peut inspecter le contenu des paquets pour s'assurer qu'ils ne contiennent pas de menaces.

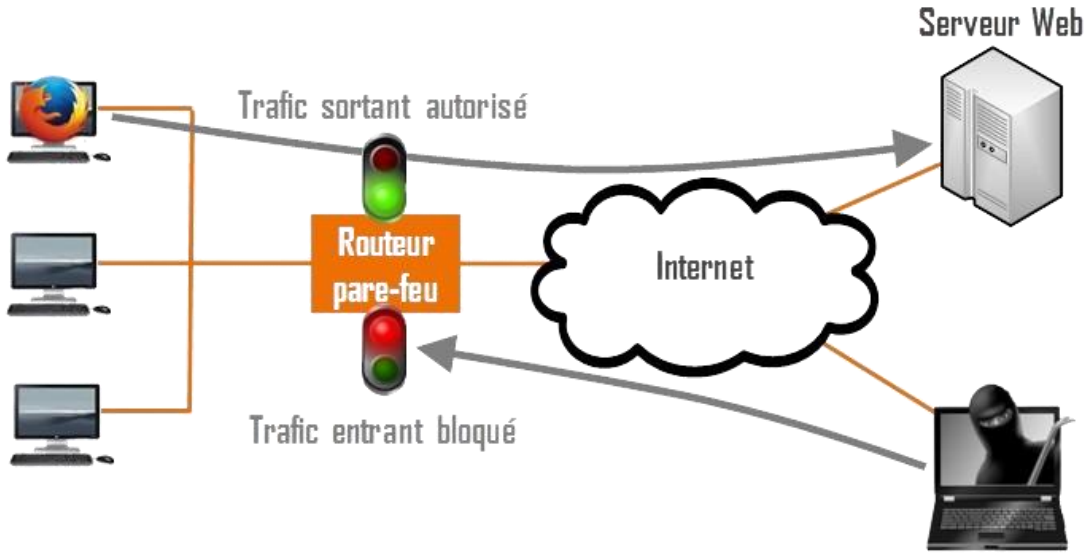



Figure 6: Fonctionnement du pare-feu

Alors que les pare-feux désignent la pierre angulaire de la défense du réseau, les menaces avancées nécessitent des mesures de sécurité supplémentaires.

### III.1.2. Les systèmes de détection d'intrusion (IDS/IPS)

Ils représentent des mécanismes de sécurité qui permettent de surveiller, analyser et réagir aux activités suspectes ou malveillantes dans un système informatique ou un réseau. En effet, les IDS (Intrusion Détection System) sont passives, elles ne font que détecter les tentatives d'intrusion, d'attaques ou d'anomalies. Dès qu'ils observent une activité, ils le comparent avec une base de signatures comme les attaques connues tels que les virus, malwares, scans de port ou avec des modèles comportementaux, anomalies par rapport au trafic habituel. Lorsqu'une activité suspecte est détectée, l'IDS envoie une alerte à l'administrateur. Les IPS (Intrusion Prevention System), de leurs atouts actifs, détectent et bloquent les intrusions en temps réel. Ils analysent d'abord le trafic, s'ils constatent un trafic malveillant, l'IPS bloque le paquet suspect, réinitialise la connexion et mets en quarantaine l'hôte attaquant. La figure 7 [9] représente une illustration comparative d'un IDS et d'un IPS.

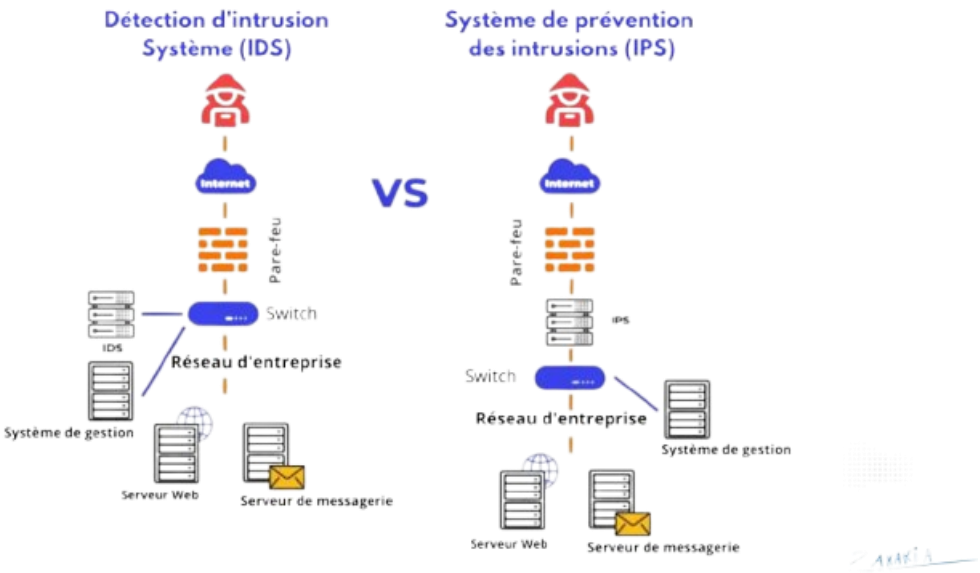



Figure 7: IDS/IPS

### III.1.3. La gestion des identités et des accès

La gestion des identités et des accès (figure 8 [10]), en anglais IAM (Identity Access Management) est une branche de la cybersécurité qui porte sur le provisionnement et la protection des identités numériques et des droits d'accès des utilisateurs dans un système informatique. Ses outils permettent de s'assurer que les bonnes personnes accèdent aux bonnes ressources pour des raisons saines et au bon moment. La mise en œuvre des IAM repose sur quatre piliers :

- **Administration des identités :** également appelé « gestion du cycle de vie des identités » consiste à créer, à gérer, et à supprimer en toute sécurité les identités des utilisateurs dans un système. Afin de faciliter et sécuriser l'accès utilisateur, les entreprises attribuent à chaque utilisateur humain et non humain une identité numérique distincte généralement stockées dans une base de données ou un annuaire central servant de source d'information unique. Ainsi, le système IAM utilisera ces informations pour valider les utilisateurs et déterminer ce qu'ils sont autorisés à faire.

- **Authentification :** c'est le processus qui permet de vérifier que l'utilisateur est bien la personne qu'il prétend être. Lorsqu'un utilisateur se connecte à un système ou demande l'accès à une ressource, il fournit des identifiants, appelés « facteurs d'authentification », pour prouver son identité. Par exemple, un utilisateur humain fournira son mot de passe ou le scan de son empreinte digitale biométrique, tandis qu'un utilisateur non humain partagera un certificat

numérique. Le système IAM vérifie ces identifiants dans la base de données centrale. Si elles correspondent, l'accès est accordé.

- **Autorisation :** C'est le processus qui permet de vérifier que les utilisateurs disposent des niveaux d'accès appropriés à une ressource. Elle est étroitement liée à l'authentification, celle-ci étant généralement une condition préalable à l'autorisation. Une fois que l'utilisateur a prouvé son identité, le système IAM vérifie les droits d'accès associés à cette identité dans la base de données centrale pour déterminer si oui ou non l'autoriser.
- **Audit :** qui consiste à s'assurer que le système IAM et ses composants (administration, authentification, autorisation) fonctionnent correctement. Elle se concentre sur le suivi et l'enregistrement de la manière dont les utilisateurs se servent de leurs droits d'accès afin d'empêcher tout accès non autorisé, y compris celui des pirates informatiques et de s'assurer que les utilisateurs autorisés n'abusent pas de leurs privilèges.




Figure 8: Gestion des identités et des accès

### III.1.4. Les Réseaux Privés Virtuels (VPN)

Un VPN (figure 9 [11]) est une technologie agissant comme un tunnel chiffré qui masque les données échangées et les adresses IP réelles des utilisateurs. Il permet d'établir une connexion sécurisée entre un utilisateur et un réseau distant en passant par internet ou un autre réseau public (non sécurisé). Son mécanisme repose sur plusieurs éléments techniques :

- **Encapsulation :** les données d'origine sont encapsulées dans de nouveaux paquets via un protocoles de tunneling (IPSec, SSL/TLS), ce qui permettra de transporter des informations privées sur un réseaux publics.
- **Chiffrement et intégrité :** IL s'établit d'abord un canal sécurisé, le client et le serveur utilisent des algorithmes de chiffrement asymétrique comme RSA pour s'authentifier mutuellement et établir ce canal ; Ensuite, ils génèrent une clé secrète symétrique qui chiffrera les données segmentées en paquets. Pour assurer l'intégrité, un protocole de hachage est utilisé, le HMAC (Hash-based Message Authentication Code).
- **Authentification :** afin d'assurer la connexion sécurisée, le VPN utilise une phase d'authentification, cela garantit que seul un utilisateur autorisé peut accéder au réseau privé.
- **Attribution d'adresse IP :** Le serveur VPN attribut au client une adresse IP virtuelle appartenant au réseau privé. Ainsi, tout le trafic du client semble provenir du serveur VPN et non de son fournisseur d'accès.
- **Acheminement du trafic :** Une fois le tunnel établit, tout le trafic de l'utilisateur est redirigé à travers le tunnel, le serveur VPN déchiffre les paquets et les transmets vers leur destination finale. Les réponses passent par le même chemin inverse.

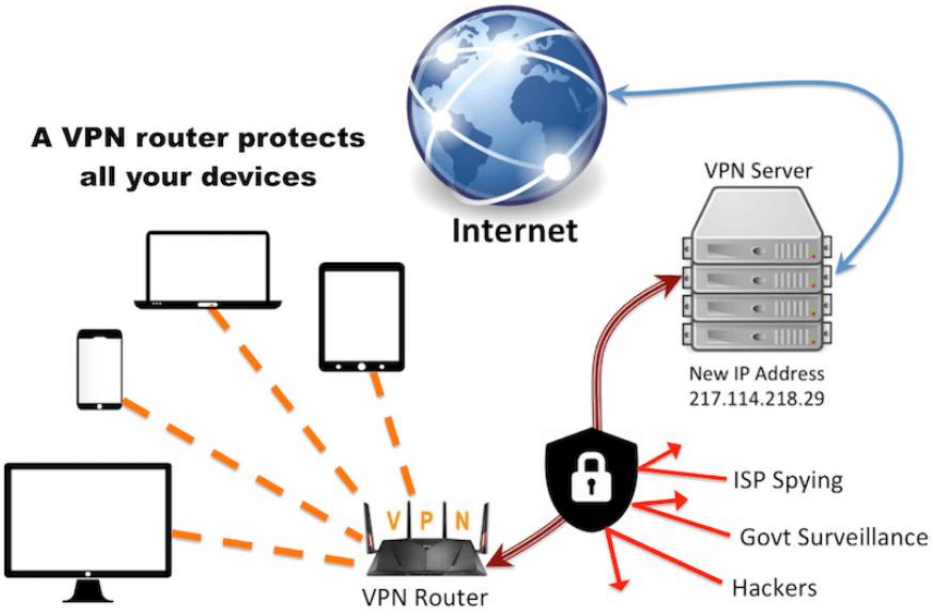



Figure 9 : Réseau Privé Virtuel

### III.1.5. Mise en place d'antivirus et antimalwares

La mise en place d'antivirus/antimalwares dans une infrastructure réseau est une combinaison de protection au niveau des postes, des serveurs et du réseau, gérée de manière centralisée et intégrée aux autres mécanismes de cybersécurité (pare-feu, IDS/IPS). Dans un premier temps des logiciels de protection sont installés aux niveaux des terminaux et des serveurs. Ils assurent une analyse en temps réel des fichiers et des processus afin de bloquer toute activité suspecte. Ensuite, la protection est étendue au niveau du réseau. Les pare-feux, les passerelles de messagerie et les proxies filtrants intègrent des moteurs antimalwares capables d'analyser le trafic, de bloquer les fichiers infectés et de filtrer les sites malveillants. Ces dispositifs permettent d'empêcher la propagation d'attaques avant même qu'elles n'atteignent les utilisateurs. L'ensemble est supervisé par une console de gestion centralisée, qui assure le déploiement des politiques de sécurité, le suivi de l'état des équipements et la détection des incidents.

### III.1.6. Les contrôles de sécurité physique

Ils ont pour objectif de protéger les équipements, les locaux et les ressources critiques contre les menaces non logicielles telles que l'intrusion, le sabotage, le vol, l'incendie ou encore les catastrophes naturelles. L'accès aux salles techniques et aux centres de données est strictement réservé au personnel autorisé. Pour cela, plusieurs technologies sont déployées : badges d'accès, dispositifs biométriques (empreinte digitale, reconnaissance faciale ou de l'iris) et systèmes de vidéosurveillance (CCTV), permettant d'assurer la traçabilité et l'audit des mouvements.

Des registres ou journaux d'accès sont tenus afin de garantir une traçabilité complète. Les équipements critiques tels que les serveurs, routeurs et commutateurs sont généralement placés dans des armoires ou racks verrouillés pour prévenir toute manipulation non autorisée. Les câbles sont installés dans des chemins sécurisés et correctement étiquetés, limitant les risques de branchements illicites. En complément, des mécanismes de redondance et des systèmes de sauvegarde renforcent la résilience en cas de panne ou de tentative de sabotage.

La protection physique inclut également la gestion des risques environnementaux. Des systèmes de détection et d'extinction d'incendie utilisant des gaz inertes plutôt que de l'eau protègent l'intégrité des infrastructures. La régulation de la température et de l'humidité est assurée par des systèmes de climatisation de précision, évitant

surchauffes et courts-circuits. Des détecteurs d'inondation ainsi que des dispositifs de protection électrique (onduleurs, parafoudres) viennent compléter l'ensemble.

Globalement, les contrôles de sécurité physique dans les infrastructures réseau reposent sur trois piliers, la maîtrise des accès humains, la protection des équipements et la sécurité environnementale.

La combinaison de ces mesures constitue une défense en profondeur essentielle à la fiabilité et à la résilience des systèmes d'information modernes. Cependant, même si ces dispositifs restent efficaces dans certains contextes, ils présentent plusieurs limites face à la sophistication croissante des menaces.

### III.2. Limites des approches traditionnelles

Les méthodes traditionnelles de sécurisation, bien qu'importantes dans une approche de défense en profondeur, montrent rapidement leurs limites face aux menaces actuelles, qui se caractérisent par leur complexité, leur évolution rapide et leur capacité à contourner les mécanismes classiques. Les antivirus et les systèmes de détection d'intrusion (IDS) s'appuient principalement sur des bases de signatures préétablies. Cette approche ne permet pas d'identifier efficacement les attaques exploitant des vulnérabilités encore inconnues. Même si des mises à jour régulières sont nécessaires, les attaquants profitent souvent du délai entre la découverte d'une menace et l'ajout de sa signature dans les bases de données. De plus, les mécanismes traditionnels reposent sur des règles statiques, généralement configurées manuellement par les administrateurs. Face à des attaques polymorphes ou en constante évolution, ces systèmes manquent de souplesse et exigent des ajustements fréquents pour rester efficaces. La gestion de ces outils demande également une expertise particulière et une surveillance continue. Les administrateurs sont souvent confrontés à un volume important d'alertes, ce qui entraîne une fatigue opérationnelle et augmente le risque d'erreurs humaines. Ces limites expliquent l'intérêt grandissant pour l'intégration de l'intelligence artificielle, qui permet d'apporter une analyse plus dynamique, prédictive et adaptative des comportements et des flux réseau.

## IV. Introduction à l'intelligence artificielle

L'intelligence artificielle n'est pas un concept nouveau. La paternité du terme est attribuée au mathématicien John McCarthy, chercheur au Massachusetts Institute of Technology (MIT). Il a introduit officiellement cette notion en 1956 lors d'un colloque

scientifique organisé sur le campus de l'Université de Dartmouth, aux États-Unis. Un de ses collègues, Marvin Minsky, alors professeur à l'Université Carnegie Mellon, a proposé l'une des premières définitions structurées de l'intelligence artificielle. Minsky, inventeur du SNARC, considéré comme le premier simulateur de réseau neuronal, définit l'IA comme « la construction de programmes informatiques capables d'effectuer des tâches qui, pour l'instant, sont mieux réalisées par des êtres humains, car elles requièrent des processus mentaux de haut niveau tels que l'apprentissage perceptuel, l'organisation de la mémoire ou encore le raisonnement critique ». Les principales branches associées à l'IA incluent le machine learning, l'apprentissage profond (deep learning), le traitement automatique du langage naturel (NLP) ainsi que la vision par ordinateur. L'émergence du machine learning dans les années 1990, puis celle du deep learning dans les années 2010, a marqué un tournant décisif. L'augmentation de la puissance de calcul, le développement des infrastructures de stockage et la disponibilité de grandes quantités de données (Big Data) ont permis à l'IA de réaliser des avancées majeures, notamment dans la reconnaissance vocale, la traduction automatique et la détection automatisée de menaces. La figure 10 [12] suivante présente des modèles d'intelligence artificielle.

