

Détecter : de la surveillance à l'évènement de sécurité

Eric DUPUIS^{1,2*}

🕒 Résumé

Ce document fournit les fondamentaux de la gestion de la menace et de sa détection. Il fait partie du cours introductif aux fondamentaux de la sécurité des systèmes d'information vue sous deux prismes quelques fois opposés dans la littérature : la gouvernance et la gestion opérationnelle de la sécurité. Le cours est constitué d'un ensemble de notes de synthèse indépendantes compilées en un document final unique. Ce document ne constitue pas à lui seul le référentiel du cours. Il compile des notes de cours mises à disposition de l'auditeur comme support pédagogique.

🔑 Mots clefs

Évènements, attaques, détection, SIEM, SOC

¹Enseignement sous la direction du Professeur Véronique Legrand, Conservatoire National des Arts et Métiers, Paris, France

²RSSI Orange Cyberdefense

*email : eric.dupuis@lecnam.net – eric.dupuis@orange.com

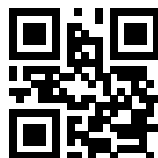
DRAFT NOTES S2 - 2020

Vérifiez la disponibilité d'une version plus récente de

SEC101-C3-ThreatMan.doc.pdf sur GITHUB CYBERDEF [↗](https://github.com/edufaction/CYBERDEF/raw/master/Builder/SEC101-C3-ThreatMan.doc.pdf)¹



2020 eduf@ction Publication en Creative Common BY-NC-ND



1. <https://github.com/edufaction/CYBERDEF/raw/master/Builder/SEC101-C3-ThreatMan.doc.pdf>



1. GERER les menaces

Je vous propose d'aborder ce chapitre lié à la surveillance de l'évènement de sécurité avec les quelques points fondamentaux de la gestion de la menace. En effet nous partons du principe qu'au sein de périmètre de surveillance donné, la gestion de la menace peut s'organiser autour de la gestion des événements à risques détectés dans ce périmètre.

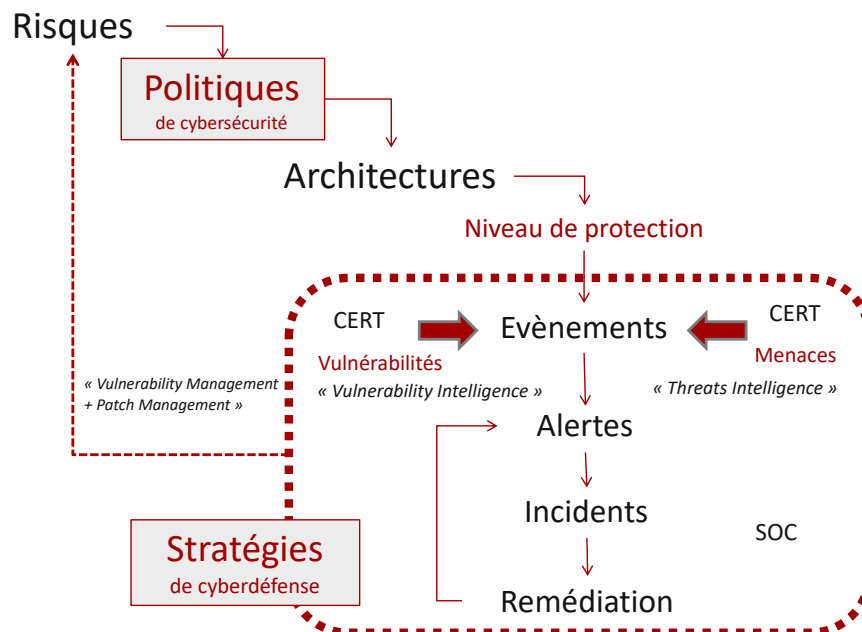


FIGURE 1. Cycle de vie de gouvernance Cyberdef

Après quelques définitions et positionnement dans l'analyse de menace, de la supervision et de l'analyse comportementale nous aborderons donc les grandes fonctions nécessaire à la **détection** comme :

- ▶ **VOIR** : capacité de voir et de capter le comportement d'un système d'information via des sources et capteurs avec le *LOG management* (Systèmes et Applicatifs). En n'oubliant pas d'évoquer l'assurance sécurité des Logs (intégrité, horodatage, valeur probante ...)
- ▶ **COMPRENDRE - PREVOIR** : Avec le *Threat Management* : Veiller, surveiller la menace dans l'environnement digital de l'entreprises, modélisation de la menace et scénarios redoutés issus d'analyse de risque ;
- ▶ **DETECTER** : Surveiller le comportement des systèmes dans le périmètre défini, faire émerger les événements, anomalies, incidents pouvant révéler une attaque en cours, une suspicion de compromission par des menaces avancées (APT), où des attaques furtives et discrètes. Nous aborderons l'outillage avec les SIEM et l'organisation avec les SOC ;



- **ALERTER** : mettre en place les mécanismes de remontée d'alerte et d'incident permettant de gérer les alertes adaptées au niveau d'impact d'une attaque.

Menaces=Veille et recherche : La gestion de la menace est au coeur des stratégies de cyberdéfense de l'entreprise. Comme pour les vulnérabilités, c'est la connaissance des menaces, de leur recherche et de leur découverte qui permet de réduire les risques ;

Menaces=Évènements : La détection d'une vulnérabilité ou d'une menace est un évènement, la question est de savoir à quel moment il est important de déclencher un mécanisme d'alerte, et comment cette alerte va devenir un incident déclenchant des mécanismes de réponse (Voir Cycle de gouvernance ?? page ??).

1.1 Modèles

Ce sont généralement les attaques externes et massives qui font l'actualité dans les médias. Un grand nombre de risques cyber quotidiens sont issus de l'intérieur même de l'entreprise . Des fuites de la part d'employés qui, de façon intentionnelle ou non, révèlent des mots de passe ou des informations sensibles. Une opération initiée par des acteurs internes malveillants (salariés, partenaires, clients) qui peuvent utiliser les informations à leur portée afin d'exploiter ou de porter dommage aux systèmes d'information de l'entreprise, mais plus globalement à l'entreprise dans sa globalité.

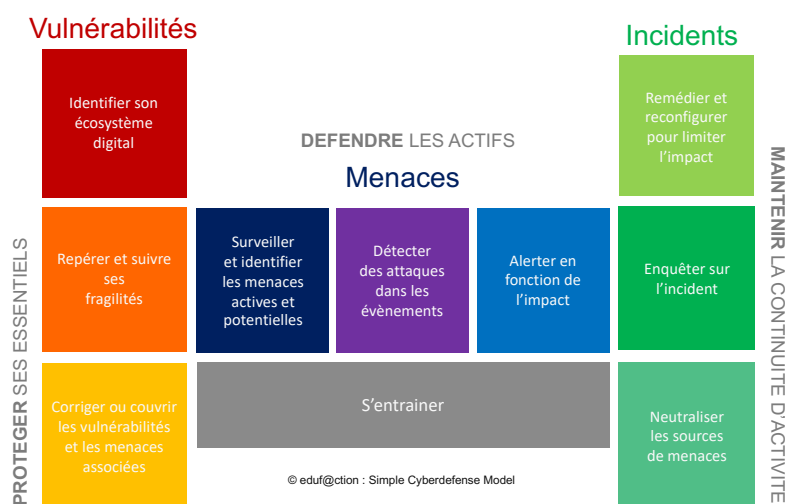


FIGURE 2. Un modèle de gestion cyberdéfense

N'importe quelle entreprise est exposée au risque. De nos jours, les connexions diverses d'une entreprise représentent de nombreuses voies pour des attaques informatiques, qui ciblent souvent les petites entreprises afin d'accéder à des plus grands acteurs (partenaires, clients ou fournisseurs). Les grandes entreprises imposent de plus en plus souvent à leurs fournisseurs et partenaires, quelle que soit leur taille, de mettre en place des mesures



de cybersécurité. Ceci est généralement structuré autour de plan d'assurance sécurité, annexé au contrat.

Les attaquants externes sont, certes, des menaces croissantes car elles recherchent sans cesse des failles de sécurité afin d'accéder à vos systèmes, mais les menaces internes ne sont pas à négliger sur le plan opérationnel. On verra d'ailleurs que pour couvrir ces menaces, la détection des événements liés à des comportements de ses propres usagers ou salariés n'est pas chose facile (Législation sur le droit des correspondance, Commission National Informatique et Liberté (CNIL) ...).

Il est important toutefois de ne pas distinguer des dynamiques internes et externes en terme d'agression cyber, c'est pour cela que les modèles de Cyberdéfense ne formalisent que très peu cette notion de provenance de l'attaque.

En outre, sur le triptyque « Vulnérabilités, Menaces, Incidents » (Voir Modèle de cyberdéfense ?? page ??), la notion d'attaque interne et/ou externe n'a de sens qu'au titre de la responsabilité du périmètre, car rien n'interdit de positionner des capteurs en dehors de son périmètre technique pour anticiper la menace. Ainsi, la veille sur internet, le renseignement cyber sont nécessairement équipés de capteurs pouvant faire remonter des événements dans des chaînes d'alertes Cyber. Par exemple, l'attaque d'une entreprise du même secteur peut être en soit un incident pouvant générer une alerte.