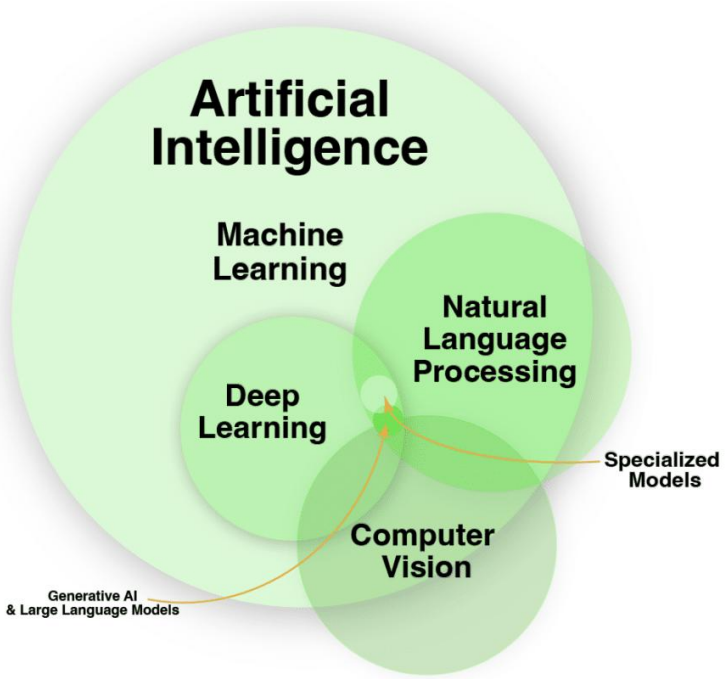



Figure 10: Les types d'intelligence artificielle

### IV.1. Le Machine Learning

Le Machine Learning (ou apprentissage automatique en français) est un domaine de l'intelligence artificielle qui s'appuie sur des méthodes mathématiques et statistiques afin de permettre aux ordinateurs « d'apprendre » à partir de données. Autrement dit, il leur permet d'améliorer leurs performances sur une tâche donnée sans être explicitement programmés pour chaque situation. Plus largement, ce champ d'étude concerne la conception, l'analyse, l'optimisation et le déploiement de telles méthodes. On parle d'apprentissage statistique car l'objectif est de construire un modèle dont l'erreur moyenne est la plus faible possible sur l'ensemble des données. De manière générale, l'apprentissage automatique comporte deux grandes phases :

#### IV.1.1. La phase d'apprentissage (ou d'entraînement)

Un modèle est estimé à partir d'un ensemble fini d'observations disponibles lors de la conception du système. Cette estimation vise à résoudre une tâche pratique comme la traduction automatique, l'estimation d'une densité de probabilité, la reconnaissance d'images (par exemple détecter la présence d'un chat) ou l'aide à la conduite d'un véhicule autonome.

#### IV.1.2. La phase de mise en production

Une fois le modèle entraîné, il peut être utilisé pour traiter de nouvelles données et produire des prédictions. Certains systèmes continuent même à apprendre après leur mise en production lorsque des retours sur leurs performances sont disponibles : on parle alors d'apprentissage en ligne ou d'apprentissage continu. Selon la nature des données utilisées pour l'apprentissage, plusieurs catégories de méthodes peuvent être distinguées :

- **Apprentissage supervisé** : l'algorithme apprend à partir de données étiquetées, pour lesquelles la sortie attendue est connue. L'objectif est de prédire correctement les sorties de nouvelles données.
- **Apprentissage non supervisé** : l'algorithme travaille sur des données non étiquetées et cherche à découvrir des structures cachées, comme des regroupements ou des motifs permettant de classifier ou d'organiser les données.

- **Apprentissage semi-supervisé** : cette approche combine une faible quantité de données étiquetées avec une grande quantité de données non étiquetées, ce qui permet d'améliorer la performance tout en réduisant les coûts d'annotation.
- **Apprentissage auto-supervisé** : forme particulière d'apprentissage non supervisé dans laquelle le modèle génère lui-même ses étiquettes à partir des données brutes. Il apprend ainsi des représentations internes utiles sans nécessiter d'annotations humaines.

Les méthodes d'apprentissage automatique peuvent s'appliquer à une grande variété de types de données : graphes, arbres, séries temporelles, ou encore vecteurs de caractéristiques composés de variables qualitatives ou quantitatives.

### IV.2. Le Deep Learning

Le Deep Learning (ou apprentissage en profondeur) est une sous-discipline du Machine Learning qui repose sur l'utilisation de réseaux de neurones artificiels constitués de nombreuses couches successives. En empilant des dizaines, voire des centaines de couches, ces modèles gagnent en capacité d'abstraction et peuvent apprendre des représentations extrêmement complexes des données. L'idée centrale du Deep Learning est de permettre à des algorithmes de reproduire certains mécanismes du cerveau humain grâce à des réseaux de neurones artificiels. Chaque couche du réseau reçoit, transforme puis transmet l'information à la couche suivante, créant ainsi un processus d'apprentissage hiérarchique où les caractéristiques simples sont progressivement combinées pour former des concepts plus élaborés. L'apprentissage en profondeur a permis des avancées majeures dans plusieurs domaines, notamment la vision par ordinateur, l'analyse d'images, la reconnaissance vocale et le traitement automatique du langage. Ces progrès rapides sont principalement dus à l'augmentation de la puissance de calcul, à la disponibilité de grandes quantités de données et à l'amélioration des architectures de réseaux neuronaux.

### IV.3. Le traitement du langage naturel

Le traitement automatique du langage naturel (ou Natural Language Processing, NLP) désigne l'ensemble des techniques permettant à une machine de comprendre, d'interpréter et de générer du langage humain. Il constitue aujourd'hui un domaine central de l'intelligence artificielle, notamment grâce à ses applications en traduction automatique, en analyse de sentiments, en résumé automatique ou encore dans les

assistants conversationnels. Le fonctionnement du NLP peut être décrit en quatre grandes étapes :

**Prétraitement** : le texte est nettoyé, normalisé et découpé en unités linguistiques (tokenisation). Il est ensuite transformé en données numériques exploitables par un modèle.

**Représentation** : les mots ou phrases sont convertis en vecteurs capables de capturer leur sens et leur contexte (word embeddings, modèles basés sur des Transformers, etc.).
**Analyse** : le modèle identifie la structure grammaticale et sémantique du texte, puis réalise diverses tâches telles que la classification, la traduction, l'extraction d'entités nommées ou l'analyse de sentiments.

**Génération** : la machine produit du texte en langage naturel, sous forme de réponses, résumés, conversations ou traductions.

En résumé, le NLP transforme le langage en données numériques, l'analyse pour en extraire le sens, puis génère à son tour un langage intelligible et cohérent.

### IV.4. La vision par ordinateur

La Computer Vision (ou vision par ordinateur) permet à une machine de « voir », c'est-à-dire d'analyser et d'interpréter des images ou des séquences vidéo. Elle est largement utilisée dans divers domaines, notamment en cybersécurité, par exemple pour détecter des logiciels malveillants à partir des images intégrées dans des courriels de phishing. La vision par ordinateur suit un processus structuré destiné à transformer des données visuelles brutes en informations interprétables. Elle commence par l'acquisition d'images ou de vidéos, qui sont ensuite soumises à un prétraitement visant à améliorer leur qualité et à les rendre exploitables numériquement (réduction du bruit, normalisation, correction du contraste, etc.). L'étape suivante consiste en l'extraction de caractéristiques visuelles pertinentes, telles que les contours, les textures ou les formes. Ces caractéristiques sont ensuite analysées par des modèles d'apprentissage automatique, en particulier les réseaux de neurones convolutifs (CNN). Ces réseaux fonctionnent en couches successives : les premières détectent des motifs simples, puis les couches plus profondes identifient des structures complexes jusqu'à reconnaître des objets ou des scènes entières. Enfin, le système procède à l'interprétation, qui peut prendre différentes formes : classification d'images, détection et localisation d'objets, segmentation sémantique ou reconnaissance faciale.

## V. Application de l'intelligence artificielle en cybersécurité

La cybersécurité, qui repose sur la mise en œuvre coordonnée de solutions matérielles, logicielles et organisationnelles pour protéger les réseaux et les systèmes contre les cybermenaces, connaît aujourd'hui une profonde transformation grâce à l'intelligence artificielle. L'IA améliore la détection des menaces, automatise la réponse aux incidents, renforce les mécanismes d'authentification et permet une analyse prédictive des risques. Le rôle de l'IA dans le renforcement de la sécurité est multidimensionnel. Elle peut automatiser des tâches répétitives telles que l'analyse des journaux ou l'évaluation des vulnérabilités, libérant ainsi les analystes pour des activités plus complexes et à forte valeur ajoutée. L'IA occupe également une place centrale dans la détection des menaces : les systèmes qui en tirent parti sont capables d'identifier des comportements anormaux en temps réel, ce qui facilite une réponse rapide et une neutralisation efficace des attaques. De plus, l'intelligence artificielle possède la capacité de s'adapter et d'évoluer en continu. En apprenant à partir de nouvelles données, elle améliore progressivement sa précision et sa capacité à anticiper ou contrer les menaces émergentes. Plusieurs approches d'application de l'IA en cybersécurité peuvent ainsi être distinguées.

### V.1. Protection et authentification par mot de passe

Grâce à l'utilisation de l'intelligence artificielle en cybersécurité, les organisations peuvent mieux protéger les mots de passe et sécuriser les comptes utilisateurs grâce à des mécanismes d'authentification renforcés. La plupart des sites Web intègrent désormais des espaces de connexion ou des formulaires permettant aux utilisateurs de saisir des informations sensibles. Ces interactions exigent des niveaux de sécurité supplémentaires afin d'empêcher que ces données ne soient compromises par des acteurs malveillants. Les outils basés sur l'IA, tels que les systèmes CAPTCHA avancés, la reconnaissance faciale ou encore les scanners d'empreintes digitales, permettent de vérifier automatiquement si une tentative d'accès provient d'un utilisateur légitime. Ces solutions contribuent à prévenir des techniques de cybercriminalité comme les attaques par force brute ou le piratage d'identifiants, qui pourraient compromettre l'ensemble du système d'information d'une organisation.

### V.2. Détection du phishing et contrôle de la prévention

Le phishing demeure l'une des menaces de cybersécurité les plus répandues et les plus dangereuses pour les organisations, quel que soit leur secteur d'activité. L'intégration de l'intelligence artificielle dans les solutions de sécurité de la messagerie électronique permet de détecter plus efficacement les anomalies et les indicateurs de messages malveillants. L'IA analyse à la fois le contenu et le contexte des courriels afin de déterminer rapidement s'il s'agit de messages indésirables, de tentatives de phishing ou de communications légitimes. Par exemple, elle peut identifier en quelques instants des signes typiques de phishing tels que l'usurpation d'adresse e-mail, la falsification de l'expéditeur ou encore des noms de domaine volontairement mal orthographiés. Les techniques d'apprentissage automatique permettent à l'IA d'améliorer progressivement la précision de ses analyses et de s'adapter aux nouvelles tactiques utilisées par les cybercriminels. L'IA apprend également des habitudes de communication des utilisateurs : leur style d'écriture, leur comportement typique ou encore la structure des échanges. Cette compréhension du contexte est essentielle pour détecter des menaces plus avancées comme le spear phishing, qui vise spécifiquement des individus de haut rang, tels que les dirigeants d'entreprise. Grâce à ces capacités d'analyse comportementale et contextuelle, l'IA peut repérer des activités suspectes et bloquer une tentative de spear phishing avant qu'elle ne compromette les réseaux et les systèmes de l'organisation.

### V.3. Gestion des vulnérabilités

À mesure que les cybercriminels adoptent des méthodes de plus en plus sophistiquées, des milliers de nouvelles vulnérabilités sont découvertes et signalées chaque année. Les organisations se retrouvent ainsi confrontées à un volume croissant de failles potentielles, rendant leur gestion particulièrement complexe. Les systèmes de sécurité traditionnels peinent à suivre ce rythme et ne permettent pas toujours d'identifier ou de prévenir en temps réel les menaces les plus critiques. Les solutions de cybersécurité renforcées par l'intelligence artificielle, telles que l'analyse du comportement des utilisateurs et des entités (UEBA), offrent une réponse plus efficace à cette problématique. Elles permettent d'observer et d'analyser l'activité des utilisateurs, des appareils et des serveurs afin de détecter des comportements anormaux ou inhabituels susceptibles d'indiquer une attaque de type « zero-day ». Grâce à ces capacités

d'analyse comportementale et prédictive, l'IA peut protéger les organisations contre des vulnérabilités dont elles n'ont pas encore connaissance, bien avant qu'elles ne soient officiellement identifiées ou corrigées.

### V.4. Sécurité réseau

La sécurité réseau comprend des processus souvent longs et complexes, notamment la création de politiques de sécurité et la compréhension de la topologie du réseau. Une fois ces politiques établies, les organisations peuvent mettre en place des mécanismes permettant de distinguer les connexions légitimes de celles qui nécessitent une inspection approfondie pour détecter d'éventuels comportements malveillants. Ces politiques constituent également un élément essentiel pour appliquer une approche de sécurité fondée sur la confiance zéro (Zero Trust). Cependant, élaborer et maintenir ces politiques à l'échelle de plusieurs réseaux demande un investissement important en temps et en ressources humaines. Les organisations n'adoptent pas toujours des conventions de nommage cohérentes pour leurs applications et leurs charges de travail, ce qui complique l'analyse et oblige les équipes de sécurité à consacrer davantage de temps à identifier à quelle application appartient une charge de travail donnée. L'intelligence artificielle permet d'alléger ces contraintes. En apprenant progressivement les modèles de trafic réseau propres à une organisation, elle est capable de recommander des politiques plus adaptées et d'associer automatiquement les charges de travail aux applications correspondantes.

### V.5. Analyse comportementale

Grâce à l'analyse comportementale, les organisations peuvent identifier plus efficacement les menaces émergentes ainsi que les vulnérabilités connues. Les mécanismes de défense traditionnels reposent principalement sur des signatures d'attaque et des indicateurs de compromission (IOC) pour détecter les menaces. Toutefois, compte tenu des milliers de nouvelles attaques lancées chaque année, cette approche devient rapidement insuffisante et difficile à maintenir. L'intégration de l'analyse comportementale permet d'améliorer considérablement les processus de chasse aux menaces. Elle s'appuie sur des modèles d'intelligence artificielle capables d'établir des profils détaillés des applications, des appareils et des utilisateurs présents sur le réseau. Les volumes importants de données collectées sont alors comparés à ces profils de référence afin de repérer tout comportement suspect ou anormal. Ce type

d’analyse permet ainsi de détecter des activités potentiellement malveillantes, même lorsqu’aucune signature d’attaque n’a encore été définie.

## Conclusion

Ce chapitre a permis d’établir les bases conceptuelles nécessaires à la compréhension de la sécurité des infrastructures réseau ainsi que des apports potentiels de l’intelligence artificielle dans ce domaine. Nous avons défini les notions fondamentales telles que l’infrastructure réseau, les failles, les vulnérabilités et les menaces, tout en présentant les principales catégories d’attaques auxquelles les organisations sont confrontées. L’examen des approches classiques de sécurisation pare-feu, VPN, IDS/IPS a permis de mettre en évidence leurs points forts, mais également leurs limites face à des menaces toujours plus sophistiquées et évolutives. L’introduction aux techniques d’intelligence artificielle, en particulier le Machine Learning et le Deep Learning, a montré comment ces outils peuvent renforcer la cybersécurité en facilitant, par exemple, la détection proactive d’intrusions et l’automatisation des réponses aux attaques. Ainsi, ce chapitre souligne la nécessité d’intégrer l’IA dans les stratégies de protection des réseaux afin de pallier les insuffisances des méthodes traditionnelles et d’anticiper les nouvelles formes de cybermenaces. Cette réflexion prépare le terrain pour la suite de notre travail, qui portera sur l’analyse de l’efficacité des mécanismes de sécurité existants et sur l’identification des failles spécifiques propres aux infrastructures réseau.

# CHAPITRE 2 : NOTIONS DE SIEM, CORRELATION ET SOAR

## Introduction

La supervision de la sécurité des systèmes d'information constitue aujourd'hui un enjeu majeur pour les organisations, en raison de la complexité croissante des infrastructures et de l'augmentation continue des menaces cybernétiques. Les entreprises doivent être en mesure de collecter, d'analyser et d'interpréter en temps réel une grande diversité d'événements de sécurité provenant de plusieurs sources : pare-feu, serveurs, applications, postes de travail ou encore équipements réseau. Cette surveillance continue vise à détecter rapidement toute activité suspecte et à réagir de manière efficace aux incidents de sécurité. Pour répondre à ces besoins, les organisations s'appuient sur des solutions de Security Information and Event Management (SIEM), qui centralisent les journaux d'événements et permettent leur analyse corrélée afin d'identifier des comportements anormaux. Le SIEM occupe ainsi une place centrale dans le dispositif de supervision, en offrant une visibilité globale sur l'état de la sécurité du système d'information. Par ailleurs, la croissance du volume d'incidents à traiter amène de plus en plus les organisations à adopter des mécanismes d'automatisation et d'orchestration (Security Orchestration, Automation and Response – SOAR), afin de réduire les temps de réaction et de standardiser les processus de réponse.

## I. Notions de SIEM (Security Information and Event Management)

### I.1 Définition et rôle

Les solutions de Security Information and Event Management (SIEM) collectent et centralisent les journaux ainsi que les événements de sécurité afin de permettre aux équipes de les analyser sous forme de visualisations, d'alertes, de recherches ou de rapports. Ces informations jouent un rôle essentiel pour détecter des menaces en temps réel, gérer la réponse aux incidents, mener des analyses forensiques sur des événements passés ou encore préparer des audits de conformité. Pour identifier les menaces et les comportements anormaux, un SIEM ingère et analyse un volume très important de données en quelques secondes, afin de repérer et signaler les activités suspectes. La surveillance en temps réel des données issues de différentes sources applications, équipements réseau, matériel, environnements cloud, etc. permet ainsi d'anticiper tant les menaces internes qu'externes.

Dans de nombreuses organisations, le SIEM sert de tableau de bord central pour les équipes de sécurité, qui y réalisent une grande partie de leurs opérations quotidiennes.

## I.2.Architecture et fonctionnement du SIEM

Afin d'assurer une surveillance efficace, un système de gestion des informations et des événements de sécurité (SIEM) doit extraire et analyser des données issues de l'ensemble de l'organisation. Son fonctionnement repose sur plusieurs étapes essentielles, chacune contribuant à renforcer la supervision de la sécurité.

### Surveillance et collecte des données en temps réel

Le SIEM collecte en continu les journaux (logs) et événements provenant de multiples sources : serveurs, routeurs, pare-feu, applications, équipements réseau, etc. Cette collecte en temps réel offre une visibilité constante sur l'activité du système d'information.

### Analyse et corrélation des événements de sécurité

Une fois les données centralisées, le SIEM les analyse et applique des algorithmes de corrélation afin de détecter des actions suspectes provenant de différentes sources. Il identifie ainsi des modèles ou comportements pouvant révéler des activités malveillantes, comme des tentatives d'intrusion, des accès non autorisés ou des comportements anormaux.

### Centralisation des logs pour une vue globale

Un log est un fichier retraçant l'historique des événements sur un système informatique : connexions utilisateur, modifications de fichiers, alertes de sécurité, etc. Dans un SIEM, l'ensemble de ces logs est regroupé dans une interface unique. Cette centralisation permet au SOC (Security Operations Center) d'obtenir une vision d'ensemble du réseau et facilite :

- L'analyse post-incident,
- Les enquêtes sur les anomalies,
- La préparation des audits de conformité (ex. : RGPD, HIPAA).

Grâce à cette vue globale, le SIEM combine surveillance, analyse et gestion proactive des événements, ce qui renforce la capacité à détecter des menaces y compris les

tentatives d'ingénierie sociale en corrélant des alertes provenant de la messagerie, des solutions de sécurité et des logs d'accès.

**Visualisation via tableau de bord**

Le SIEM intègre un tableau de bord offrant une vue unifiée des informations collectées. Il présente les alertes de sécurité, classées par niveau de criticité, et facilite ainsi la collaboration et la prise de décision au sein des équipes de sécurité. La Figure 11 [13] ci-dessous montre les composants d'un SIEM.

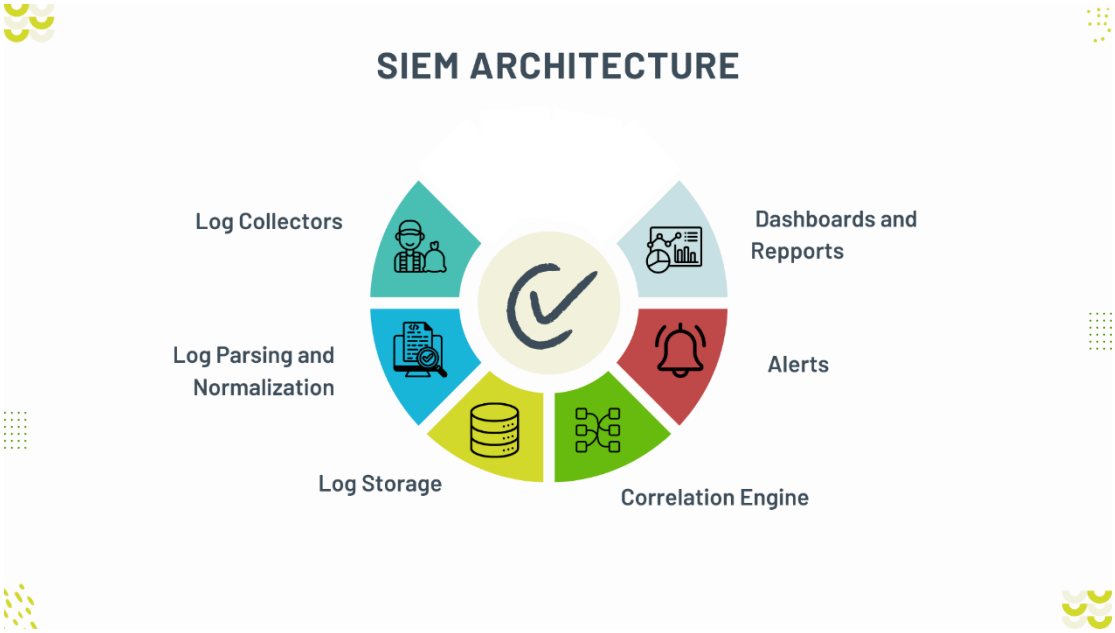



Figure 11: Architecture d'un SIEM

## I.3. Avantages et limites du SIEM

### I.3.1. Avantages

Avec un SIEM moderne, doté d'un bon fonctionnement et capable d'évoluer, les organisations profitent de :

**Visibilité holistique**

Le fait de disposer d'un emplacement centralisé à partir duquel les équipes peuvent monitorer, analyser en continu et intervenir sur leur environnement est essentiel. Cela fournit à l'équipe de sécurité une source unique de vérité.

**Récit unifié**

Un SIEM correctement configuré normalise les types de données disparates afin

d’offrir une vue cohérente (snapshot) du vaste environnement informatique d’une organisation.

**Détection automatisée des menaces**

Les spécialistes de la sécurité peuvent automatiser la détection des menaces et des anomalies, puis interroger rapidement les données pour enquêter sur une série d’événements, accéder aux données historiques afin d’identifier des tendances ou d’obtenir du contexte, et bien plus encore.

**Gestion des risques**

Grâce au SIEM, les équipes peuvent mettre en évidence des menaces inconnues via la détection d’anomalies alimentée par des modèles de Machine Learning prédéfinis, et ainsi obtenir des informations sur les entités les plus à risque.

### I.3.2. Limites

Bien que les solutions SIEM apportent une aide précieuse, elles ne sont pas exemptes de limites :

**Complexité d’implémentation**

Mettre en place un SIEM peut être complexe et nécessite une expertise approfondie en sécurité informatique. Il est également indispensable d’ajuster et de reconfigurer régulièrement les règles de corrélation afin de limiter les faux positifs et les faux négatifs.

**Coût élevé**

Les solutions SIEM sont souvent onéreuses, aussi bien en termes de licences que d’infrastructure et de personnel qualifié pour les administrer. Cela peut représenter un frein important pour les petites organisations.

**Gestion des faux positifs**

Un SIEM mal configuré peut générer un nombre élevé de faux positifs, c’est-à-dire des alertes qui ne correspondent pas à de réelles menaces. Cette surcharge peut détourner les équipes de sécurité de leurs priorités et réduire leur efficacité opérationnelle.

## II. Notions de Corrélation des événements de sécurité

### II.1. Définition

La corrélation désigne le processus d'analyse et de rapprochement de plusieurs événements ou alertes provenant de différentes sources (pare-feu, antivirus, IDS, serveurs, etc.) afin de détecter des attaques complexes, des comportements anormaux ou encore d'identifier l'origine d'un incident qu'un seul événement ne permettrait pas de révéler. Elle constitue le cœur du moteur analytique d'un SIEM (Security Information and Event Management), car elle joue un rôle essentiel dans la transformation de données brutes en informations exploitables. En pratique, un SIEM collecte en continu une grande quantité de logs et d'événements de sécurité issus de multiples composants du système d'information : pare-feu, serveurs, routeurs, systèmes d'exploitation, applications, etc. Pris isolément, ces événements offrent une vision partielle et souvent insuffisante de la situation. C'est grâce au processus de corrélation que le SIEM peut analyser, comparer et relier ces événements pour faire apparaître des liens pertinents. Cette analyse corrélée permet ainsi d'identifier des comportements anormaux ou des chaînes d'actions suspectes qui, considérées séparément, ne sembleraient pas malveillantes. Par exemple, une tentative de connexion échouée n'a rien d'alarmant ; mais si elle est suivie d'une connexion réussie depuis une adresse IP inhabituelle, puis d'un transfert massif de fichiers, la corrélation de ces événements met en évidence un scénario d'attaque potentiel.

### II.2. Les types de corrélation

La corrélation des événements est au cœur du fonctionnement d'un SIEM. Elle permet d'établir des liens logiques entre plusieurs événements de sécurité afin de détecter des comportements suspects ou des attaques complexes qui ne seraient pas identifiables à partir d'un seul log. On distingue plusieurs types de corrélation selon les critères et la logique utilisés pour relier les événements.

#### Corrélation temporelle

La corrélation temporelle repose sur l'analyse de la succession d'événements dans une fenêtre de temps définie. L'objectif est de détecter des séquences d'actions qui, isolément, paraissent anodines, mais qui, combinées dans un laps de temps restreint, révèlent un comportement anormal ou malveillant.

Elle est particulièrement utile pour identifier les différentes étapes d'une attaque, telles que la reconnaissance, l'exploitation ou l'exfiltration de données.

**Corrélation contextuelle**

La corrélation contextuelle tient compte des informations relatives à l'environnement dans lequel un événement survient : identité de l'utilisateur, rôle, localisation, type d'actif, classification de la ressource, etc. Par exemple, une tentative de connexion administrative (root, admin) depuis un poste utilisateur standard ou depuis une adresse IP inhabituelle peut être automatiquement considérée comme suspecte. Le SIEM utilise des données d'inventaire (Active Directory, CMDB, annuaire LDAP, etc.) pour contextualiser chaque log et ajuster sa pertinence analytique. Cette approche réduit considérablement les faux positifs et améliore la priorisation des alertes, car elle relie la détection à la réalité opérationnelle de l'organisation.

**Corrélation logique**

La corrélation logique repose sur des règles prédéfinies créées par les analystes SOC afin de détecter des scénarios d'attaque connus. Chaque règle combine plusieurs conditions logiques (AND, OR, NOT) pour établir un lien entre différents événements. Elle est très personnalisable et efficace pour détecter des situations précises, mais nécessite des ajustements réguliers pour rester pertinente face à l'évolution des menaces.

**Corrélation comportementale (statistique et IA)**

La corrélation comportementale repose sur la détection d'écarts significatifs entre un comportement observé et un comportement considéré comme normal. Elle s'appuie généralement sur des modèles statistiques ou des algorithmes d'apprentissage automatique (Machine Learning). Cette approche est particulièrement efficace pour détecter des menaces inconnues (zero-day), des comportements internes suspects ou des anomalies invisibles aux règles statiques. Elle complète les corrélations traditionnelles en offrant une analyse adaptative et évolutive.

**Corrélation topologique**

La corrélation topologique considère la localisation physique ou logique des systèmes et des événements. Elle permet d'identifier des schémas d'attaque affectant une zone ou un segment spécifique du réseau. Cette approche facilite la cartographie de l'impact d'un incident et la priorisation des segments à isoler.

**Corrélation par enrichissement (CTI)**

Cette forme de corrélation consiste à enrichir les événements collectés avec des informations externes issues de la Cyber Threat Intelligence (CTI). Elle vise à contextualiser et prioriser les alertes selon la réputation des entités impliquées (adresses IP, noms de domaine, hachages, URL, etc.).

L'enrichissement CTI augmente la pertinence analytique du SIEM et permet une détection proactive des menaces identifiées au niveau mondial.

La diversité des types de corrélation permet au SIEM de combiner règles statiques, analyses comportementales et informations issues de la CTI pour construire une vision complète et dynamique des menaces. En les exploitant conjointement, une organisation améliore la précision de ses alertes, réduit les faux positifs et accélère la détection des incidents.

### II.3. Les défis de la corrélation

Malgré son importance centrale dans la supervision de la sécurité, la corrélation des événements présente plusieurs défis opérationnels et analytiques qui peuvent limiter son efficacité. Le premier défi concerne la génération de faux positifs. En raison du nombre élevé de règles de corrélation et de la diversité des environnements informatiques, il n'est pas rare qu'un SIEM signale comme suspectes des activités pourtant légitimes. Cette situation entraîne une surcharge d'alertes pour les analystes du SOC, qui doivent trier manuellement les notifications pertinentes parmi un volume important de détections non critiques. Ce phénomène, souvent appelé alert fatigue, peut retarder la détection d'incidents réels et diminuer la réactivité des équipes. Un autre défi majeur réside dans le manque de contexte entourant les événements détectés. Une alerte isolée, dépourvue d'informations sur l'utilisateur, le système concerné ou la réputation d'une adresse IP, reste difficile à interpréter. Sans ce contexte, les analystes peinent à évaluer la gravité réelle de la situation et à prioriser les actions à mener. C'est précisément pour répondre à ces limites que l'intégration de la Cyber Threat Intelligence (CTI) devient essentielle. En enrichissant les événements collectés avec des données externes indicateurs de compromission, campagnes d'attaque connues, profils d'acteurs malveillants la CTI donne du sens et de la profondeur aux alertes générées par le SIEM. Elle contribue ainsi à réduire les faux positifs, à améliorer la

pertinence des corrélations et à orienter plus efficacement les efforts de réponse aux incidents.

## III. Automatisation et Orchestration de la réponse aux incidents

### III.1. Définition

Selon Gartner (2020), une plateforme SOAR se définit comme un ensemble d'outils permettant aux équipes SOC de collecter, d'analyser et de répondre automatiquement aux incidents de sécurité, en s'appuyant sur des playbooks prédéfinis et des intégrations avec divers outils de sécurité. Ces solutions assurent une coordination entre le SIEM, les systèmes de détection (IDS/IPS, EDR), les pare-feux et les plateformes de ticketing, afin d'exécuter des actions correctives de manière rapide, cohérente et automatisée.