Si on se focalise sur la gestion de la menace, il existe de nombreuses manières de présenter ce processus. Un modèle (Voir Threat Management Cycle ?? page ??) issu des travaux de l'Organisation du Traité de l'Atlantique Nord (OTAN) est intéressant car il propose 4 axes d'analyses, toutefois Threat Mitigation et Threat Protection sont un peu ambiguës.

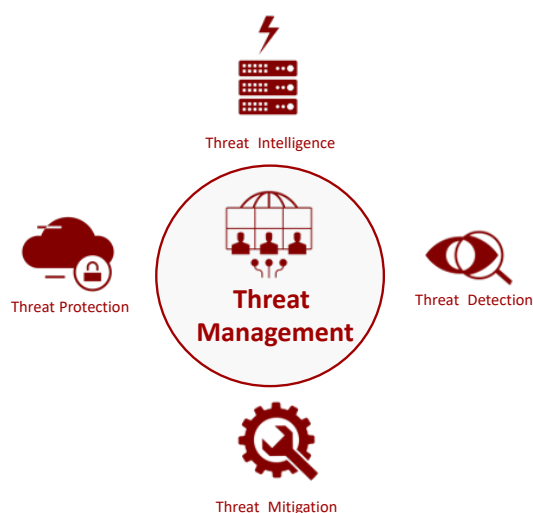


FIGURE 3. les 4 axes de la gestion de la menace

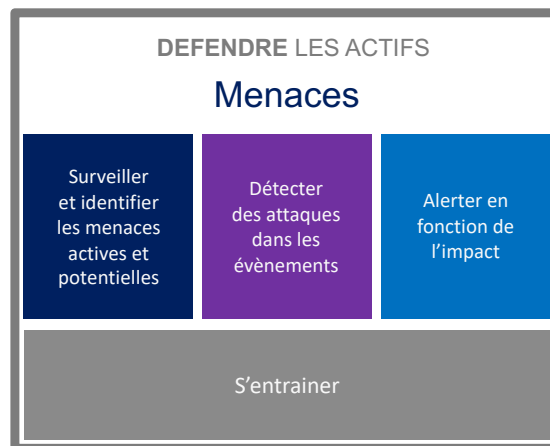
Je propose de continuer d'utiliser le modèle organiser autour de quatre volets (Surveiller,



Détecter, Alerter, s'entraîner) (Voir Défendre les actifs ?? page ??) pour structurer la présentation.

La nature de ces menaces est en constante évolution. On y trouve les plus courantes :

- ▶ Attaques par **déni de service distribuées** (DDoS). Un réseau d'ordinateurs inonde un site Web ou un logiciel avec des informations inutiles. L'exemple, le plus classique est celui d'un serveur WEB. Quand la charge sur les services est trop importante et que le système n'est pas dimensionné ou filtré pour ce type de volume de demande, ce débordement de requêtes provoque une indisponibilité du système inopérant.
- ▶ **Codes malveillants** : Bots et virus. Un logiciel malveillant qui s'exécute à l'insu de l'utilisateur ou du propriétaire du système (bots), ou qui est installé par un employé qui pense avoir affaire à un fichier sain (cheval de Troie), afin de contrôler des systèmes informatiques ou de s'emparer de données.



© eduf@ction : Simple Cyberdefense Model

FIGURE 4. la gestion de la menace

- ▶ **Piratage**. Lorsque des acteurs externes exploitent des failles de sécurité afin de contrôler vos systèmes informatiques et voler des informations, en utilisant ou pas un code malveillant. Par exemple, un changement régulier des mots de passe et la mise à niveau des systèmes de sécurité est fondamentale pour limiter les impacts.
- ▶ **Hameçonnage** ou dévoiement. Tentative d'obtenir des informations sensibles en se faisant passer frauduleusement pour une entité digne de confiance. Le hameçonnage se fait généralement par e-mail, mais il ne faut pas oublier les SMS et les services utilisant du message (Webmail, mail intégré comme LinkedIn, ...),

C'est la combinaison d'actions élémentaires, d'attaques élémentaires qui font des scénarios de menaces.



Il est important aussi de repositionner la définition des menaces par rapport à la notion d'attaque, mais aussi la notion de risques et de vulnérabilités (Voir Cycle du risque ?? page ??).

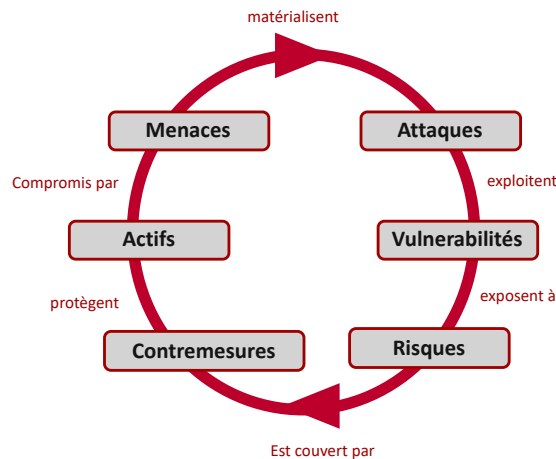


FIGURE 5. la gestion de la menace

- ▶ Des attaques **matérialisent** des menaces,
- ▶ des menaces **exploitent** des vulnérabilités,
- ▶ des vulnérabilités **exposent** à des risques,
- ▶ des risques **sont convertis** par des contre-mesures,
- ▶ des **contre-mesures** protègent des actifs,
- ▶ des actifs **sont soumis** à des menaces.

Nous voyons donc ici qu'il est important de ne pas séparer en terme de gouvernance et de pilotage opérationnel de la sécurité la gestion des vulnérabilités, la gestion des menaces et la gestion des risques.

1.2 Les processus de gestion de la menace

Gérer la menace comporte deux donc domaines d'activités :

- ▶ La veille, au sens renseignement sur la menace (Threat Intelligence) ;
- ▶ La détection d'attaque, ou de menaces potentielles au sein de l'environnement (Threat Detection).

Ces deux domaines d'activités se base sur la remontée d'information et l'automatisation des détections d'évènements à risques. Pour automatiser la détection d'évènement, il donc imprimant de disposer de sources d'information de type « évènements » que des techniques historique informatiques et réseaux apportent grâce au *LOG Management* (gestion des traces et journaux). Les journaux informatiques des systèmes sont au coeur de la détection, mais il existe de nombreuses autres sources d'évènements (informations,



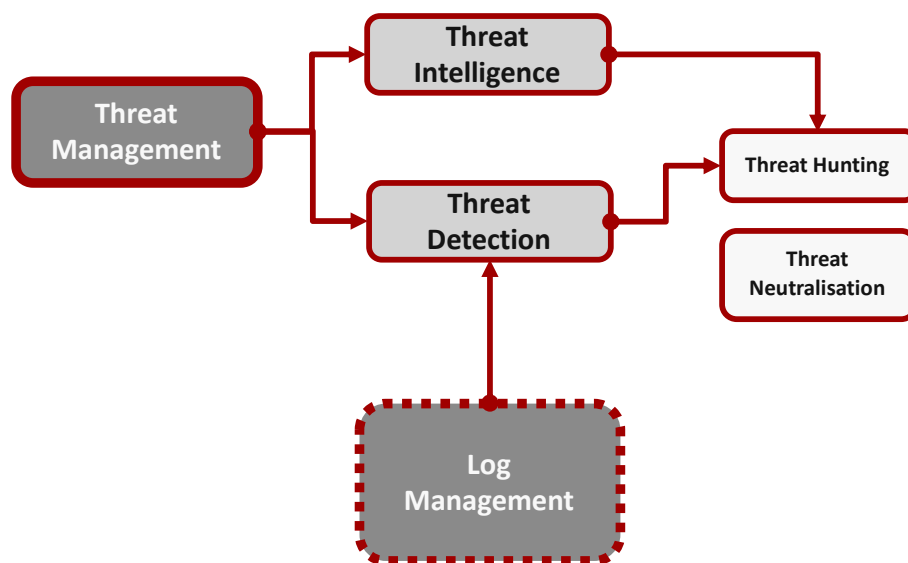


FIGURE 6. la gestion de la menace

renseignements) formalisées qui peuvent apporter de l'information pertinente pour la détection ou l'anticipation d'attaques.

1.3 Détecter, la surveillance du SI

« Détecter oui, mais détecter quoi et pourquoi » est la phrase maitresse de la première étape de réflexion autour de la gestion de la détection d'incident de sécurité. La première question à se poser est de définir ce qu'est un incident de sécurité pour l'entreprise. S'il est vrai qu'il existe un certain nombre de menace « standard » que l'on considère très rapidement comme un incident, le déploiement d'outil de gestion d'incident de sécurité ne serait être limité qu'à cet usage standard.

Il y a de nombreuses manières de détecter des tentatives d'attaques dans un système. Les IPS/IDS (*Intrusion Prevention System / Intrusion Detection System*), Firewall réseaux et firewall applicatif, ... Toutefois l'imagination des attaquants est suffisamment grande, pour que des attaques complexes ne puisse être détectées par ces seuls outils et produits de sécurité protégeant les flux informationnels.