### III.2. Fonctionnement d'une plateforme SOAR

Une architecture SOAR typique repose sur trois piliers principaux :

**Orchestration**

Coordination de plusieurs outils de sécurité (SIEM, EDR, pare-feu, sandbox, etc.) afin de centraliser les données et d'exécuter des actions synchronisées. Par exemple, lorsqu'une alerte critique est générée par le SIEM, le SOAR peut automatiquement interroger un outil EDR pour isoler la machine concernée.

**Automatisation**

Exécution automatique de tâches répétitives ou chronophages, telles que la collecte d'informations, la vérification d'indicateurs dans une base CTI ou la création de tickets d'incident. Le SOAR peut également télécharger automatiquement les métadonnées d'un fichier suspect et interroger des plateformes comme VirusTotal ou MISP pour en évaluer la dangerosité.

**Réponse et reporting**

Génération de rapports, notification aux analystes et exécution de contre-mesures conformément aux politiques de sécurité définies. Par exemple, le système peut bloquer automatiquement une adresse IP malveillante sur le pare-feu après corrélation des événements et validation par les données CTI.

### III.3. Apports et bénéfices du SOAR

Les bénéfices apportés par l'intégration d'une plateforme SOAR au sein d'un centre opérationnel de sécurité sont aujourd'hui largement reconnus et documentés dans la

littérature spécialisée. L'un des apports les plus significatifs réside dans la réduction considérable du temps de réponse aux incidents. Grâce à l'automatisation des actions correctives, le délai entre la détection d'une menace et la mise en œuvre d'une mesure de remédiation peut être réduit de plusieurs heures à seulement quelques secondes. Cette rapidité permet non seulement de limiter l'impact d'une attaque, mais aussi de renforcer la résilience globale du système d'information. Par ailleurs, le SOAR contribue à la standardisation des procédures de sécurité. En appliquant des scénarios d'intervention préétablis (playbooks), chaque incident est traité selon un processus défini et reproductible. Cette homogénéité garantit la cohérence opérationnelle des réponses, tout en réduisant les risques d'erreurs humaines souvent observées dans les environnements de supervision manuelle. Un autre avantage majeur réside dans l'allègement de la charge de travail des équipes du SOC. En automatisant les tâches répétitives et chronophages, telles que la collecte de logs, la vérification d'indicateurs de compromission ou la génération de tickets d'incident, le SOAR libère les analystes de sécurité, leur permettant de se concentrer sur les investigations à forte valeur ajoutée et les incidents critiques nécessitant une expertise humaine. Le renforcement de la traçabilité constitue également un atout essentiel. Chaque action effectuée par la plateforme est enregistrée et documentée, offrant ainsi une visibilité complète sur l'historique des traitements réalisés. Cette capacité d'audit est particulièrement utile dans le cadre des exigences de conformité et de gouvernance, car elle permet de justifier les décisions et les mesures adoptées lors d'un incident de sécurité. Enfin, le SOAR se distingue par sa flexibilité et sa capacité d'intégration continue. Grâce à ses connecteurs basés sur des interfaces de programmation (API) et à sa compatibilité avec les formats standards tels que STIX et TAXII, il peut s'adapter facilement à la diversité des environnements techniques. Cette interopérabilité assure une communication fluide entre les différents outils de sécurité (SIEM, CTI, pare-feu, etc.) et permet une évolution progressive de l'architecture sans rupture opérationnelle. Ainsi, l'adoption d'un SOAR transforme en profondeur la dynamique du centre opérationnel de sécurité, en passant d'une approche essentiellement réactive à une stratégie proactive, automatisée et cohérente, où la rapidité, la précision et la traçabilité deviennent les piliers d'une supervision de sécurité moderne et efficiente.

### III.4. Limites et défis du SOAR

Malgré les nombreux avantages qu'offre une plateforme SOAR, son déploiement et son exploitation présentent plusieurs défis techniques, organisationnels et méthodologiques qu'il convient d'anticiper. L'un des principaux obstacles réside dans la complexité d'intégration au sein d'environnements informatiques hétérogènes. En effet, la mise en place d'un SOAR nécessite la connexion d'une multitude d'outils de sécurité tels que les SIEM, les EDR, les pare-feux, les bases CTI ou encore les systèmes de ticketing via des connecteurs ou des interfaces de programmation (API). Cette interconnexion exige une configuration rigoureuse et souvent des développements spécifiques pour assurer la compatibilité et la synchronisation des flux d'information. Un autre défi majeur concerne la qualité et la pertinence des scénarios d'automatisation, communément appelés playbooks. Ces scénarios constituent le cœur du fonctionnement du SOAR, mais leur conception requiert une connaissance approfondie des menaces, des processus internes et des politiques de sécurité de l'organisation. Un playbook mal défini ou trop générique peut entraîner des conséquences indésirables, telles que le blocage injustifié de services légitimes, la suppression de fichiers critiques ou la génération de nouvelles alertes non pertinentes. Il est donc essentiel de calibrer avec précision les conditions d'exécution et de validation de chaque action automatisée. La dépendance aux données issues de la Cyber Threat Intelligence (CTI) représente également une contrainte significative. L'efficacité d'un SOAR repose en grande partie sur la fiabilité et la fraîcheur des indicateurs de compromission utilisés pour prendre des décisions automatiques. Si les flux CTI sont obsolètes, incomplets ou peu fiables, les actions entreprises peuvent s'avérer inefficaces, voire contre-productives. Une gestion rigoureuse des sources de renseignement et un processus de validation des données s'avèrent donc indispensables. Sur le plan organisationnel, la mise en œuvre d'une solution SOAR implique un changement culturel au sein du centre opérationnel de sécurité. L'introduction de l'automatisation modifie en profondeur les modes de travail, les responsabilités et les interactions entre les analystes. Certains processus auparavant manuels deviennent automatisés, ce qui peut susciter une résistance au changement ou une perte de visibilité sur certaines étapes du traitement des incidents. Pour pallier ce risque, il est recommandé d'adopter une approche progressive, en commençant par l'automatisation de tâches simples et répétitives, telles que l'enrichissement d'IOC ou la création de tickets, avant d'étendre les capacités du SOAR à des scénarios plus complexes.

## Conclusion

Ce chapitre a exploré les fondements de la supervision de la sécurité des systèmes d'information, en mettant en évidence le rôle central de la corrélation des événements dans les SIEM. Cette corrélation permet d'identifier des comportements suspects à partir de flux massifs de données, mais elle reste limitée si elle n'est pas enrichie par un contexte pertinent. Parallèlement, la complexité croissante des infrastructures et le volume des alertes rendent nécessaire l'automatisation et l'orchestration de la réponse via les plateformes SOAR. Celles-ci standardisent les procédures, exécutent automatiquement des actions correctives et assurent la traçabilité, tout en réduisant la charge opérationnelle du SOC. Leur efficacité repose cependant sur une intégration soignée et des playbooks bien conçus. En somme, l'articulation entre SIEM et SOAR constitue un écosystème intégré et complémentaire, capable de détecter, analyser et traiter les menaces de manière proactive et cohérente. Cette approche assure une supervision plus précise, rapide et résiliente des systèmes d'information, répondant aux exigences des environnements numériques actuels.

# CHAPITRE 3 : CONCEPTION DU SYSTEME INTELLIGENT

## Introduction

Ce chapitre présente la concrétisation du système proposé, en explorant la manière dont les concepts et méthodes théoriques développés précédemment se traduisent en une solution opérationnelle. Il s'agit de montrer comment un ensemble d'outils et de technologies peut être orchestré pour surveiller, analyser et répondre aux incidents de sécurité de manière efficace. L'objectif principal est de démontrer le fonctionnement pratique du système, en mettant en lumière les mécanismes qui permettent de transformer des données brutes issues du réseau en informations exploitables, et de prendre des mesures automatiques face aux menaces identifiées. Cette mise en œuvre illustre non seulement les capacités techniques du système, mais également sa valeur en termes de sécurité proactive et de gestion intelligente des incidents.

## I. Mise en place du SIEM

Une solution SIEM (Security Information and Event Management), dédiée à la gestion des informations et des événements de sécurité, offre un aperçu global de l'infrastructure IT concernée à tout moment. Capable d'analyser les données issues de toutes les sources (applications réseau, matériels, cloud et solutions SaaS) en temps réel, le SIEM est incontournable pour aider les entreprises à anticiper les menaces internes et externes. C'est dans cette optique que nous avons déployé un SIEM basé sur la suite Elastic Stack, nous permettant de centraliser, corréler et visualiser les événements de sécurité provenant de différentes sources, facilitant ainsi la détection précoce des incidents. Pour ce faire, nous combinons Suricata IDS avec Elasticsearch, Kibana et Filebeat.

**Suricata** est un système open-source avancé de protection des réseaux informatiques. Il s'appuie sur un moteur de règles puissant et effectue une inspection approfondie des paquets (Deep Packet Inspection). Il est reconnu pour sa capacité à analyser le trafic réseau à la recherche d'événements suspects et à bloquer les paquets invalides. Nous l'avons déployé sur un serveur configuré en mode IDS (Intrusion Detection System).

Dans le fichier de configuration `/etc/suricata/suricata.yaml`, nous avons associé Suricata à l'interface réseau `ens3` en mode écoute, afin qu'il inspecte le trafic y circulant. Cette configuration est définie dans la section `af-packet`. Ce module

assure notamment une capture performante grâce à l'équilibrage de charge et au traitement parallèle. La configuration établie indique que Suricata capture les paquets sur l'interface **ens3**, permettant ainsi une analyse en temps réel du trafic réseau circulant entre les différents segments du système.

```
af-packet:
  - interface: ens3
    # Number of receive threads. "auto" uses the number of cores
    #threads: auto
    # Default clusterid. AF_PACKET will load balance packets based on flow.
    cluster-id: 99
    # Default AF_PACKET cluster type. AF_PACKET can load balance per flow or per hash.
    # This is only supported for Linux kernel > 3.1
    # possible value are:
    #  * cluster_flow: all packets of a given flow are sent to the same socket
    #  * cluster_cpu: all packets treated in kernel by a CPU are sent to the same socket
    #  * cluster_qm: all packets linked by network card to a RSS queue are sent to the same
    #  socket. Requires at least Linux 3.14.
    #  * cluster_ebpf: eBPF file load balancing. See doc/userguide/capture-hardware/ebpf-xdp.rst for
    #  more info.
    # Recommended modes are cluster_flow on most boxes and cluster_cpu or cluster_qm on system
    # with capture card using RSS (requires cpu affinity tuning and system IRQ tuning)
    # cluster_rollover has been deprecated; if used, it'll be replaced with cluster_flow.
    cluster-type: cluster_flow
    # In some fragmentation cases, the hash can not be computed. If "defrag" is set
    # to yes, the kernel will do the needed defragmentation before sending the packets.
    defrag: yes
```

Figure 12: Définition de l'interface de capture

Nous définissons ensuite la variable HOME_NET qui permet à Suricata de distinguer le réseau interne du réseau externe. Cette distinction est essentielle pour la détection d'attaques, car de nombreuses règles IDS s'appuient sur le sens du trafic (entrant ou sortant).

```
vars:
  # more specific is better for alert accuracy and performance
  address-groups:
    HOME_NET: "[192.168.0.0/16,10.0.0.0/8,172.16.0.0/12]"
    #HOME_NET: "[192.168.0.0/16]"
    #HOME_NET: "[10.0.0.0/8]"
    #HOME_NET: "[172.16.0.0/12]"
    #HOME_NET: "any"

    EXTERNAL_NET: "!$HOME_NET"
    #EXTERNAL_NET: "any"
```

Figure 13: Définition du réseau surveillé

Nous précisons le format de la sortie dans la section Outputs qui est la partie la plus importante dans un SIEM. En effet, elle détermine comment les événements générés par Suricata sont exportés vers Filebeat. Dans notre cas, les fichiers générés par suricata sont soit en format fast.log soit en format eve.json.

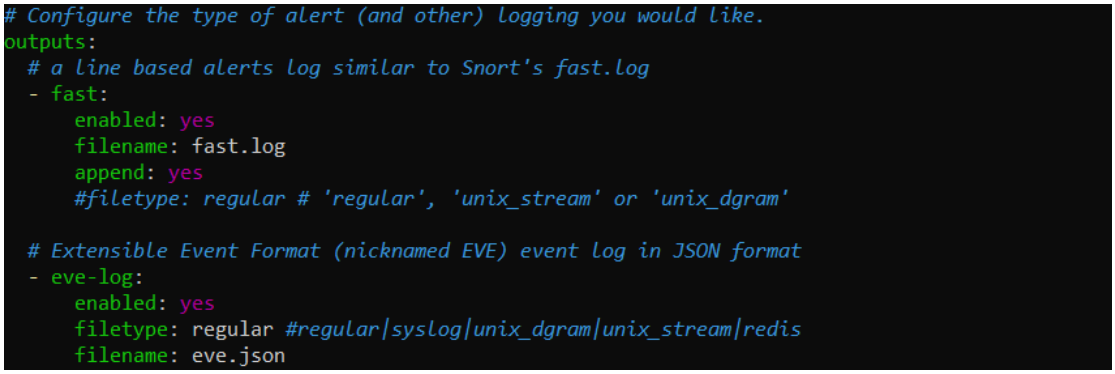



Figure 14: Configuration de l'interface de sortie

Suricata repose sur un ensemble de règles de détection permettant d'identifier les attaques connues. Le moteur IDS analysera chaque paquet ou flux réseau selon ces règles afin de générer des alertes en cas de comportement suspect.

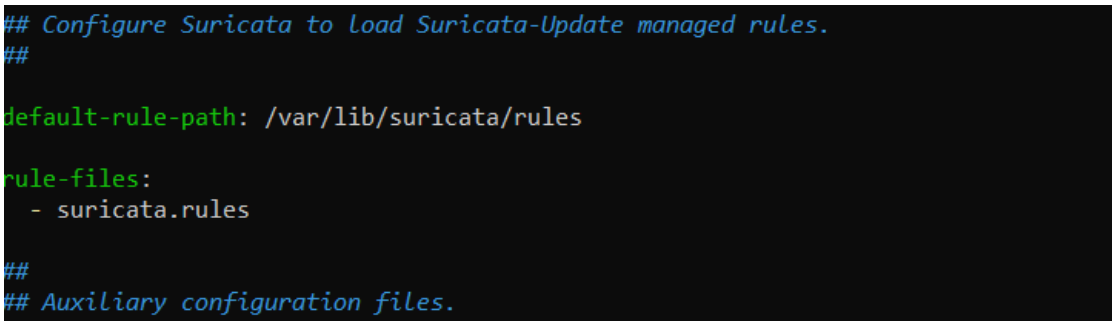



Figure 15: Chargement des règles IDS

**Filebeat** est l'agent de collecte de logs développé par Elastic. Il est conçu pour assurer la transmission fiable et structurée des événements vers la plateforme Elastic Security. Filebeat surveille en continu les fichiers de logs et détecte toute nouvelle entrée afin de la transformer en un événement exploitable. Installé sur le même serveur que Suricata, il collecte, parse et normalise le fichier /var/log/suricata/eve.json généré par ce dernier. Filebeat utilise un module suricata pour lire les fichiers eve.json en continu, enrichir les champs (IP, ports, catégories, signatures) puis les convertit en format ECS (Elastic Common Schema). Il va enfin envoyer les données vers Elasticsearch par indexation.

La première étape de la configuration de Filebeat consiste à lui indiquer quel fichier il doit lire (eve.json).

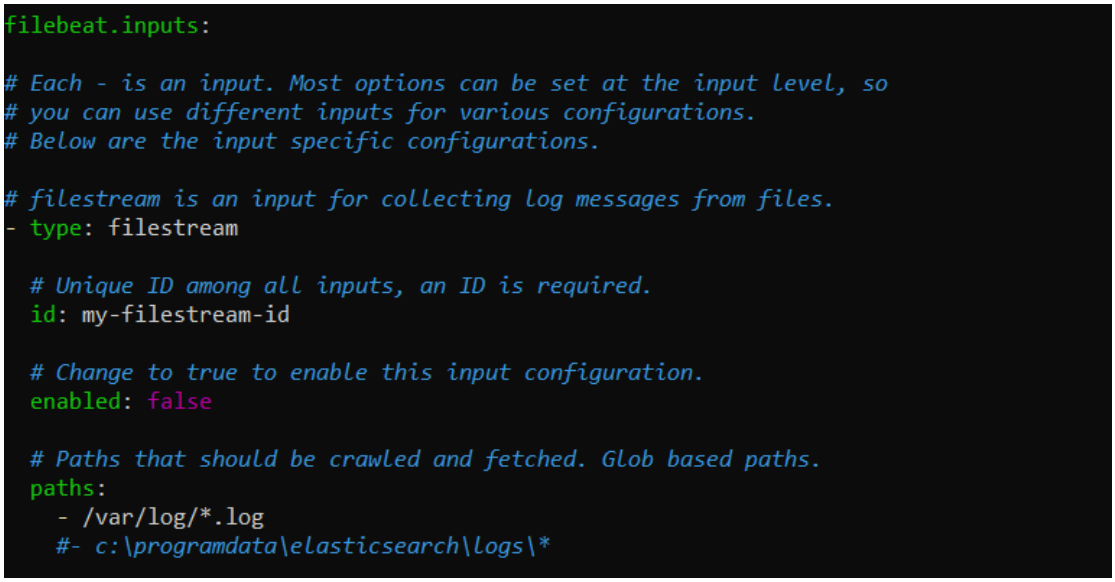



Figure 16: Définition du fichier de log à surveiller

Filebeat propose un module permettant l'enrichissement des données, une mise au format ECS (Elastic Common Schema), la création automatique de dashboards dans Kibana. L'utilisation du module améliore considérablement la lisibilité et l'analyse des événements dans Kibana

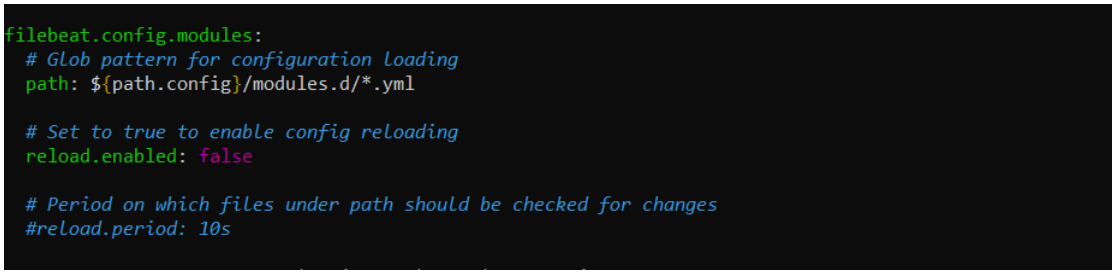



Figure 17: Activation des modules

Une fois les données collectées, Filebeat les envoie directement à Elasticsearch grâce à la section output.elasticsearch. Dans notre environnement, la communication se fait via HTTPS avec des restrictions d'accès.

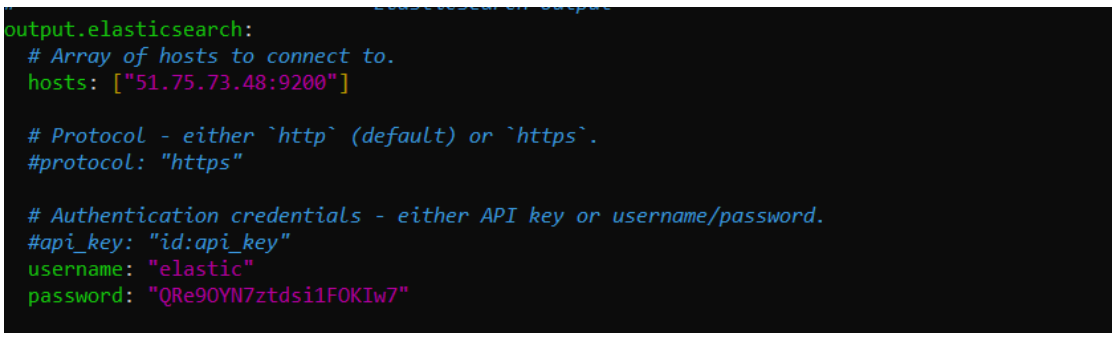



Figure 18: Envoie des logs vers Elasticsearch

**Elasticsearch** est le moteur de recherche et d'analyse distribué nous permettant de traiter les grands volumes de données en temps réel. Toujours sur le même serveur que suricata, il stocke toutes les données du SIEM. Elasticsearch va collecter, structurer, indexer et interroger efficacement les logs produits par Suricata. Une fois qu'il reçoit les événements, il les stocke au sein d'un index. Chaque document ingéré est analysé et transformé en un ensemble de tokens qui alimentent un index inversé. Constituant une plateforme robuste et performante pour stocker, indexer et interroger efficacement de larges volumes d'événements, Elasticsearch nous permet ici d'effectuer des recherches extrêmement rapides sur des champs variés, qu'il s'agisse d'adresses IP, d'identifiants, de timestamps ou de messages textuels.

Dans le fichier de configuration d'Elasticsearch, /etc/elasticsearch/elasticsearch.yml nous définissons d'abord les paramètres réseau dont network.host qui permet d'autoriser les connexions depuis Filebeat et Kibana ainsi que le port 9200 utilisé par les API REST.

```
network.bind_host: ["127.0.0.1", "51.75.73.48"]
#
# By default Elasticsearch listens for HTTP traffic on the first free port it
# finds starting at 9200. Set a specific HTTP port here:
#
network.host: 0.0.0.0
http.port: 9200
#
# For more information, consult the network module documentation.
#
```

Figure 19: Paramètres réseaux

Nous avons configuré Elasticsearch sur un nœud unique afin d'éviter les alertes liées au clustering puis activé le système de sécurité (compte utilisateurs, mots de passe, protections) avant de permettre la génération des clés d'accès sécurisées dans Elastic.

```
# ---------------------------------- Security ----------------------------------
#
#                                *** WARNING ***
#
# Elasticsearch security features are not enabled by default.
# These features are free, but require configuration changes to enable them.
# This means that users don't have to provide credentials and can get full access
# to the cluster. Network connections are also not encrypted.
#
# To protect your data, we strongly encourage you to enable the Elasticsearch security features.
# Refer to the following documentation for instructions.
#
# https://www.elastic.co/guide/en/elasticsearch/reference/7.16/configuring-stack-security.html
discovery.type: single-node
xpack.security.enabled: true
xpack.security.authc.api_key.enabled: true
```

Figure 20: Sécurisation d'Elasticsearch

**Kibana** est l'interface Web pour la représentation visuelle de nos données collectées. Avec Elasticsearch et Filebeat, il forme ce qu'on appelle la pile ELK (The Elastic Stack). Cette suite open source nous permet de collecter les données de Suricata (et cela dans n'importe quel format), de les organiser et de les préparer à des fins analytiques. En plus de la possibilité de visualiser les données traitées par Elasticsearch, Kibana nous offre également une analyse automatique en temps réel, un algorithme de recherche très flexible et différents types de visualisations (histogrammes, graphiques, diagrammes circulaires, etc.) pour chaque donnée recherchée.

Dans le fichier de configuration de Kibana /etc/kibana/kibana.yml, nous avons d'abord activé xpack qui est une fonction de sécurité générant des clés secrètes. Kibana utilise ces clés pour stocker les données dans Elasticsearch.

```
# Specifies locale to be used for all localizable strings, dates and number formats.
# Supported languages are the following: English - en , by default , Chinese - zh-CN .
#i18n.locale: "en"
xpack.encryptedSavedObjects.encryptionKey: fde5c5f6f89cf7802605176c6f07af75
xpack.reporting.encryptionKey: e44a4100ee662db71c047cdd5d41a8db
xpack.security.encryptionKey: acb6fcfc0aeb12351bd540a40a030a4c
```

Figure 21: Activation de la xpack

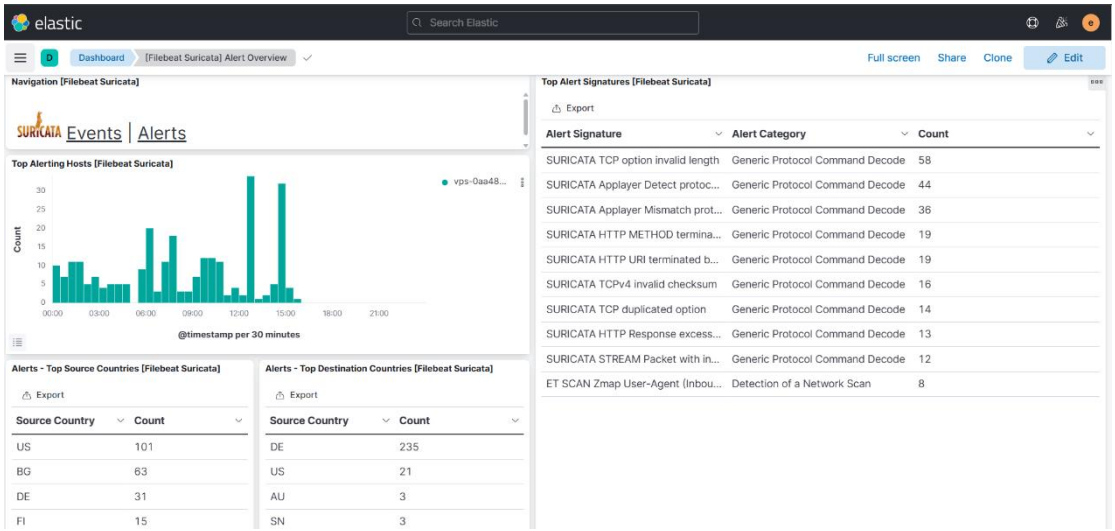



Figure 22: Tableau de bord Kibana

La construction de notre SIEM basé sur l'écosystème Elastic repose sur un enchaînement cohérent de ces composants complémentaires. Suricata IDS constitue la source primaire d'événements : il analyse le trafic réseau, applique ses règles de détection et génère des alertes au format EVE-JSON. Ces journaux représentent les données brutes de sécurité. Installé sur le même serveur, Filebeat assure leur collecte et leur transfert. Grâce à son module Suricata, il lit en continu les fichiers de logs, les

normalise et les expédie de manière fiable vers Elasticsearch, garantissant une chaîne de collecte sans perte d'événements. Elasticsearch joue le rôle de cœur analytique du SIEM. Il indexe et stocke les logs afin de permettre des recherches rapides, tout en offrant des possibilités de corrélation, d'enrichissement et d'analyse temporelle. Il devient ainsi le socle central où convergent et s'organisent les données de sécurité. Enfin, Kibana fournit la couche de visualisation et d'investigation. Grace à ses tableaux de bord, ses outils d'exploration et son interface Elastic Security, il nous permet de consulter les alertes, d'examiner les événements Suricata et d'obtenir une visibilité globale sur l'activité du réseau.

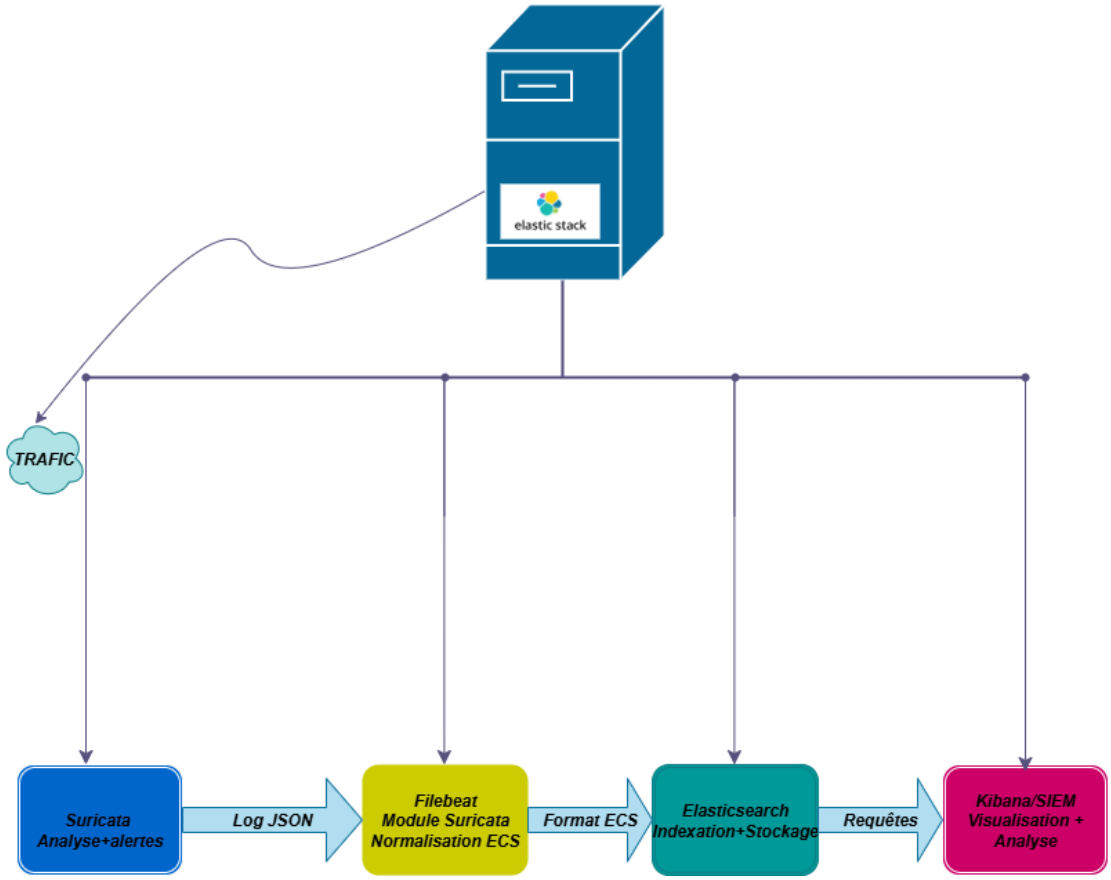



Figure 23: Architecture physique SIEM

## II. Mise en place du SOAR(Thehive+Cortex) et connexion au SIEM

### II.1. Mise en place du SOAR (Thehive+Cortex)

Le SOAR (Security Orchestration, Automation, and Response - Orchestration, automatisation et réponse en matière de sécurité) nous permet de standardiser et de rationaliser la réponse face aux attaques et aux incidents. Le SOAR optimise les workflows à l'intérieur et au-delà du centre des opérations de sécurité et standardise les

processus SOC, ce qui assure un examen et une réponse cohérents tout en permettant aux analystes en sécurité d'améliorer leurs compétences. En automatisant les workflows d'un certain nombre de tâches manuelles simples en général associées à une réponse aux incidents (enregistrement des incidents de sécurité, alerte des parties concernées, envoi et mise à jour des tickets de rapport), le SOAR réduit nettement le temps moyen de résolution. C'est avec cette logique que nous mettons en place un SOAR Thehive+Cortex afin de concrétiser l'automatisation de la réponse pour notre système.

Nous avons mis en place notre SOAR en utilisant la méthode Docker afin d'installer Thehive et Cortex à l'intérieur de conteneur puis définit l'architecture du SOAR ainsi que les interactions entre ses composant (Thehive, Cortex, Elasticsearch) dans le fichier de configuration docker-compose.yml.

**TheHive** est une plateforme open source de réponse aux incidents de sécurité (SIRP) nous permettant de gérer efficacement les alertes et les enquêtes. Elle centralise les informations issues du SIEM et les organise sous forme de cas (cases) et de tâches, afin de structurer le travail d'investigation. Lorsqu'une alerte est reçue, elle peut être examinée puis transformée en cas, qui devient l'unité principale de gestion de l'incident. Chaque cas sera enrichi par des notes, des pièces jointes, des indicateurs de compromission (IoCs) et des tâches assignées aux analystes. TheHive facilite ainsi la collaboration, la traçabilité et la coordination des actions.

Les spécifications du fonctionnement de Thehive sont définies dans le fichier docker-compose.yml.

```
thehive:
  image: strangebee/thehive:5.3.0
  container_name: thehive
  depends_on:
    - cortex
  environment:
    - CORTEX_URL=http://cortex:9001
  ports:
    - "9000:9000"
  networks:
    - thehive
```

Figure 24: Configuration de Thehive

Dans cette configuration du conteneur Thehive,

- L'image strangebee/thehive:5.3.0 fournit l'environnement complet du SOAR.