Nous pouvons en effet considérer par exemple que la détection d'un rançon-logiciel dans l'entreprise est un incident complexe, qu'un IPS/IDS ne détectera pas, qui va par ailleurs nécessiter une alerte et une remédiation rapide si ce n'est immédiate. Par ailleurs, une fuite d'information sur un système métier par des mécanismes discrets sera souvent étudiée spécifiquement. Globalement le déploiement d'une fonction d'alerte va nécessiter la définition des « menaces » redoutées par l'entreprise. Ces dernières sont généralement issues des analyses de risques. En effet, il est important au delà des menaces dites standards de revenir aux origines du déploiement des fonctions de sécurité qui sont de gérer et couvrir les risques.



En premier lieu, il convient de chercher à détecter les menaces non couvertes par les mesures de sécurité, les fameuses menaces résiduelles.

Dans l'environnement de l'entreprise, les scénarios complexes issus de l'analyse de risques lors de l'étude des événements redoutés vont donner les événements corrélés à détecter. On y trouvera l'application concrète des arbres d'attaques popularisés par un des plus célèbre cyber expert Bruce Schneier (**schneier1999attack**) qui est présentée de manière un peu plus détaillée dans le chapitre ??

1.4 Attaques

Le cycle de vie d'une cyber-attaque, qu'elle soit complexe ou sophistiquée, reste le même depuis des années. Elle se déroule en 3 étapes :

- ▶ la première : **la phase de reconnaissance**. Elle va permettre d'identifier sa cible et de rechercher l'ensemble des vulnérabilités. Contrairement à l'audit de sécurité, dans cette étape, l'assaillant n'a aucune contrainte de périmètre ni de cadre contractuel. Cette absence de contrainte va lui permettre d'exploiter tout type de vulnérabilité ;
- ▶ la seconde : l'**attaque elle-même**. Les attaques, aujourd'hui plus sophistiquées, permettent aux assaillants, une fois entré dans le système d'information de l'entreprise, d'effectuer diverses actions comme l'élévation de privilèges, la création d'une porte dérobée, la mise en sommeil des agents dormants et surtout l'effacement de toute trace de son passage ;
- ▶ la dernière : l'« atteinte de son objectif et l'action ». Cela peut se traduire par une simple perturbation des systèmes ou encore l'exfiltration d'informations sensibles dans le but d'actes de manipulation.

Ces mécanismes sont souvent modélisés sous la forme d'arbres d'attaques ou de scénario issus d'analyses de risques.

1.4.1 Arbre d'attaques

Détecter la menace dans un système d'information c'est aussi connaître les méthodes, stratégies des attaquants. Ces scénarios d'attaque ou d'opération peuvent être modélisés avec des outils au coeur des analyses de risques. Bien que très largement en arrière plan des méthodes et des outils de gestion de la menace, les arbres d'attaque restent au coeur des mécanismes de détection.

Les arbres d'attaques sont une représentation des scénarios d'attaques. La racine représente le but final de l'attaque, les différents noeuds sont les buts intermédiaires et les feuilles les actions élémentaires à effectuer. Ces actions seront évaluées par exemple avec les potentiel d'attaque des critères communs (cf.CC et ISO)

Globalement, ces arbres sont basés sur trois types de noeuds :

- ▶ Noeud **disjonctif** OR : OU logique. Cela signifie que pour que le noeud soit réalisé, il faut qu'au moins un de ses fils soit réalisé.



- ▶ Nœud **conjonctif** AND : ET logique. Pour sa réalisation, il faut que l'ensemble de ses fils soit réalisé.
- ▶ Nœud **conjonctif séquentiel** SAND : Pour sa réalisation, il faut que l'ensemble de ses fils soit réalisé dans un ordre séquentiel c'est-à-dire les fils sont effectués les uns après les autres dans l'ordre indiqué.

En fonction de ces nœuds les valeurs des feuilles seront remontées pour obtenir le potentiel d'attaque de la racine. C'est sur la base de ce type de technique que sont construit un certain nombre d'outil de détection.

Gartner, et Lockheed Martin ont dérivé le concept de ces arbres d'attaque dans des modèles dit de « Kill Chain » issus de modèle militaire établis à l'origine pour identifier la cible, préparer l'attaque, engager l'objectif et le détruire.

Ce modèle analyse une fragilité potentielle en dépistant les phases de l'attaque, de la reconnaissance précoce à l'exfiltration des données. Ce modèle de chaîne ou processus cybercriminel aide à comprendre et à lutter contre les ransomware, les failles de sécurité et les menaces persistantes avancées (APT). Le modèle a évolué pour mieux anticiper et reconnaître les menaces internes, l'ingénierie sociale, les ransomware avancés et les nouvelles attaques.

1.4.2 Le déploiement d'une menace en 8 étapes

- ▶ **Phase 1 : Reconnaissance.** Comme dans un « casse » classique, vous devez d'abord repérer les lieux. Le même principe s'applique dans un cyber-casse : c'est la phase préliminaire d'une attaque, la mission de recueil d'informations. Pendant la reconnaissance, le cybercriminel recherche les indications susceptibles de révéler les vulnérabilités et les points faibles du système. Les pare-feu, les dispositifs de prévention des intrusions, les périmètres de sécurité (et même les comptes de médias sociaux) font l'objet de reconnaissance et d'examen. Les outils de repérage analysent les réseaux des entreprises pour y trouver des points d'entrée et des vulnérabilités à exploiter.
- ▶ **Phase 2 : Intrusion.** Après avoir obtenu les renseignements, il est temps de s'infiltrer. L'intrusion constitue le moment où l'attaque devient active : les malwares (y compris les ransomwares, spywares et adwares) peuvent être envoyés vers le système pour forcer l'entrée. C'est la phase de livraison. Celle-ci peut s'effectuer par e-mail de phishing ou prendre la forme d'un site Web compromis ou encore venir du sympathique café au coin de la rue avec sa liaison WiFi, favorable aux pirates. L'intrusion constitue le point d'entrée d'une attaque, le moment où les agresseurs pénètrent dans la place.
- ▶ **Phase 3 : Exploitation.** Le hacker se trouve de l'autre côté de la porte et le périmètre est violé. La phase d'exploitation d'une attaque profite des failles du système, à défaut d'un meilleur terme. Les cybercriminels peuvent désormais entrer dans le système, installer des outils supplémentaires, modifier les certificats de sécurité et créer de nouveaux scripts à des fins nuisibles.



- ▶ **Phase 4 : Escalade de privilèges.** Quel intérêt y a-t-il à entrer dans un bâtiment si vous restez coincé dans le hall d'accueil ? Les cybercriminels utilisent l'escalade de privilèges pour obtenir des autorisations élevées d'accès aux ressources. Ils modifient les paramètres de sécurité des GPO, les fichiers de configuration, les permissions et essaient d'extraire des informations d'identification.
- ▶ **Phase 5 : Mouvement latéral.** Vous avez carte blanche, mais vous devez encore trouver la chambre forte. Les cybercriminels se déplacent de système en système, de manière latérale, afin d'obtenir d'autres accès et de trouver plus de ressources. C'est également une mission avancée d'exploration des données au cours de laquelle les cybercriminels recherchent des données critiques et des informations sensibles, des accès administrateur et des serveurs de messagerie. Ils utilisent souvent les mêmes ressources que le service informatique, tirent parti d'outils intégrés tels que PowerShell et se positionnent de manière à causer le plus de dégâts possible.
- ▶ **Phase 6 : Furtivité, camouflage, masquage.** Mettez les caméras de sécurité en boucle et montrez un ascenseur vide pour que personne ne voit ce qui se produit en coulisses. Les cyber-attaquants font la même chose. Ils masquent leur présence et leur activité pour éviter toute détection et déjouer les investigations. Cela peut prendre la forme de fichiers et de métadonnées effacés, de données écrasées au moyen de fausses valeurs d'horodatage (time-stamping) et d'informations trompeuses, ou encore d'informations critiques modifiées pour que les données semblent ne jamais avoir été touchées.
- ▶ **Phase 7 : Isolation et Dénî de service.** Bloquez les lignes téléphoniques et coupez le courant. C'est là où les cybercriminels ciblent le réseau et l'infrastructure de données pour que les utilisateurs légitimes ne puissent obtenir ce dont ils ont besoin. L'attaque par déni de service (DoS) perturbe et interrompt les accès. Elle peut entraîner la panne des systèmes et saturer les services.
- ▶ **Phase 8 : Exfiltration.** Prévoyez toujours une stratégie de sortie. Les cybercriminels obtiennent les données. Ils copient, transfèrent ou déplacent les données sensibles vers un emplacement sous leur contrôle où ils pourront en faire ce qu'ils veulent : les rendre contre une rançon, les vendre sur eBay ou les envoyer à BuzzFeed. Sortir toutes les données peut prendre des jours entiers, mais une fois qu'elles se trouvent à l'extérieur, elles sont sous leur contrôle.

Différentes techniques de sécurité proposent différentes approches de la chaîne cyber-criminelle. De Gartner à Lockheed Martin, chacun définit les phases de manière légèrement différente.

C'est un modèle quelque peu critiqué pour l'attention qu'il accorde à la sécurité périmétrique et focalisé sur la prévention des malwares. Cependant, quand elle est combinée à l'analyse avancée et à la modélisation prédictive, la chaîne cyber-criminelle devient essentielle à une sécurité complète.