- Le conteneur est nommé explicitement pour faciliter la gestion.
- depends_on : garantit que Cortex est disponible avant TheHive et permet l'intégration fluide des analyseurs.
- La variable d'environnement CORTEX_URL connecte TheHive au moteur d'analyse Cortex pour l'automatisation des réponses aux incidents.
- Le port 9000 est exposé pour accéder à l'interface web.
- L'utilisation d'un réseau Docker dédié assure une communication interne sécurisée et isolée entre TheHive et Cortex.

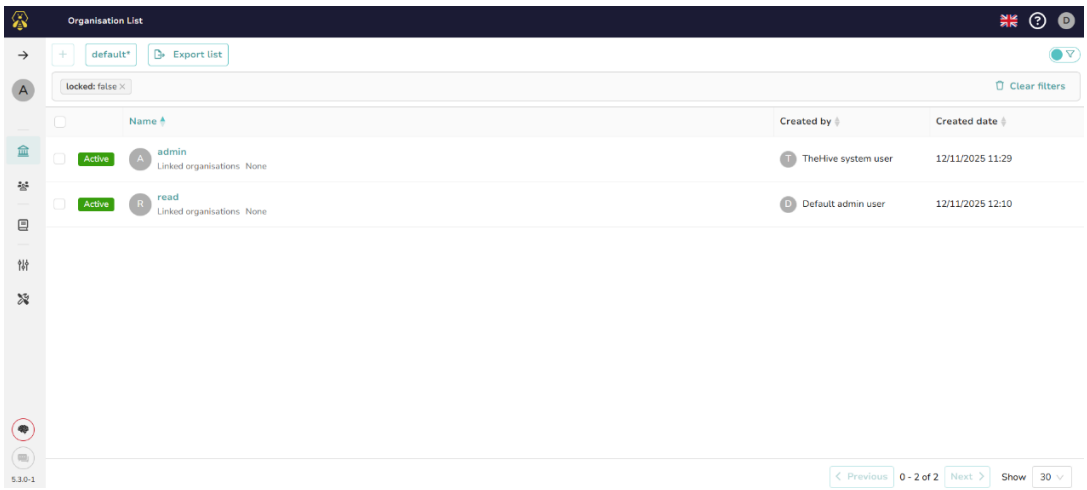



Figure 25: plateforme Thehive

**Cortex,** une plateforme open source d'analyse automatisée enrichit et examine rapidement les indicateurs ou les artefacts liés aux incidents. Il agit comme un moteur centralisé capable d'exécuter de nombreux modules d'analyse, appelés analyzers et responders, développés pour effectuer des tâches spécifiques comme l'analyse de fichiers, la recherche d'adresses IP, la consultation de bases de menaces ou encore l'extraction de métadonnées. Lorsqu'on soumet un élément à Cortex (fichier, URL, adresse IP, hash…), la plateforme sélectionne et exécute automatiquement les analyzers appropriés. Les résultats générés sont normalisés, détaillés et renvoyés à TheHive, afin d'enrichir l'enquête en cours. Cette automatisation réduit le temps nécessaire pour collecter des informations essentielles et améliore la précision des analyses. Cortex va également effectuer des actions de réponse grâce aux responders, qui permettent par exemple d'isoler un hôte, de bloquer une adresse IP ou d'envoyer une alerte à un autre système.

La configuration du conteneur Cortex est la suivante :

```yaml
cortex:
  image: thehiveproject/cortex:3.1.8
  container_name: cortex
  depends_on:
    - elasticsearch
  environment:
    - CORTEX_ELASTICSEARCH_URL=http://elasticsearch:9200
    - CORTEX_PORT=9001
  ports:
    - "9001:9001"
  networks:
    - thehive
```

Figure 26: Configuration du conteneur cortex

- L'image thehiveproject/cortex:3.1.8 fournit l'environnement complet du moteur d'analyse.
- Le conteneur est nommé explicitement pour faciliter la gestion avec Docker.
- depends_on : Elasticsearch garantit que Cortex peut se connecter à Elasticsearch au démarrage.
- La variable CORTEX_ELASTICSEARCH_URL configure la connexion à Elasticsearch pour stocker et indexer les résultats des analyseurs.
- Le port 9001 est exposé pour accéder à l'interface web et à l'API de Cortex.
- Le réseau Thehive permet une communication sécurisée et directe avec TheHive et Elasticsearch.

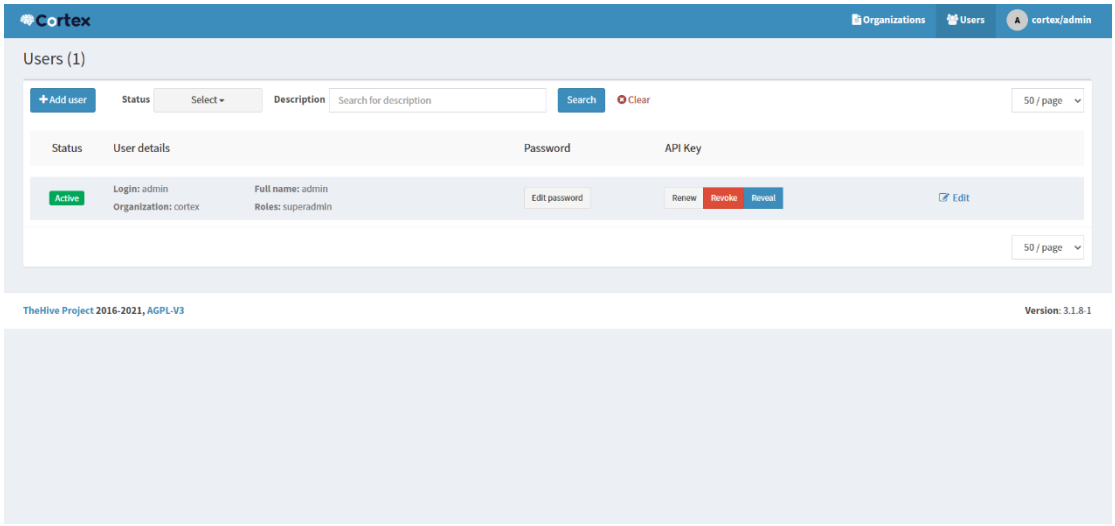



Figure 27: plateforme Cortex

Afin d'éviter toute confusion de port, nous avons installé un conteneur elasticsearch propre au SOAR. Sa configuration dans le fichier docker-compose.yml se présente ainsi :

```yaml
elasticsearch:
  image: docker.elastic.co/elasticsearch/elasticsearch:7.17.14
  container_name: cortex-elasticsearch
  environment:
    - discovery.type=single-node
    - ES_JAVA_OPTS=-Xms512m -Xmx512m
  volumes:
    - esdata:/usr/share/elasticsearch/data
  ports:
    - "9210:9200"
  networks:
    - thehive
```

Figure 28 : Configuration du conteneur Elasticsearch dédiée à cortex

- L'image elasticsearch:7.17.14 fournit un moteur d'indexation et de stockage compatible avec Cortex.
- Le conteneur est nommé explicitement pour simplifier la gestion Docker.
- discovery.type=single-node assure un fonctionnement sur un seul nœud adapté aux tests ou environnement mémoire limité.
- La mémoire Java est limitée à 512 Mo pour éviter de saturer le serveur.
- Le volume esdata garantit la persistance des données et résultats des analyseurs.
- Le port 9210 exposé permet à Cortex d'accéder au moteur Elasticsearch.
- L'utilisation du réseau Docker thehive assure une communication interne sécurisée entre Cortex et Elasticsearch.

Constituant le SOAR, l'interaction entre thehive et cortex repose sur une intégration directe via l'API de Cortex, permettant une automatisation fluide des analyses et des actions de réponse au sein de notre système. Pour se faire, dans la menue administration de Thehive nous avons ajouté un serveur cortex en renseignant le nom(cortex), l'URL (http://cortex:9001) et la clé générée par cortex.

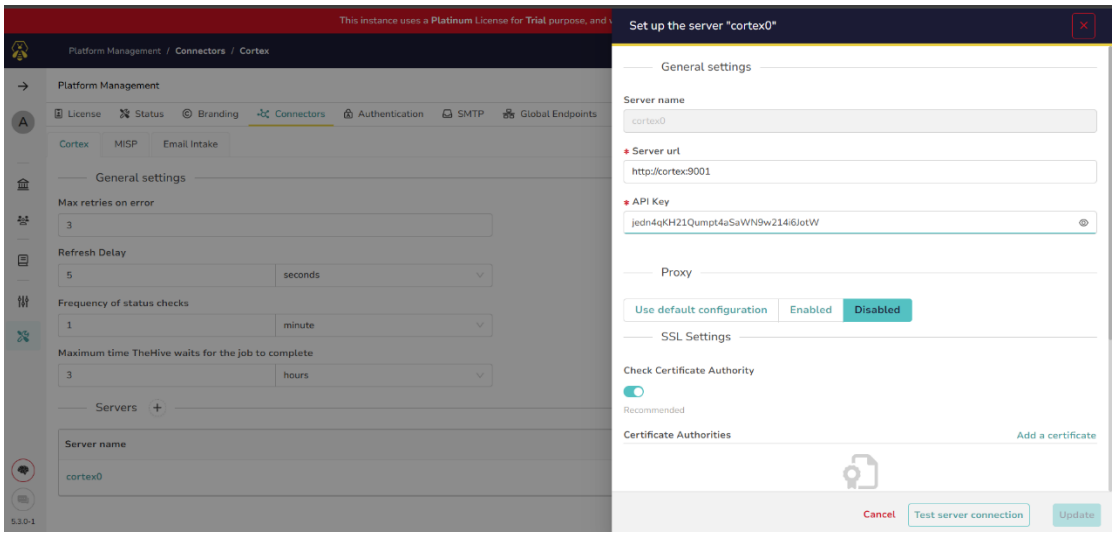



Figure 29: Connexion thehive et cortex

TheHive représente la plateforme centrale où l'on reçoit, examine et structure les incidents sous forme de cas. Souhaitant enrichir un élément d'un cas (fichier suspect, URL, adresse IP, hash, domaine…), nous allons déclencher depuis TheHive une analyse automatisée vers Cortex. Cette requête sera envoyée via l'API de Cortex, qui sélectionne ensuite les analyzers adaptés au type d'artefact soumis. Une fois l'analyse terminée, Cortex renvoie directement les résultats vers TheHive sous un format structuré. Ces informations enrichissent alors le cas en cours, permettant d'obtenir rapidement des données additionnelles (réputation, indicateurs de compromission, analyses comportementales, métadonnées, etc.). Cette boucle d'échange augmente la rapidité et la précision de l'investigation tout en réduisant les actions manuelles. En plus des analyzers, TheHive peut également déclencher des responders disponibles dans Cortex pour effectuer des actions de remédiation, comme bloquer une adresse IP ou envoyer un artefact à un autre outil. Ainsi, l'intégration TheHive–Cortex forme une chaîne complète d'analyse et de réponse automatisée.

### II.2. Connexion du SOAR au SIEM

Après avoir mise en place notre SOAR, nous allons le connecter à notre SIEM ElasticSecurity, afin que ce dernier envoie directement ses alertes de manière automatique à Thehive via webhook qui est une méthode de communication qui relie deux applications pour qu'elles échangent des informations de manière autonome dès qu'un événement se produit, évitant ainsi d'avoir à vérifier manuellement l'état du système. Sachant que Thehive accepte les alertes via son API, nous avons récupéré son URL d'ingestion (http://51.75.73.48:9000/api/alert) avant d'y créer un utilisateur API du nom de Elastic Connector.

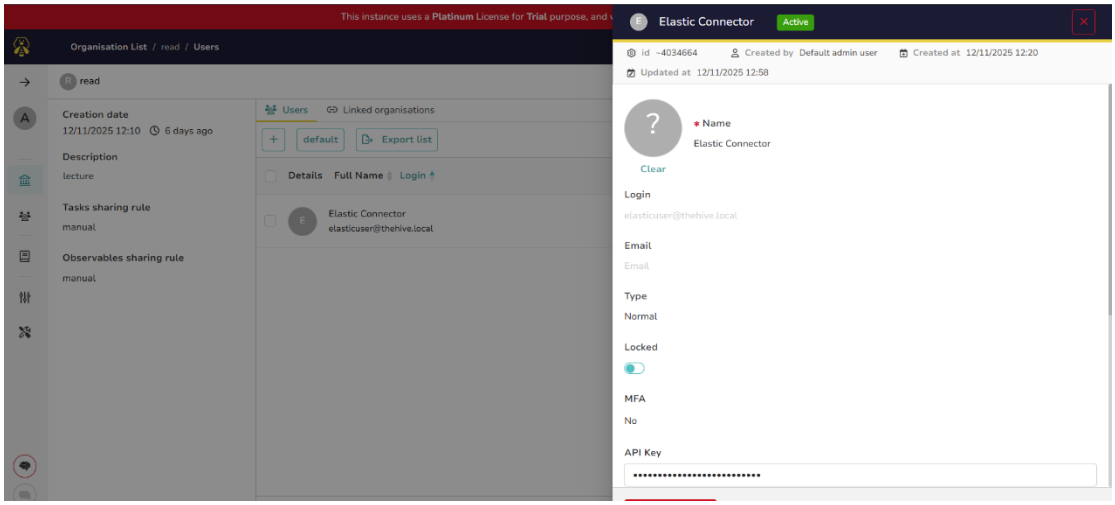


Figure 30: Création de l'utilisateur API

Une fois créer l'utilisateur dans thehive, dans Elastic Security (Kibana), nous avons créé un connecteur du nom de yarame et de type Webhook, ainsi ajouter à ses paramètres l'URL d'ingestion de thehive et valider son autorisation d'envoi avec la clé générée par thehive pour l'utilisateur Elastic Connector. En dernier lieu, dans la partie Rules d'Elastic, nous avons associé le connecteur Webhook à une règle afin de rendre précise la détection de menaces.

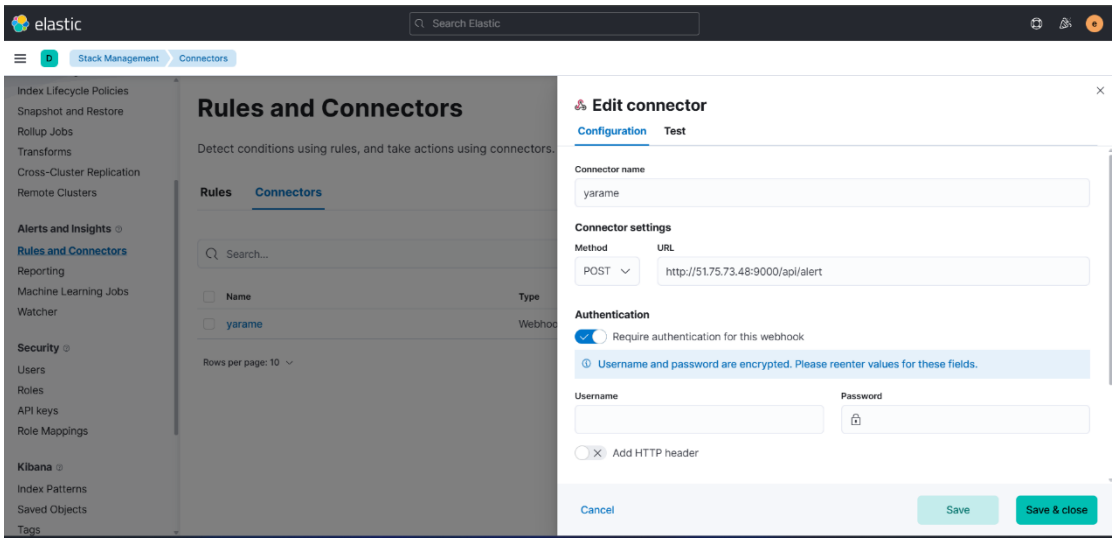


Figure 31: Création du connecteur webhook

Ces étapes faites, une fois Elastic Security détectera une alerte, il enverra automatique une requête webhook à Thehive et ce dernier créera un incident (case) avant d'appeler Cortex pour analyser les IoCs (IP, hash, domaine, …). Ainsi, nous obtenons une chaine complète de détection, d'analyse et de réponse.