1.4.3 UBA : User Behavior Analytics

L'analyse du comportement des utilisateurs (UBA) apporte des informations détaillées sur les menaces liées à chaque phase de la chaîne criminelle. Et elle contribue à prévenir et arrêter les attaques avant que les dommages ne soient causés. En effet, le volume d'activités suspectes, dont des faux positifs, inhérents aux outils de sécurité traditionnels sont très chronophage à surveiller et donc sources d'erreurs. Pour y pallier, les entreprises doivent être en mesure d'analyser le comportement des utilisateurs, souvent via des outils d'apprentissage automatique, afin de donner un sens aux informations remontées par la lecture des activités sur le réseau.

Cette analyse du comportement des utilisateurs (UBA : User Behavior Analytics) aide à comprendre et hiérarchiser les alertes filtrant celles qui sont suspectes en comparaison avec des comportements habituel des utilisateurs.

1.4.4 La surveillance des terminaux

La surveillance des terminaux (EndPoint) est aujourd'hui un point important dans la prise en charge de la menace du côté l'utilisateur (mais aussi serveurs)

Le terme « Endpoint (Threat) Detection and Response » (EDTR ou EDR). Un système EDR met l'accent sur la détection d'activités suspectes directement sur les hôtes de traitement du système d'information au delà de l'infrastructure.

1.5 Gestion de la menace

Nous avons évoqué dans le chapitre sur l'anticipation, la veille sur la menace. Opérer la détection d'attaques ou de menaces dormantes dans l'environnement de l'entreprise nécessite une connaissance précise des mécanismes d'exécution ou d'opération de ces menaces. La connaissance de ces mécanismes d'action, de protection, de déploiement, de réplication, de survivabilité, de déplacement des codes malveillants par exemple est la base de leur détection. Il est en de même sur les scénarios mixant des actions sur les réseaux ou sur les systèmes informatiques ou numériques. Ces connaissances sont généralement structurées dans des bases de connaissances dont les sources sont gratuites ou payantes.

1.6 Bases de connaissance et menaces

1.6.1 Sources identifiées menaçantes

Nous parlerons ici de sources de menaces comme les indicateurs permettant d'identifier l'origine technique d'une menace. Cela peut être une adresse mail, un serveur/service de mail, une adresse IP de provenance d'un code malveillant, d'une attaque, ou d'un comportement anormal. On peut citer par exemple :

- ▶ Une adresse mail connue pour envoyer des codes malveillants;
- ▶ des adresses IP ou des adresses de serveur Mail pour Spam.




En face, il y a des attaquants qui bien entendu vont changer leur position pour émettre ou attaquer d'ailleurs, ou avec une autre forme (furtivité). Ces bases d'informations peuvent donc devenir rapidement obsolètes. Ceci dénote l'importance de disposer de base de connaissance sur les sources de menaces à jour et en temps réel.

1.6.2 Cibles de menaces

Les cibles de menace peuvent être connues à un instant T. Ces cibles peuvent être sectorielles (Banques, sites étatiques ...).

1.6.3 Threat Intelligence Database

Dans la notion de partage de l'information sur la menace, le projet MISP ² (Open Standards For Threat Information Sharing) fournit les modèles de données et des indicateurs.

Il ne faut pas oublier que les données dans ces bases peuvent être "éphémères". (Adresse IP malveillante, machines infectées nettoyées...). Il est donc important de disposer de sources fiables et mise à jour en temps réel.

Fiche TECHNO : Bases de Threat Intelligence

Le marché des bases publiques et commerciales de *Threat Intelligence* est un sujet intéressant pour une fiche TECHNO . On peut rechercher les sources les plus pertinentes, celles qui fournissent des informations techniques pour les SIEM, celles spécialisées dans des menaces sectorielles ou technologiques ...

1.6.4 Exemple de TI

On peut donc trouver dans ces base de threat intelligence des données du type suivant :

Nom	Command&Control URL	HASH : sha256
caracal.raceinspace.astronaut	http://api.lulquid.xyz	f1d32c17a169574369088...
com.caracal.cooking	http://api.namekitchen9.xyz/api/subscription	46e41ef7673e34ef72fb3a9718...
com.leo.letmego	http://api.leopardus.xyz/api/subscription	b21cb5ebfb692a2db1c5cbbc20e00d90a...
com.caculator.biscuitent	http://api.lulquid.xyz	734418efafd312e9b3e96adaac6f86cc1a...
com.pantanal.aquawar	http://api.pantanal.xyz	8fec77c47421222cc754b32c60794e...
com.pantanal.dressup	http://api.pantanal.xyz	64e2c905bcef400e861469e114bf...
inferno.me.translator	http://api.molatecta.icu	ebe3546208fd32d3f6a9e5daf21a7240...
translate.travel.map	http://api.nhudomainuon.xyz	f805e128b9d686170f51b1add35...
travel.withu.translate	http://api.molatecta.icu	b7670b5d9a6643a54b800b4c...
allday.a24h.translate	http://api.royalchowstudio.xyz	29f2fd6cc0f632e45dd1f15ec72985...
banz.stickman.runner.parkour	http://api.lulquid.xyz/api/	e1027b6681e93d9763f19ea7e5ab2...

TABLE 1. Exemple de données de TI (Threat Intelligence)

avec le nom du package malveillant, l'url du centre de commande et de contrôle, et le Hash (empreinte) permettant de rechercher le fichier dans les données d'un système.

2. ANTICIPER les menaces

2. <https://www.misp-project.org>



2.1 Surveiller et anticiper : Threat Intelligence

La «threat intelligence» est un service de renseignement concernant les cyber-menaces. Les solutions SIEM par exemple, possèdent leurs propres sources, il n'y a pas que les SIEM qui peuvent utiliser ces sources, il est possible de connecter de nombreux outils de détections à d'autres sources externes de «threat intelligence». De manière non exhaustive un service de «threat intelligence» fournit des éléments comme :

- ▶ **Malicious / Phishing IP / URL** : une liste d'URL utilisée pour délivrer un fichier malicieux ou procéder à une attaque par hameçonnage ;
- ▶ **Botnet C&C URL** : une liste d'URL utilisée pour héberger des serveurs de commande et contrôle de logiciel malveillant ou de réseaux de machine zombie ;
- ▶ **Malicious Hash** : liste d'empreintes de logiciels malveillants connus et ayant déjà été analysés ;
- ▶ **IP Reputation** : liste d'adresses IP suspectées dans des attaques informatiques ou en lien avec une cyber-menace (pouvant être utilisée en Black List sur les systèmes de filtrage, comme historiquement).

La surveillance et le renseignement de la menace au sens général du terme (Threat Intelligence) devrait contenir les 2 niveaux :

- ▶ Le renseignement à **vocation** cyber qui comprend toutes les analyses et information permettant d'anticiper et de caractériser une menace qui pourrait s'exprimer dans le monde numérique ;
- ▶ Le renseignement **d'origine** Cyber, dont les données techniques liées à des attaques, menaces qui permettent de configurer des systèmes de détection et de réponse.

Il est vrai qu'encore aujourd'hui parler de « threat intelligence » nous dirige systématiquement sur la deuxième assertion.

Veiller et surveiller les menaces, détecter les attaques nécessite d'analyser deux axes :

- ▶ Les menaces génériques, ou ciblant un domaine particulier (Santé, Industrie, Banque ...) que l'on trouve généralement en utilisant des technologies de « threat Intelligence » ;
- ▶ Les menaces ciblées, dont les indices d'émergence peuvent être détecter en analysant la menace ou en recherchant des indices de compromissions quand ces menaces sont actives dans le périmètre de l'entreprise. « threat Detection, Hunting ... »

et ceci de deux manières :

- ▶ Surveillance de l'écosystème de la menace (IOC, DarkWeb, Threat Intelligence...)
- ▶ Recherche de compromission, ou d'infection (Threat Hunting, ...)



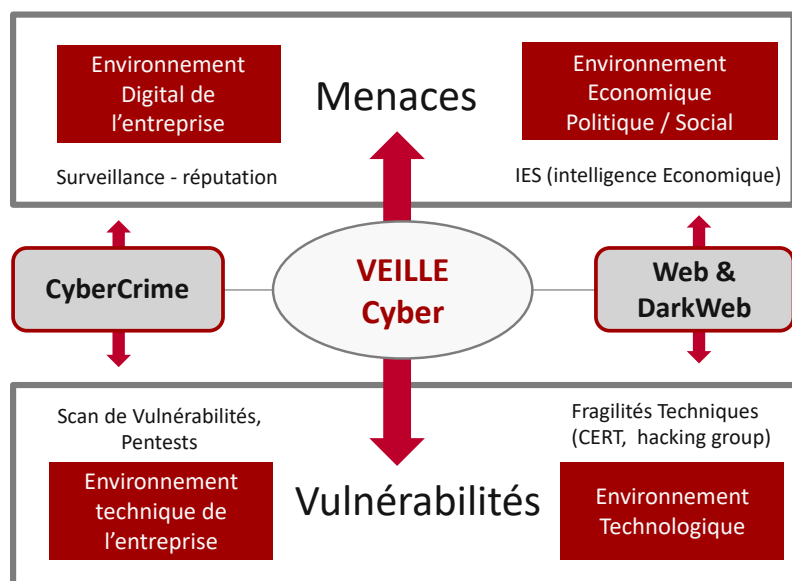


FIGURE 7. Veille cyber, une veille sur les risques

Ce sont des sujets que nous aborderons dans le processus de gestion de la menace. La surveillance des menaces génériques relève d'action de veille comme cela est fait pour les vulnérabilité. Les scénarios de message sont vus comme des éléments de signature d'une attaque ou d'une tentative ou de préparation d'attaque.

Je vous propose de présenter la gestion de la menace sous la forme de 3 thèmes ((Voir Gestion de la menace ?? page ??)).

- ▶ Log Management ;
- ▶ Threat Intelligence (au sens renseignement) ;
- ▶ Threat Detection.

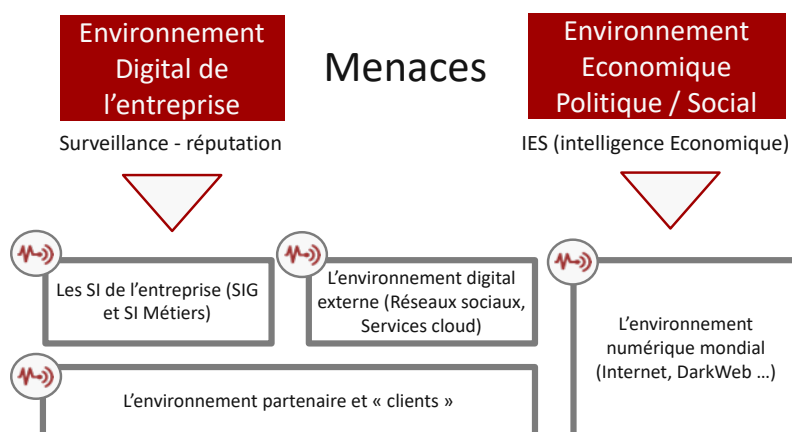


FIGURE 8. Les sources



2.1.1 Surveillance de la compromission

Un des domaines de la surveillance est donc celui de la compromission. C'est à dire la surveillance dans le fameux Darkweb de l'émergence de données volées, « perdues » par une entreprise ou par un particulier.

Ce domaine, dénommé par certains "Leak Intelligence" ou "Leak Management", correspond à la gestion des fuites de données, au sens de leur détection et la recherche de "la source" de fuite. C'est souvent dans cette dynamique de découverte d'information « interne » dans des espaces « malveillants » ou pas que l'on découvre des compromissions techniques ou non techniques issues d'attaques.

La compromission la plus connue reste encore de nos jours la fuite du couple Utilisateur/mot de passe sur des sites hackés. Il existe des bases de données publiques qui publient ces données "compromises" comme par exemple : Have i been pwned [↗](https://haveibeenpwned.com)³, qui permettent à partir d'une adresse mail ayant servie d'identifiant de savoir si on a été compromis sur un site qui aurait été piraté.

2.1.2 Surveillance des fragilités

Comme nous l'avons vu dans le chapitre sur les vulnérabilités, des « scans » de vulnérabilités sur des plages d'IP de l'entreprise permet de déterminer les fragilités de services ouverts ou accessibles. Généralement organisées dans une dynamique d'audit, ces évaluations de sécurité sont conduites avec un cadre contractuel et légal. Il existe pourtant des entreprises qui fournissent des informations de fragilités sur des entreprises ou des plages d'IP. On trouvera par exemple sur le site de SHODAN [↗](https://www.shodan.io)⁴ des informations intéressantes (et payantes) sur des fragilités de systèmes dont une grande partie de systèmes d'objets connectés.

Fiche TECHNO : Base de données de fragilités publiques

Les techniques et les sites qui proposent des outils ou l'accès à des bases de données de « sites » vulnérables. C'est un sujet intéressant pour une fiche TECHNO de synthèse sur ce qui existe sur le marché.

2.1.3 Surveillance du ciblage

La surveillance du ciblage, que les anglo-saxons appellent le « TARGETING » est aussi un élément d'anticipation. En effet, ces éléments sont souvent les premiers signaux d'un événement « cyber » qui pourrait toucher l'entreprise.

On y trouve l'émergence de la collecte d'information sur une cible donnée. La mise en oeuvre dans les codes malveillants de ciblage d'IP spécifique, etc...

Il y a deux types d'outils pour ce se faire :

- La surveillance classique du web de type « cyberveille », qui permet de découvrir

3. <https://haveibeenpwned.com>

4. <https://www.shodan.io>



des éléments compromis appartenant à l'entreprise (soient les données, soient des informations permettant de déduire que l'entreprise a été compromise).

- ▶ L'analyse en temps réel des codes malveillants qui peut permettre en regardant de manière détaillée l'évolution du code pour comprendre et connaître les modalités des attaques et les nouvelles cibles.

2.1.4 Que faire des ces informations

Disposer des fragilités de l'entreprise, et connaître les scénarios potentiels permet de évaluer un niveau de risque.

2.2 de l'outillage sur la menace

2.2.1 la gestion des menaces CTI

Le projet OpenCTI (Open Cyber Threat Intelligence), développé par l'ANSSI en partenariat avec le CERT-EU, est un outil de gestion et de partage de la connaissance en matière d'analyse de la cybermenace (Threat Intelligence). Initialement conçue pour structurer les informations de l'agence relatives à la menace informatique, la plateforme facilite aussi les interactions entre l'ANSSI et ses partenaires.

L'outil, intégralement libre, est disponible à l'usage de l'ensemble des acteurs de la « threat intelligence ». L'application permet ainsi de stocker, organiser, visualiser et partager leurs propres connaissances en la matière.

Le projet OpenCTI ⁵ a été initié en septembre 2018 par l'ANSSI et co-développé avec le CERT-EU en l'absence de solutions complètement appropriées pour structurer, stocker, organiser, visualiser et partager la connaissance de l'ANSSI en matière de cybermenace, à tous les niveaux.

2.2.2 STIX

Les modélisations des attaques est un large champs de recherche et d'outillage. Juste à titre d'illustration, nous pouvons parler d'un modèle comme STIX™ (Structured Threat Information Expression) langage et format de donnée permettant de modéliser et échanger des informations techniques sur les processus d'attaque cyber. Je vous propose d'explorer pour cela sur le site STIX sur GitHub ⁶.

3. DETECTER les attaques

La détection et la surveillance de système d'information d'entreprise, passent souvent pas l'utilisation des informations qui remontent déjà des services de protections périmétriques comme les Firewall (Réseau, Web application...), proxy, routeurs, IPS/IDS. Tous les équipements de sécurité installés sur un environnement sont des sources d'évènements

5. <https://www.ssi.gouv.fr/actualite/opencti-la-solution-libre-pour-traiter-et-partager-la-connaissance-de-la-cybermenace>

6. <https://github.com/OpenCTI-Platform/opencti>



qui permettent pour peu qu'elles soient analysés, de détecter des attaques ou le déploiement de menaces variées (comportements déviants, flux d'APT, flux de scans, ...) Ces sources peuvent être complétées par tout équipement ou système produisant des journaux d'activité (LOGs).

Par ailleurs, certains équipement dédiés de « surveillance » appelés « sondes » sont aussi aptes à remonter des événements de détection sur la base de « signatures » ou « algorithmie » dont des IOC (« indice of compromission »).

Fiche TECHNO : Sondes souveraines

La Loi de Programmation Militaire (LPM) appelle les OIV à s'équiper de sondes dites souveraines, qualifiées par l'ANSSI et disposant d'un niveau de sécurisation élevé. C'est un sujet technique intéressant pour une fiche TECHNO.

Dans cette partie nous allons traiter donc de la détection de menaces sur la base de différentes techniques de détections et de surveillance :

- ▶ Logs et sources d'évènements ;
- ▶ Corrélation d'évènement avec les SIEM ;
- ▶ Évènements provenant des terminaux.

et nous explorerons quelques éléments d'organisation autour de la détection et de la surveillance.

3.1 Log Management

3.1.1 Traces, journaux, logs

Dans le domaine informatique et télécom, le terme log (ou ses synonymes traces, journaux) est généralement un fichier, une base de données ou tout autre moyen de stocker des informations. On parle donc de stockage d'un historique d'événements qu'un logiciel ou un système souhaite « tracer ». Ce mot qui est le diminutif de *logging*, est traduit en français par « journal ». Le log est donc un journal horodaté, qui stocke temporellement les différents événements qui se sont produits sur un logiciel, un ordinateur, un serveur, etc. Il permet ainsi d'analyser avec une fréquence programmée (heure par heure, minute par minute, etc) l'activité d'un processus technique.

Vous trouverez des éléments très intéressants dans le Guide LOG MANAGEMENT ⁷ édité par NetIQ.

Dans notre cas d'usage, les log sont les éléments techniques bruts d'un capteur d'événements. L'objectif est d'assurer que l'ensemble des journaux d'événements contiennent suffisamment d'information pour assurer des corrélations permettant constituer des signatures d'attaques.