## III. Développement du système intelligent de corrélation

Dans le cadre de ce mémoire, nous avons développé un module de corrélation automatique des alertes reposant sur un modèle de langage (LLM) de type TinyLlama. L'objectif du module est de déterminer automatiquement si des alertes Suricata sont corrélées, c'est-à-dire si elles appartiennent à un même incident ou à une même chaîne d'attaque, en tenant compte de leurs informations techniques (IP source, port, signature…), de leurs descriptions et des relations implicites entre événements. La corrélation intelligente repose ici sur un LLM qui analyse les alertes de façon sémantique, ce qu'un système de corrélation classique ne peut faire.

L'architecture du modèle est constituée de trois grandes parties dont :

**Chargement et préparation du modèle LLM :**

Le modèle TinyLlama-1.1B-Chat est chargé depuis HuggingFace, grâce à notre clé d'accès Suricata-bot.

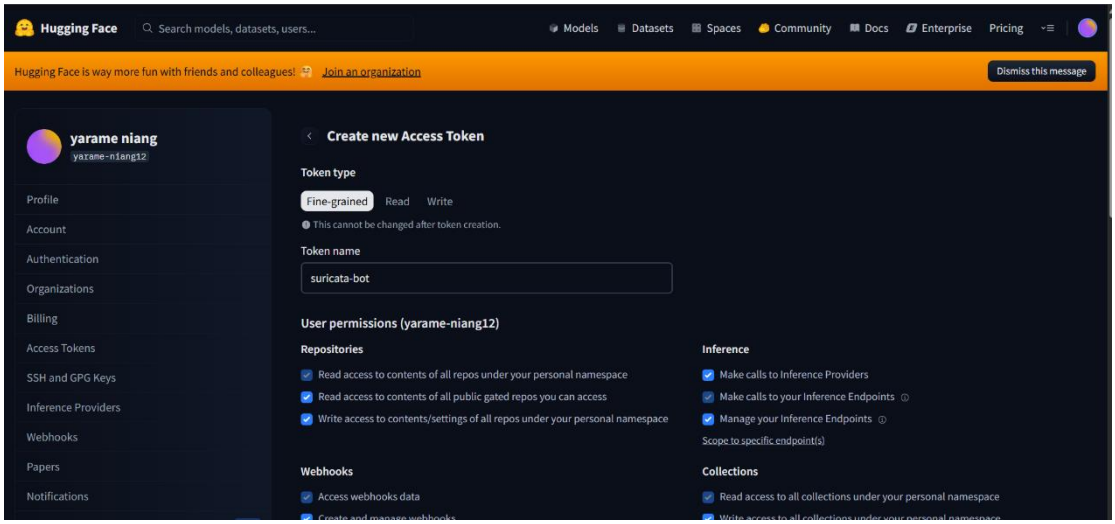


Figure 32: Obtention de la clé d'accès sur Hugging Face

**Un prompt de corrélation :**

Un template standardisé est utilisé pour structurer le raisonnement du modèle. Ce prompt présente deux alertes et demande au LLM de répondre uniquement YES ou NO.

**Une API REST FastAPI :**

Le module expose un endpoint /correlate_batch permettant d'envoyer un lot d'alertes. L'API retourne ensuite une liste de groupes corrélés, c'est-à-dire des ensembles d'alertes que le modèle juge appartenir au même incident.

Notre code développé en python présente 5 compartiments dont :

**Chargement du modèle :**

Le modèle TinyLlama/TinyLlama-1.1B-Chat-v1.0 est un petit modèle de langage basé sur l'architecture Transformer, conçu pour comprendre du texte et générer des réponses de manière conversationnelle. Il fonctionne en découpant d'abord l'entrée en tokens, puis en analysant ces tokens grâce au mécanisme d'attention, qui lui permet d'identifier les liens et relations entre les mots. Ensuite, il prédit progressivement les mots suivants pour construire une réponse cohérente. Cette version "Chat" a été entraînée spécifiquement sur des dialogues et des instructions, ce qui lui permet de suivre des consignes, d'expliquer, de résumer et d'interpréter des données. Bien que léger (1,1B de paramètres), il reste suffisamment performant pour des tâches d'analyse et

d'automatisation comme la corrélation d'alertes, tout en étant optimisé pour fonctionner sur des serveurs avec peu de ressources.

Dans cette partie, le tokenizer transforme un texte en tokens puis indique que le modèle génère la réponse. La ligne device_map="auto" place automatiquement le modèle sur le GPU si disponible.

```
model_path = "TinyLlama/TinyLlama-1.1B-Chat-v1.0"

print("Chargement du tokenizer...")
tokenizer = AutoTokenizer.from_pretrained(model_path)

print("Chargement du modèle...")
model = AutoModelForCausalLM.from_pretrained(
    model_path,
    torch_dtype=torch.float16 if torch.cuda.is_available() else torch.float32,
    device_map="auto"
)
```

Figure 33: Chargement du modèle TinyLlama

**Le prompt de corrélation :**

Cette partie reste simple mais stricte, il impose un format unique pour toutes les requêtes, une réponse fermée (YES/NO), un raisonnement basé sur le contenu textuel des deux alertes, Ceci permet au LLM de se concentrer uniquement sur la question de corrélation.

```
PROMPT_TEMPLATE = """
<s>
[INST]
Check correlation for these two Suricata alerts. For this task, you need to:

• Examine the content and attributes of the alerts to understand their nature and potential correlation.
• Identify similarities such as source IP, destination IP, timestamp, signature ID, and other relevant metadata.
• Utilize correlation techniques such as temporal analysis, statistical analysis, and behavioral analysis to determine the likelihood of correlation.
• Document the findings and analysis performed to support the determination of correlation.
• Respond with 'Yes' if sufficient evidence suggests a correlation; otherwise, respond with 'No'.

Here are the alerts:
alert1:
(content of first alert)

alert2:
(content of second alert)
[/INST]
LABEL (YES or NO)
</s>
"""
```

Figure 34: Définition du prompt de corrélation

**Fonctionnement du module de corrélation (pair à pair) :**

Ici se fait la construction du prompt contenant les deux alertes, l'encodage des données avec le tokenizer, la génération LLM avec une température très basse (0.1) pour garantir une réponse stable, le décodage de la sortie du modèle et la détection de "YES" ou "NO". Le modèle agit donc comme un classifieur binaire intelligent.

```python
def correlate_pair(alert1, alert2):
    prompt = PROMPT_TEMPLATE.format(alert1=str(alert1), alert2=str(alert2))

    inputs = tokenizer(prompt, return_tensors="pt").to(model.device)

    output = model.generate(
        **inputs,
        max_new_tokens=50,
        temperature=0.1
    )

    decoded = tokenizer.decode(output[0])

    if "YES" in decoded:
        return True
    elif "NO" in decoded:
        return False
    else:
        return False  # fallback safe
```

Figure 35: Module de corrélation pour 2 alertes

**Corrélation globale sur une liste d'alertes :**

Ce module, toujours piloté par le LLM, parcourt toutes les alertes non encore groupées, pour chaque alerte, il teste la corrélation avec toutes les suivantes, dès qu'une alerte est corrélée, elle est ajoutée au même groupe, les alertes déjà groupées sont ignorées pour éviter les doublons et à la fin, il retourne une liste de groupes d'alertes corrélées.

```python
def correlate_alerts_list(alerts: list):
    """
    alerts : liste de dictionnaires JSON représentant les alertes.
    Retourne une liste de groupes corrélés.
    """

    correlated_groups = []
    already_grouped = set()  # indices des alertes déjà regroupées

    for i, alert in enumerate(alerts):

        if i in already_grouped:
            continue

        group = [alert]
        already_grouped.add(i)

        for j in range(i + 1, len(alerts)):
            if j in already_grouped:
                continue

            correlated = correlate_pair(alert, alerts[j])

            if correlated:
                group.append(alerts[j])
                already_grouped.add(j)

        correlated_groups.append(group)

    return correlated_groups
```

Figure 36: Module de corrélation pour un lot d'alertes

API REST FastAPI :

Notre API accessible grâce via l'URL http://51.75.73.48:8000/docs#/default/correlate_batch_correlate_batch_post ,expose

deux endpoints, GET / qui retourne un message indiquant que le module est opérationnel et POST /correlate_batch qui prend en entrée un JSON contenant une liste d'alertes. Ce dernier retourne des groupes d'alertes corrélées.

```python
app = FastAPI()

class BatchAlerts(BaseModel):
    alerts: list  # Liste de JSON

@app.get("/")
def home():
    return {"message": "Module de corrélation LLM opérationnel !"}

@app.post("/correlate_batch")
def correlate_batch(data: BatchAlerts):
    groups = correlate_alerts_list(data.alerts)

    # Convertir les groupes en JSON naturel
    json_groups = []
    for g in groups:
        json_groups.append(g)

    return {"groups": json_groups}
```

Figure 37: Structure de l'API

Le LLM joue le rôle d'un **analyseur expert** capable de comprendre le sens des alertes. Il corrèle les alertes en fonction de la :

- **Corrélation sémantique :** analyse de la description textuelle de l'alerte.
- **Corrélation logique :** Comparaison des IP, ports, protocoles, timestamps.
- **Corrélation contextuelle :** Compréhension de relations implicites entre des alertes corrélées.
- **Corrélation temporelle :** Lien entre événements rapprochés dans le temps. Grâce à ces capacités, le modèle peut corréler des alertes que des règles classiques ne détecteraient pas.

Notre interface ci présente est générée automatiquement par le Framework moderne FastAPI utilisé pour développer des API de manière pragmatique. Cet outil nous propose une interface dynamique permettant de consulter, tester et valider les différents endpoints exposés par notre module de corrélation.

L'interface est normalisée pour l'exposition des services, son en-tête indique les versions de FastAPI et d'OpenAPI utilisées. Cette conformité garantit une interopérabilité élevée avec les autres composants du système, tels que le SIEM et le SOAR. Grâce à cette documentation générée automatiquement, l'API est exploitable

aussi bien par des outils automatisés que par des analystes humains, ce qui facilite grandement l'intégration dans une architecture de sécurité moderne.

Deux endpoints principaux sont affichés :

**GET / :** Permet de tester la disponibilité du service. Il sert généralement de point d'entrée simple pour vérifier que l'API répond correctement.

**POST /correlate_batch :** Représente le cœur fonctionnel de notre module de corrélation. Cet Endpoint reçoit un ensemble d'alertes au format JSON (BatchAlerts) et déclenche le moteur de corrélation. Celui-ci analyse chaque alerte et détermine les relations possibles avec les autres (similitudes d'IP, de ports, de signatures, de timestamps, etc.).

La partie Schémas met en évidence les modèles de données utilisés par l'API :

**BatchAlerts :** structure décrivant l'ensemble des alertes envoyées au moteur de corrélation.

**ValidationError/HTTPValidationError :** schémas d'erreurs générés automatiquement lorsque les données soumises ne respectent pas les formats attendus.

Cette définition explicite des modèles apporte une rigueur technique, obligeant le client à respecter la structure demandée. Cela renforce la fiabilité du service et réduit les erreurs lors de l'intégration dans un pipeline de traitement d'alertes.

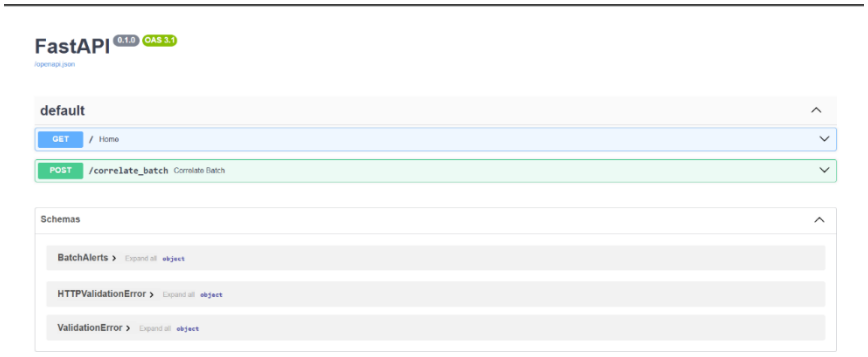



Figure 38: Interface FastAPI

**IV. Système complet de corrélation intelligente et d'automatisation de la réponse**

Afin d'effectuer notre système complet assurant la corrélation intelligente des alertes issues du SIEM et la réponse automatique face aux incidents grâce à notre SOAR,

nous ajoutons de nouvelles fonctions dans notre script python pour que Cortex puisse répondre grâce à un de ses répondeurs, que Thehive puisse créer des cas sur les alertes corrélées et isolées et que notre module LLM charge les alertes toutes les 15 minutes depuis Elasticsearch. Nous détaillons dans la suite chaque fonction ajoutée.

**Configuration de TheHive et Cortex :**

Cette partie du code définit la configuration du SOAR, c'est-à-dire les URLs et les clés API de TheHive et Cortex. Ces informations permettent au script de créer de nouveaux cas d'incident dans TheHive et de déclencher automatiquement des responders dans Cortex, comme l'action SentinelOne destinée à isoler un hôte compromis.

```
# ------------------------------------------------------------
#  CONFIGURATION THEHIVE + CORTEX
# ------------------------------------------------------------

THEHIVE_URL = "http://51.75.73.48:9000"
CORTEX_URL = "http://51.75.73.48:9001"

THEHIVE_API_KEY = "0OtAeQhiAUdq1lMFwiRuTBU2QxX5EskM"
CORTEX_API_KEY = "jedn4qKH21Qumpt4aSaWN9w214i6JotW"

CORTEX_RESPONDER_ISOLATE = "SentinelOne"
```

Figure 39: Configuration Thehive+Cortex

**Configuration Elasticsearch :**

Ci-après, le code configure l'accès à Elasticsearch en précisant l'adresse du serveur, les identifiants de connexion et le nom de l'index dans lequel Suricata stocke les alertes (Suricata-*). Cette section assure la liaison avec le SIEM : elle permet de récupérer les alertes json qui serviront à l'analyse.

```
# ------------------------------------------------------------
#  CONFIG ELASTICSEARCH SÉCURISÉ (HTTP)
# ------------------------------------------------------------

ELASTIC_URL = "http://localhost:9200"
ELASTIC_USER = "elastic"
ELASTIC_PASSWORD = "QRe9OYN7ztdsi1FOKIw7"
ELASTIC_INDEX = "suricata-*"

es = Elasticsearch(
    ELASTIC_URL,
    basic_auth=(ELASTIC_USER, ELASTIC_PASSWORD)
)
```

Figure 40: Configuration d'elasticsearch

**Proposition d'actions exécutables par Cortex :**

Pour chaque groupe d'alertes corrélées, le système propose un plan d'action comme isoler les hôtes associés, faire tourner des analyses Cortex, automatiser la réponse, avertir l'équipe SOC.

```
def propose_actions_for_group(group):
    return {
        "isolate_host": list({a.get("src_ip") for a in group if "src_ip" in a}),
        "send_to_cortex_analyzer": ["VirusTotal", "IPReputation"],
        "auto_respond": True,
        "notify_soc": True
    }
```

Figure 41: Ajout du champ action

**Exécution du répondeur SantinelOnes :**

La section suivante du code contient la fonction responsable de l'appel au répondeur SentinelOne dans Cortex. Cette fonction envoie l'adresse IP concernée directement à Cortex, qui se charge ensuite de l'exécution réelle de l'isolation sur l'agent SentinelOne installé sur la machine.

```
def cortex_isolate_host(ip: str):
    url = f"{CORTEX_URL}/api/responder/{CORTEX_RESPONDER_ISOLATE}/run"
    headers = {"Authorization": f"Bearer {CORTEX_API_KEY}"}
    data = {"data": ip, "dataType": "ip"}

    try:
        r = requests.post(url, json=data, headers=headers, timeout=10)
        return r.json()
    except Exception as e:
        return {"error": str(e)}
```