7. https://www.microfocus.com/media/white-paper/the_complete_guide_to_log_and_event_management_wp_fr.pdf



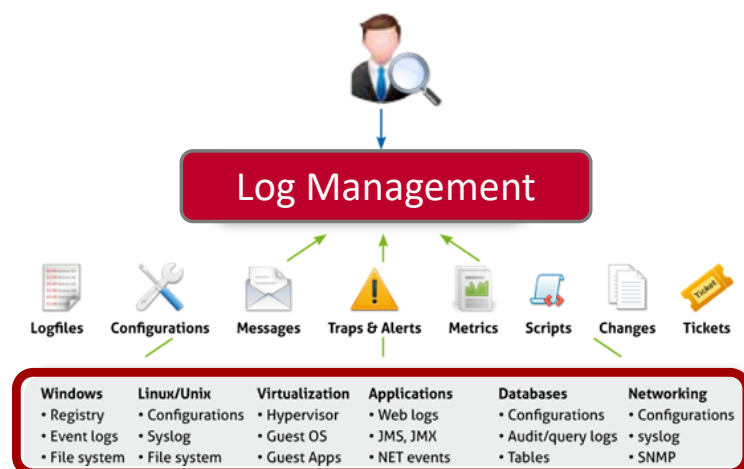


FIGURE 9. Sources de log

Sources choisies des logs

Pour détecter des attaques en temps réel, il faut disposer des informations caractéristiques de ces attaques captées dans le système d'information. Cela nécessite donc de sélectionner les bonnes sources (équipements réseaux, ou informatiques), les bons journaux, les bonnes traces dans ces journaux. Ces choix sont donc primordiaux. Pour faire ces choix, il est nécessaire de connaître la capacité des équipements et des systèmes logiciels de générer ces traces. Au coeur de ces événements il sera alors possible par corrélation de détecter des scénarios complexes d'attaques

Une grande majorité des équipements (réseau, serveurs, terminaux (endpoint)), des bases de données ou des applications d'un système d'information peuvent aujourd'hui générer des logs ou traces. Ces fichiers contiennent, pour chaque équipe, la liste de tous événements « traçables » qui se sont déroulés pendant l'exécution : réussite ou échec d'une connexion, redémarrage, utilisation des ressources (mémoire, ...).

L'exploitation de ces traces est souvent complexe car chaque équipement dispose de ses propres fonctions de gestion des traces, avec encore dans de nombreux cas un format d'enregistrement et de stockage propriétaire. Il faut consulter ces logs équipements par équipements. Heureusement, il existe des outils qui permettent de centraliser et de « normaliser » ces traces.

On peut citer par exemple, SYSLOG, qui par ailleurs n'est pas le seul type d'outil pour assurer la collecte et la normalisation des traces. SYSLOG reste pourtant un outil de référence dans les architectures de collecte.



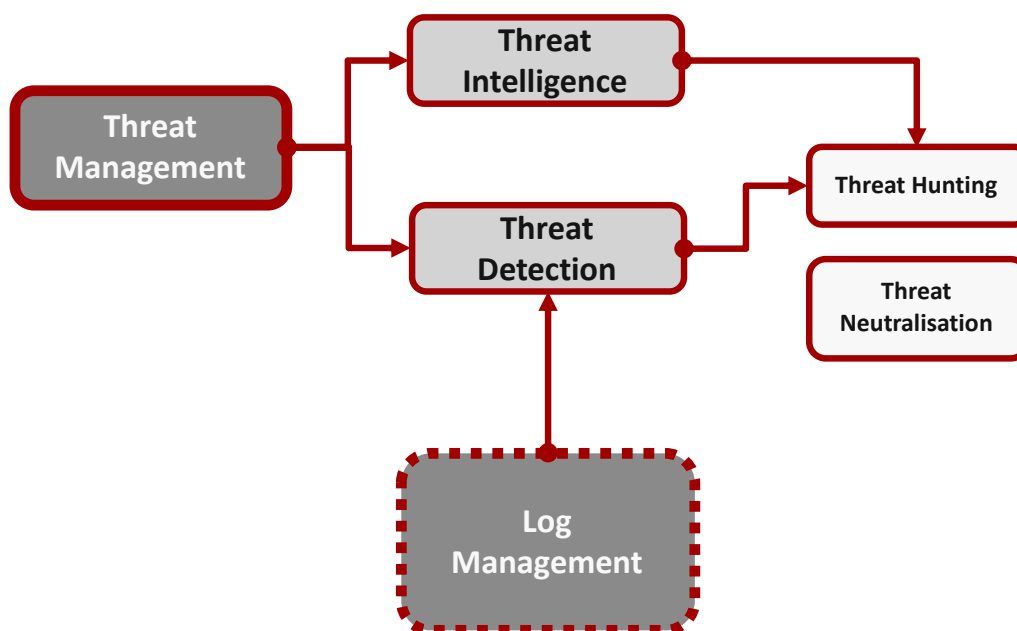


FIGURE 10. Les logs au coeur de la détection

3.1.2 Services et protocole SYSLOG

Le protocole Syslog est un protocole réseau qui permet de transporter les messages de journalisation générés par les applications vers une machine hébergeant un serveur Syslog.

Quand un système veut conserver les traces d'un événement, il est possible, d'utiliser syslog pour communiquer les détails de l'événement à un daemon syslog (serveur syslog) qui va le conserver dans une base de données.

Le protocole Syslog est structuré autour de la notion de périphérique, de relais et de collecteur dans une architecture Syslog.

- ▶ Un **périphérique** est une machine ou une application qui génère des messages Syslog ;
- ▶ Un **relais** est une machine ou une application qui reçoit des messages Syslog et les retransmet à une autre machine ;
- ▶ Un **collecteur** est une machine ou une application qui reçoit des messages Syslog mais qui ne les retransmet pas.

Tout périphérique ou relais sera vu comme un émetteur lorsqu'il envoie un message Syslog et tout relais ou collecteur sera vu comme un récepteur lorsqu'il reçoit un message Syslog.

L'intérêt d'un serveur/collecteur Syslog est donc de permettre une centralisation de ces journaux d'événements, permettant de repérer plus rapidement et efficacement



les défaillances de machines présentes sur un réseau. On trouvera par exemple, sur le site homputersecurity.com ⁸ des éléments pour déployer un serveur SYSLOG, et une description détaillée sur le site [Developpez.com](https://ram-0000.developpez.com/tutoriels/reseau/Syslog/) ⁹.

3.1.3 L'usage des LOGs

Dans un usage de cybersécurité, les traces, journaux et logs informatiques et réseaux structurent les sources d'évènements et de stockages des informations. Ils sont donc :

- ▶ Un outil indispensable au processus de **détection de menace**. La gestion des logs (ou d'événements) s'avère en effet un outil primordial pour les analyses a posteriori, mais peut aussi servir dans la détection en temps réel pour peu que les outils d'analyse puisse le faire ; Nous verront cela dans la partie sur la détection de la menace.
- ▶ **Une couverture légale** . Confrontée à une plainte, une entreprise peut utiliser ces traces pour gérer un litige avec un tiers en attestant de la non-implication de son système d'information ou, a contrario, assumer le litige tout en remontant jusqu'à l'utilisateur concerné. La société peut également utiliser ces traces pour fournir des éléments aux services de polices. La fourniture d'élément probant à valeur légale nécessite quelques précautions.
- ▶ **Le dépistage des malversations internes ou de comportements déviants**. Les flux illégaux, les flux de données déviants (copies de fichiers en masse avant qu'un salarié quitte l'entreprise par exemple)

Evidement, un outil de gestion de LOG ne serait à lui seul et sans fonction de corrélation avoir la capacité de détecter des évènements liés entre eux.

Fiche TECHNO : Syslog et cybersécurité

L'environnement SYSLOG possède une richesse fonctionnelle qui nécessiterait une présentation détaillée pour en appréhender les capacités et la puissance d'usage. C'est un sujet pour une fiche TECHNO de référence. Il existe de nombreuses documentations sur internet, toutefois une présentation détaillée d'une architecture SYSLOG pour un usage de cybersécurité est sujet à explorer.

3.1.4 Puits de logs

La construction d'un « puits de log » est une première brique de réponse : il s'agit de collecter, à l'aide d'un outil automatisé du marché, l'ensemble des journaux d'équipements dans un espace de stockage unique. L'un des critères de sélection de cet outil est justement sa capacité à reconnaître différents formats de logs (syslog, traps SNMP, formats propriétaires...).

Le volume d'information centralisée peut vite exploser : il est important d'éviter la collecte de données inutiles. Par ailleurs, le système peut également être gourmand en

8. <https://homputersecurity.com/2018/03/01/comment-mettre-en-place-un-serveur-syslog/>

9. <https://ram-0000.developpez.com/tutoriels/reseau/Syslog/>



puissance de calcul en fonction des périmètres de recherches effectuées.

On parle de log management à partir du moment où les données contenues dans ces puits sont traitées et exploitées, par exemple pour retrouver un élément dangereux (virus, problème de sécurité...), ou un comportement malveillant (fuite d'information, suppression de données...). Il est nécessaire de cadrer en amont les finalités du projet, qui peuvent être multiples.

3.1.5 Outils d'analyses

Après avoir **collectés, stockés** des événements dans un format compréhensible (structuré ou non), il est nécessaire de disposer d'outils de recherche et d'analyses de ces logs. Il existe de nombreux outils dont beaucoup de codes open source pour ce faire.

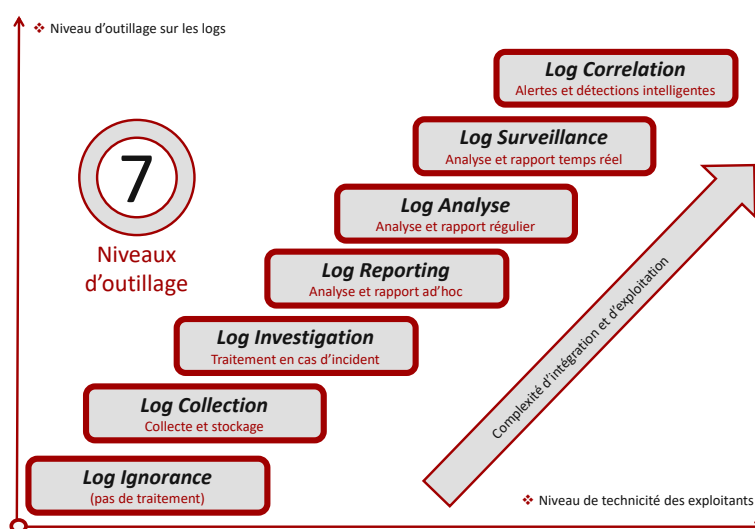


FIGURE 11. Les niveaux d'outillage

Nous pouvons citer par exemple Graylog ¹⁰ basé sur MongoDB pour la gestion des métadonnées et Elasticsearch pour le stockage des logs et la recherche textuelle. Graylog permet de mieux comprendre l'utilisation d'un système d'information tant dans l'amélioration la sécurité (comportements, événements à risques, indicateurs de compromission (IOC)) que dans l'usage des applications et services.

Fiche TECHNO : Analyse de logs

Les outils d'indexation et d'analyse de logs sont nombreux et chacun possède des avantages et des inconvénients, des facilités d'usage de déploiement et des fonctionnalités différentes. C'est un sujet de réalisation pour une fiche TECHNO

10. <https://www.graylog.org>



3.2 SIEM, une technologie

3.2.1 SIEM et Logs

Puisque nous disposons maintenant d'une capacité de capter de l'information pertinente pour détecter des attaques, que nous avons une architecture de collecte, ainsi qu'une architecture de stockage et de filtrage, nous pouvons injecter des informations dans des outils de recherche et de corrélation d'attaque. Ceci pour peu que l'architecture et les outillages puissent suivre la charge d'analyse en temps réel.

Il ne faudra pas aussi oublier que les traces informatiques et réseaux ne sont pas les seules sources d'information nécessaires à la détection d'attaques (en temps réel ou différé), il faut aussi connecter des sources de menaces :

- ▶ **Threat Intelligence Database** : IOC et identifiants des sources malveillantes (IP, noms de domaine, serveur mail ...)
- ▶ **Leak** : Fuite de données détectés par la surveillance du Web et du Darknet.

3.2.2 Un peu d'histoire

Le SIEM (*Security Information and Event Management*) est aujourd'hui l'aboutissement d'un vœu très ancien des responsables sécurité qui supervisent depuis bien des décennies des systèmes de contrôle périmétrique : Corréler tous les événements de fonctionnement sur l'ensemble des équipements.

Le SIEM se définit donc comme la collecte, la surveillance, la corrélation et l'analyse en temps réel d'événements fonctionnels et techniques provenant de sources disparates. Les solutions SIEM d'aujourd'hui permettent à l'entreprise de réagir rapidement et avec suffisamment de traçabilité et d'éléments techniques pour accélérer l'analyse et la réponse à incident. Une solution SIEM assure la gestion, l'intégration, la corrélation et l'analyse généralement en seul lieu, ce qui facilite la surveillance et la résolution des problèmes des infrastructures hétérogènes, complexes et distantes. Sans ce type d'outillage, un analyste de sécurité doit passer en revue des millions de données non comparables, stockées dans des « silos » pour chaque matériel, chaque logiciel et chaque source de sécurité. Un SIEM est donc avant tout un outil de fédération des événements de sécurité.

L'acronyme SIEM ou «gestion des informations de sécurité» fait référence à des technologies combinant à la fois la gestion des **informations** de sécurité et la gestion des **événements** de sécurité. Comme ils sont déjà très similaires, le terme générique plus large peut être utile pour décrire les outils récents qui conjuguent. Là encore, il est essentiel de différencier la surveillance des événements de la surveillance des informations générales. Un autre moyen essentiel de distinguer ces deux méthodes consiste à considérer la gestion des informations de sécurité comme une sorte de processus à long terme ou plus large, dans lequel des ensembles de données plus diversifiés peuvent être analysés de manière plus méthodique. En revanche, la gestion des événements de sécurité examine à nouveau les types d'événements utilisateur pouvant constituer des signaux d'alerte ou indiquer aux administrateurs des informations spécifiques sur l'activité du réseau.



C'est souvent l'usage d'un SIEM dans une ambiguïté de gestion long terme de la sécurité en tant que propriété d'un système d'une part, et la gestion court terme de l'urgence d'une attente à la sécurité qui pose problème dans les projets et dans les opérations.

Ce genre d'outillage est passé par différentes étapes de maturation avec des SIM et SEM et enfin des SIEM. Il s'agit de combiner les fonctions de gestion des informations (SIM, Security Information Management) et des événements (SEM, Security Event Management) en un seul système de management de sécurité.

- ▶ dans la gestion des informations de sécurité (SIM) , la technologie consiste à collecter des informations à partir des journaux d'équipement de sécurité, qui peut consister en différents types de données. Globalement on peut dire qu'un SIM est aimantant important pour des équipes de supervision de la sécurité périmétrique. d'une part pour la traçabilité et le reporting de sécurité.
- ▶ technologies spécialement conçues pour rechercher des authentifications suspectes, des ouvertures de session sur un compte ou des accès de gestion de haut niveau à des heures précises du jour ou de la nuit.

Bien qu'outillant des processus très similaires mais distincts, les trois acronymes SEM, SIM et SIEM ont tendance à être confus ou à causer de la confusion chez ceux, peu familiarisés avec les processus de sécurité. La similitude entre la gestion des événements de sécurité ou SEM et la gestion des informations de sécurité ou SIM est au cœur du problème.

Ces deux types de collecte d'informations concernent la collecte d'informations de journal de sécurité ou d'autres données similaires en vue d'un stockage à long terme, ou l'analyse de l'environnement de sécurité d'un réseau. Quoi qu'il en soit, de nos jours le terme SIEM est utilisé quelque soit d'ailleurs l'usage.

Plus concrètement, un système de type SEM centralise le stockage et l'interprétation des logs en temps réel et permet une analyse. Les experts en cyber sécurité peuvent ainsi prendre des mesures défensives plus rapidement. Un système de type SIM collecte pour sa part des données et les place dans un référentiel à des fins d'analyse de tendances. Dans ce cas, la génération de rapports de conformité est automatisée et centralisée.

Le SIEM, qui regroupe ces 2 systèmes, accélère donc l'identification et l'analyse des événements de sécurité, atténue les conséquences d'attaques et facilite la restauration qui s'ensuit. Pour y parvenir, il collecte les événements, les stocke (avec normalisation) et agrège des données pertinentes mais non structurées issue de plusieurs sources. L'identification des écarts possibles par rapport à la moyenne / norme nourrit la prise de décision. En outre, les tableaux de bord générés contribuent à répondre aux exigences légales de conformité de l'entreprise.

En d'autres termes, avec le SIEM les équipes de sécurité opérationnelle industrialisent la surveillance tout en simplifiant l'analyse de multiples sources d'événements de sécurité (antivirus, proxy, Web Application Firewall. . .). La corrélation des événements provenant



d'applications ou d'équipements très variés est aussi facilitée. De quoi détecter des scénarios de menaces avancées.

Dans la pratique, Il existe 3 grandes manières d'opérer ou de faire opérer un SIEM :

- ▶ SIEM déployé et intégré dans l'entreprise ;
- ▶ SIEM basé dans le cloud ;
- ▶ SIEM géré / managé en mode MSSP

S'équiper d'une solution de type SIEM nécessite encore un investissement conséquent en raison de la complexité de la mise en œuvre des gros outils du marché. Toutefois, bien qu'initialement destiné aux grandes entreprises, le SIEM peut être déployer dans tous les types d'organisations, même les plus petits, avec des solutions de plus en plus automatisées. Toutefois, il ne faudra pas oublier que la configuration d'un SIEM (création de scénario de détection (USE CASE) demande de la ressource et des compétences spécifiques. Par ailleurs, l'important sera par la suite de définir quel type de réaction à la suite de la détection d'un événement à risque.

Un SIEM s'avère capable de détecter des incidents de sécurité qui seraient passés inaperçus. Pour une raison simple : les nombreux hôtes qui enregistrent des événements de sécurité ne disposent pas de fonctions de détection d'incidents.