Figure 42: Exécution du répondeur de cortex

**Création d'une case dans TheHive :**

Une fois les actions effectuées, le script génère un nouveau cas dans TheHive. Le contenu du cas regroupe les alertes corrélées, les détails des actions réalisées et les résultats renvoyés par les responders Cortex. Cela permet aux analystes SOC d'avoir une vision complète de l'incident, avec la chronologie des événements et les mesures déjà prises automatiquement par le système.

```python
def create_thehive_case(group, actions, cortex_results):
    url = f"{THEHIVE_URL}/api/v1/case"
    headers = {"Authorization": f"Bearer {THEHIVE_API_KEY}"}

    description = "Correlated Alerts:\n\n" + "\n".join(str(a) for a in group)
    description += "\n\nIsolation Results:\n" + str(cortex_results)

    data = {
        "title": "Corrélation automatique + isolation SentinelOne",
        "description": description,
        "severity": 2,
        "tlp": 2,
        "pap": 2,
        "tags": ["LLM", "Correlation", "SOAR", "Isolation", "SentinelOne"],
    }

    try:
        r = requests.post(url, json=data, headers=headers)
        return r.json()
    except Exception as e:
        return {"error": str(e)}
```

Figure 43 : Création d'un case thehive

**Récupération des alertes du SIEM :**

Le script inclut également une fonction qui interroge Elasticsearch pour récupérer les alertes. Ce filtrage temporel assure que seules les alertes récentes sont analysées et que l'automatisation se fait en quasi-temps réel.

```python
def fetch_recent_alerts():
    now_utc = datetime.now(timezone.utc)
    fifteen_mins_ago = now_utc - timedelta(minutes=15)
    query = {
        "size": 1000,
        "query": {
            "range": {
                "timestamp": {
                    "gte": fifteen_mins_ago.isoformat(),
                    "lte": now_utc.isoformat()
                }
            }
        }
    }
    res = es.search(index=ELASTIC_INDEX, body=query)
    alerts = [hit["_source"] for hit in res["hits"]["hits"]]
    return alerts
```

Figure 44: Récupération des alertes depuis Elasticsearch

**Tâche de récupération périodique :**

Une tâche asynchrone tourne en continu : toutes les quinze minutes, elle récupère les nouvelles alertes.

```python
async def periodic_alert_fetch():
    while True:
        try:
            alerts = fetch_recent_alerts()
            if alerts:
                groups = correlate_alerts_list(alerts)
                for group in groups:
                    actions = propose_actions_for_group(group)
                    cortex_results = [cortex_isolate_host(ip) for ip in actions["isolate_host"]]
                    create_thehive_case(group, actions, cortex_results)
            print(f"[{datetime.now(timezone.utc).isoformat()}] Checked {len(alerts)} alerts.")
        except Exception as e:
            print(f"[ERROR] {e}")
        await asyncio.sleep(900)  # 15 minutes
```

Figure 45 : Définition de la récupération périodique

Toutes ces étapes établies, nous obtenons un système qui fonctionne comme une chaîne entièrement intégrée. Le SIEM constitue la source principale d'informations : Suricata détecte les événements réseau et les enregistre dans Elasticsearch. Toutes les quinze minutes, le module Python vient récupérer ces alertes et les transmet au modèle TinyLlama. Celui-ci analyse leur contenu et détermine lesquelles sont liées entre elles, formant ainsi des incidents cohérents plutôt qu'une simple liste d'événements isolés. Une fois les groupes formés, les actions automatisées sont déclenchées via Cortex. Les responders, comme SentinelOne, appliquent directement des mesures de défense, par exemple l'isolation d'un poste compromis. Enfin, toutes les informations relatives à l'incident alertes corrélées, détails des actions, résultats des responders sont regroupées dans un nouveau case TheHive, permettant au SOC d'avoir une vision complète, structurée et prête à être investiguée. Ainsi, le SIEM détecte, le LLM corrèle intelligemment, et le SOAR exécute les réponses et documente l'incident : les trois composantes interagissent en complémentarité pour former un système de défense automatisé et intelligent.

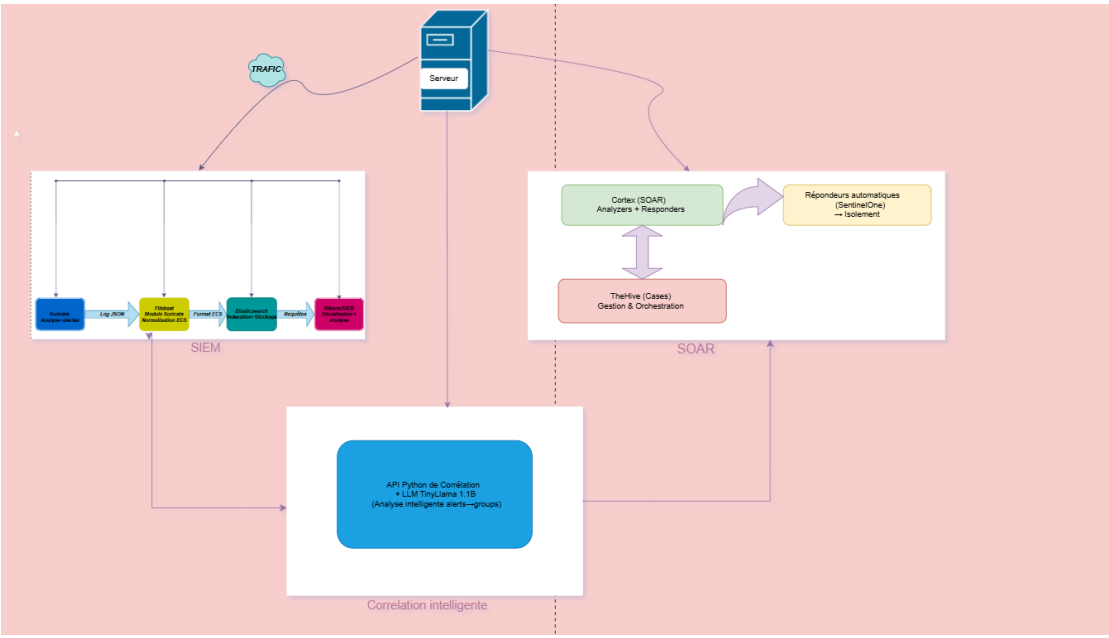



Figure 46: Système intelligent de corrélation et d'automatisation de la réponse

**Conclusion**

L'implémentation technique présentée dans ce chapitre a permis de construire une architecture complète, cohérente et opérationnelle pour la détection, l'analyse et l'automatisation de la réponse aux incidents de sécurité. L'intégration du SIEM basé sur Elastic Stack a d'abord fourni une base solide pour la collecte centralisée, la normalisation et la visualisation des événements réseau. Elle assure une surveillance continue et une traçabilité fiable des activités.

L'ajout du SOAR, via TheHive et Cortex, est ensuite venu enrichir cette architecture en apportant des capacités avancées d'orchestration et d'automatisation, indispensables pour réduire le temps de réaction face aux menaces.

La conception et l'intégration du module de corrélation intelligente, fondé sur un modèle de langage TinyLlama, ont constitué une étape déterminante. Ce composant introduit une couche d'analyse contextuelle capable d'identifier automatiquement les relations entre différentes alertes. Il dépasse ainsi les limites des approches classiques qui reposent sur des règles statiques. Cette capacité de corrélation sémantique améliore la compréhension globale des incidents et accroît la précision du triage.

Enfin, la mise en place du système complet a permis d'interconnecter les trois éléments SIEM, LLM et SOAR afin de former une chaîne entièrement automatisée allant de la détection des événements jusqu'à la réponse active. Grâce à cette intégration, le système peut désormais récupérer les alertes, les analyser de manière intelligente,

déclencher des actions de remédiation via Cortex et créer automatiquement un dossier d’incident dans TheHive.

## Conclusion Générale

Au terme de ce travail, nous avons démontré qu'il est aujourd'hui possible de dépasser largement les approches traditionnelles de supervision en cybersécurité en intégrant une intelligence capable d'interpréter les événements, d'établir des liens pertinents entre eux et d'automatiser de manière raisonnée la réponse aux incidents.

L'objectif de ce mémoire était de développer un système intelligent de corrélation et d'automatisation de la réponse aux incidents, fondé sur l'analyse des logs et des alertes. Les différents chapitres ont posé les bases théoriques, techniques et méthodologiques nécessaires à la construction d'une telle solution.

Le premier chapitre a rappelé les fondements essentiels de la cybersécurité et a mis en évidence les limites des approches classiques face à la sophistication croissante des menaces. C'est dans ce contexte que l'intelligence artificielle trouve sa pertinence, car elle apporte une capacité d'analyse avancée, une rapidité d'exécution et une compréhension contextuelle qui améliorent de manière significative l'efficacité opérationnelle des équipes de défense.

Le deuxième chapitre a approfondi les concepts clés sur lesquels s'appuie la solution développée. Il a présenté le SIEM utilisé pour la collecte, l'agrégation et la normalisation des événements, la corrélation indispensable pour réduire le bruit et comprendre la logique des attaques, et enfin le SOAR qui permet de standardiser les workflows, d'automatiser les actions et de réduire le temps de réponse. L'étude de ces technologies a permis d'identifier leur complémentarité et de poser les bases d'une architecture intégrée, modulable et orientée vers l'automatisation.

Le troisième chapitre a décrit la conception du mécanisme de développement du système intelligent. Il a détaillé la méthodologie adoptée pour construire une solution capable d'analyser automatiquement des alertes, d'en extraire le sens, de produire une corrélation intelligente à l'aide d'un modèle de langage, puis de proposer ou de déclencher des actions automatisées via le SOAR. La modélisation des composants, les choix technologiques et la structure de la communication entre les modules ElasticStack, API de corrélation, Cortex et TheHive, ainsi que SentinelOne, ont permis de définir une architecture cohérente, évolutive et extensible.

Enfin, l'implémentation technique est venue concrétiser l'ensemble de cette conception. Elle a inclus la mise en place d'un SIEM basé sur ElasticStack, le

déploiement d'un SOAR avec Cortex et TheHive, le développement d'une API de corrélation intelligente capable de traiter des lots d'alertes, d'éviter les problèmes de mémoire et de fournir des analyses contextualisées, ainsi que l'intégration d'actions automatiques de réponse comme l'isolement d'un hôte au moyen du responder SentinelOne. L'ensemble aboutit à un système opérationnel capable de détecter, d'analyser et de répondre automatiquement aux incidents.

En définitive, ce mémoire montre que l'association IA, SIEM et SOAR constitue aujourd'hui une voie incontournable pour renforcer la posture de sécurité des organisations. Elle permet une meilleure compréhension des menaces, une prise de décision plus rapide et une réduction significative de la charge opérationnelle. Les résultats obtenus ouvrent plusieurs perspectives d'évolution : amélioration continue du modèle, extension des types d'incidents pris en charge, intégration de nouvelles sources de données, ou ajout de mécanismes d'apprentissage destinés à renforcer la précision de la corrélation.

Ainsi, ce travail apporte une contribution concrète à la mise en place d'un système moderne, intelligent et automatisé de défense cyber, capable de s'adapter aux exigences actuelles des environnements numériques.

**Bibliographie**

[1]https://fr.wikipedia.org/wiki/S%C3%A9curit%C3%A9_des_syst%C3%A8mes_d%27information

Consulté le 08 décembre 2025

[2] https://www.okta.com/fr-fr/identity-101/privacy-vs-security/

Consulté le 08 décembre 2025

[3] https://icesi.fr/images/blog/3_ddos_illustration.jpg

Consulté le : 12 décembre 2025

[4]https://encryptedtbn0.gstatic.com/images?q=tbn:ANd9GcTVgglOFKacvpI8DpYjdITgVF5UH1AeSYtQCA&s

Consulté le : 12 décembre 2025

[5] https://www.pandasecurity.com/fr/mediacenter/src/uploads/2022/04/man-in-the-middle-attack-FR_02-Phases-of-man-in-middle-attack.png

Consulté le : 12 décembre 2025

[6] https://www.cloudflare.com/img/learning/security/threats/sql-injection-attack/sql-injection-infographic.png

Consulté le : 12 décembre 2025

[7] https://www.fortinet.com/content/dam/fortinet/images/cyberglossary/insider-threats.jpg

Consulté le : 12 décembre 2025

[8] https://www.wixxim.fr/uploads/images/fonctionnement-pare-feu.jpg

Consulté le : 12 décembre 2025

[9] https://media.licdn.com/dms/image/v2/D4E22AQGDAy_6WdQbVg/feedshare-shrink_800/B4EZQE2YnSGcAw/0/1735248153775?e=2147483647&v=beta&t=Xw7zWeA8FIVCp5ObGnU11rMTZkXmq100xYvVP41LFNY

Consulté le : 12 décembre 2025

[10]https://encryptedtbn0.gstatic.com/images?q=tbn:ANd9GcQlT1MM50Hcwz0RQTJ-Q9wb0FzE3BYOPVhDSQ&s

Consulté le : 12 décembre 2025

[11]https://encryptedtbn0.gstatic.com/images?q=tbn:ANd9GcRe62_25j30hSFjPYnMM3XzeC9_c3r5iH_ddg&s

Consulté le : 12 décembre 2025

[12]https://www.getronics.com/nitropack_static/UWNfeAMAKKXlWKpzbejLWlUzjkXSZYXd/assets/images/optimized/rev-621e644/www.getronics.com/wp-content/uploads/2024/04/types-of-ai-1024x892.png

Consulté le : 12 décembre 2025

[13]https://miro.medium.com/v2/resize:fit:1400/1*w81lhgWLhiDhyJrAwo6QKA.png

Consulté le : 12 décembre 2025

[14]https://csirt-universitaire.sn/

Consulté le : 15 novembre 2025

[15]https://www.fortinet.com/fr/resources/cyberglossary/man-in-the-middle-attack

Consulté le : 15 novembre 2025

[16]https://www.fortinet.com/fr/resources/cyberglossary/insider-threats

Consulté le : 15 novembre 2025

[17]https://www.fortinet.com/fr/resources/cyberglossary/artificial-intelligence-in-cybersecurity

Consulté le : 15 novembre 2025

[18]M.Diop, année académique 2024-2025, Cours droit ntic

[19]https://www.elastic.co/fr/what-is/siem

Consulté le : 10 novembre 2025

[20]https://www.mailinblack.com/ressources/glossaire/quest-ce-quun-siem/

Consulté le : 12 décembre 2025

[21]https://www.rubrik.com/fr/insights/cyber-threat-intelligence

Consulté le : 16 novembre 2025

[22]https://www.howtoforge.com/how-to-install-and-configure-suricata-ids-along-with-elastic-stack-on-ubuntu-22-04/

Consulté le : 10 novembre 2025

[23] https://www.elastic.co/fr/elasticsearch

Consulté le : 10 novembre 2025

[24]https://www.ionos.fr/digitalguide/web-marketing/analyse-web/tutoriel-kibana/

[25] IPCCC_et_CSNET, Mouhamadou Lamine Diakhame∗, Cherif Diallo∗, Mohamed Mejri, ∗Gaston Berger University, Saint-Louis, Senegal

[26]"MCM-CASR: Novel Alert Correlation Framework for Cyber Attack Scenario Reconstruction Based on NLP, NER, and Semantic Similarity", M. Diakhame, C. Diallo, and M. Mejri in 2023 7th Cyber Security in Networking Conference (CSNet), IEEE, 2023

[27] Systematic Literature Review of Security Event, Correlation Methods, IGOR KOTENKO, (Senior Member, IEEE), DIANA GAIFULINA, AND IGOR ZELICHENOK. St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), St. Petersburg Institute for Informatics and Automation of the Russian, Academy of Sciences (SPIIRAS), 199178 St. Petersburg, Russia

[28] Systematic Literature Review of Security Event Correlation Methods, IGOR KOTENKO, DIANA GAIFULINA, IGOR ZELICHENOK, 2022

[29] Mise en œuvre de la gestion des informations et des événements de sécurité (SIEM) (Bibliothèque Network Pro), 1re édition par David R. Miller (Auteur), Shon Harris (Auteur), Allen Harper (Auteur), Stephen VanDyke (Auteur), Chris Blask (Auteur)

[30] https://www.paloaltonetworks.fr/cortex/xsoar-state-of-soar-report-2020

[31] Systèmes de Détection D'intrusion : Réduction des Faux Positifs à l'aide de la Corrélation des Événements, El mostapha CHAKIR, Youness IDRISSI KHAMLICHI, Mohammed MOUGHIT

[32]https://www.verizon.com/business/resources/fr/reports/2023-dbir-executive-summary.pdf

Consulté le : 25 novembre 2025