Le SIEM dispose d'une fonction de détection grâce à sa capacité de corrélation des événements. Contrairement à un système de prévention d'intrusion qui identifie une attaque isolée, le SIEM regarde au-delà. Les règles de corrélations lui permettent d'identifier un événement ayant causé la génération de plusieurs autres (hack via le réseau, puis manipulation sur un équipement précis. . .).

Dans de tels cas de figure, la plupart des solutions ont la capacité d'agir indirectement sur la menace. Le SIEM communique avec d'autres outils de sécurité mis en place dans l'entreprise (Exemple pare-feu) et pousse une modification afin de bloquer l'activité malveillante. Résultat, des attaques qui n'auraient même pas été remarquées dans l'entreprise sont contrecarrées.

- ▶ la première fonction d'un SIEM est déjà de corréler les événements provenant des composants de sécurité ;
- ▶ la deuxième fonction de corréler des événement de comportement du SI ;
- ▶ la troisième fonction de corréler avec des événements externes au SI sur la base de capteurs externes (threats intelligence de type renseignement).

Pour aller encore plus loin, une organisation peut choisir d'intégrer à son SIEM, une Cyber Threat Intelligence (CTI ou Flux de renseignement sur les menaces).

Selon la définition de Gartner, la Cyber Threat Intelligence (CTI) est la connaissance fondée sur des preuves, y compris le contexte, les mécanismes, les indicateurs, les implications et des conseils concrets, concernant une menace nouvelle ou existante ou un



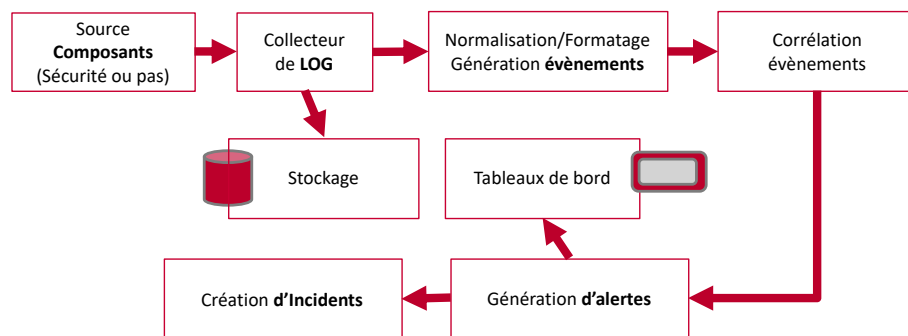


FIGURE 12. architecture d'un SIEM

risque pour les actifs d'une organisation qui peuvent être utilisés afin d'éclairer les décisions concernant la réponse du sujet à cette menace ou un danger.

La CTI consiste donc à collecter et organiser toutes les informations liées aux menaces et cyber-attaques, afin de dresser un portrait des attaquants ou de mettre en exergue des tendances (secteurs d'activités visés, les méthodes d'attaque utilisées, etc.). Résultat, une meilleure anticipation des incidents aux prémices d'une attaque d'envergure.

3.3 la détection d'incidents

3.3.1 Collecter les logs

COLLECTE : Le premier acte est d'alimenter le système avec les «logs» des composants du système d'information (outils de sécurité comme les IDS, IPS, «HoneyPot», Pare-feu, etc. ... que ceux des autres composants comme les équipements réseaux, les serveurs mandataires, de fichier, de mail, les services dans le nuage, etc. ...);

3.3.2 Normaliser et indexer les logs

NORMALISATION : La normalisation des logs est un processus d'uniformisation des données collectées qui, quelque soit la source, détecte le type de donnée pour intégrer chaque information dans une base de données. Par exemple plusieurs équipements différents écrivent chacun dans un de leur journal d'activité l'adresse IP source, la date, l'heure et l'activité. Chaque équipement aura son format particulier, il va donc falloir les analyser pour détecter et labelliser chaque information utile. Sur ce point, les constructeurs d'équipement et les éditeurs de solutions SIEM travaillent ensemble afin de proposer des modules de normalisation qui couvrent la plupart des usages clients. De plus de nombreux formats de solutions de log sont nativement pris en charge par les SIEM. Les informations sont ensuite indexées pour optimiser les temps de traitement de l'information.

3.3.3 Corréler les logs

« CORRELATION » : La valeur ajoutée du SIEM vis-à-vis d'un simple serveur de « log», se joue en grande partie sur la corrélation. En effet, ici il est question de pouvoir mettre en évidence des activités en analysant conjointement plusieurs évènements de sources



différentes. Bien que ces événements puissent paraître anodins de manière isolé, ensemble ils forment un indicateur de compromission. Cette corrélation se base sur des règles forgées à partir de cas d'usage (*USE CASE*)

3.3.4 Analyser et alerter

ALERTE : Ensuite il faut analyser le contenu corrélé pour faire le tri entre les faux positifs et les vrais positifs. Il y a aussi la détection de menace dont le but est de comparer des IOC à des listes de sources malveillantes provenant d'adresse IP, de nom de domaine, de signature de code malveillant, etc. Chaque solution propose ses propres algorithmes d'évaluation du risque lié aux événements analysés. Une fois les analyses réalisées, des mécanismes peuvent déclencher des alertes si un seuil a été franchi.

3.3.5 Réponse graduelle

Il est possible de proposer des actions d'isolation ou de réponse automatisée. Par exemple : Si les «logs» d'un Anti-virus analyse la présence d'un rançongiciel assimilable au logiciel malveillant «Wannacry» et que le poste en question fait des tentatives de connexions extérieures vers le port 445, il est possible de déclencher un isolement de la machine du réseau à l'aide d'un script ou de l'agent présent sur le poste.

La cyber-protection d'une entreprise est principalement basée sur les outils de protection périmétriques que ceux ci soit des équipements physiques ou qu'ils soient dans le cloud : systèmes de détection d'intrusion (IDS), scanners de vulnérabilités, antivirus ainsi que systèmes de gestion et corrélation d'événements sécurité (SIEM). Lorsqu'il s'agit de superviser un système informatique à grande échelle réparti sur plusieurs sites, il devient vite très difficile de corréler et analyser toutes les sources d'information disponibles en temps réel afin de détecter les anomalies et les incidents suffisamment vite pour réagir efficacement. Cette complexité est due à la quantité d'information générée, au manque d'interopérabilité entre les outils ainsi qu'à leurs lacunes en matière de visualisation.

3.4 De l'usage d'un SIEM pour la gouvernance

À l'heure où les normes et certifications de cyber-sécurité sont de plus en plus nombreuses, le SIEM devient un élément clé de tout système d'information. C'est un moyen relativement simple de répondre à plusieurs exigences de sécurité (Exemple : historisation et suivi des logs, rapports de sécurité, alerting, ...) et de prouver sa bonne foi aux autorités de certification ou de suivi. D'autant que le SIEM peut générer des rapports hautement personnalisables selon les exigences des différentes réglementations.

Ce seul bénéfice suffit à convaincre des organisations de déployer un SIEM. La génération d'un rapport unique traitant tous les événements de sécurité pertinents quelle que soit la source des logs (générés en outre dans des formats propriétaires) fait gagner un temps précieux. En outre la création de tableaux de bord issus des usages de supervision ou de surveillance du SIEM offre des outils de pilotage de la sécurité.



3.4.1 Supervision

En collectant, corrélant et analysant l'activité du système d'information, le SIEM peut présenter des indicateurs clés de performance (KPI) et des indicateurs clé de risque (KRI). Les fonctions du SIEM ne sont pas totalement dédiées à la sécurité, on peut donc remonter des informations plus générales concernant l'usage du système d'information. Par exemple, l'outil de source ouverte Elastic Stack, peut être utilisé pour faire du «Business Intelligence» ou du «Big Data». A travers les indicateurs dynamiques et personnalisables, on peut développer un outil de gouvernance du système d'information. Par exemple, connaître l'utilisation de ses boîtes mails, la quantité de pièces jointes échangées comparativement à un service de partage de fichiers.

3.4.2 Surveillance

Le SIEM est capable de faire de la surveillance générale. Le but est de définir des règles de corrélation en capacité de déclencher une alerte ou des règles spécifiques à destination des équipes techniques du SOC. Ces règles peuvent se baser sur l'analyse des menaces, sur des comportements anormaux, ou sur un indicateur de véracité d'une attaque. En termes de surveillance, l'atout du SIEM est de centraliser la gestion des alertes sur une seule interface. L'autre atout du SIEM sur les solutions de surveillance d'infrastructure classique, c'est qu'il ne nécessite pas directement d'agent spécifique. L'arrivée de l'analyse des terminaux pour compléter l'analyse réseau modifie les cultures et offre de nouvelles manière de détecter des codes malveillants en action.

3.4.3 Les contreparties du SIEM

Déployer un SIEM ne suffit pas pour autant à sécuriser complètement votre organisation. Les solutions SIEM présentent des limites qui les rendent inefficaces sans un accompagnement à la hauteur et sans solutions tierces. Contrairement à une solution de sécurité de type IDS ou Firewall, un SIEM ne surveille pas les événements de sécurité mais utilise les données de logs enregistrées par ces derniers. Il est donc essentiel de ne pas négliger la mise en place de ces solutions.

3.4.4 Une configuration pointue

Les SIEM sont des produits complexes qui appellent un accompagnement pour assurer une intégration réussie avec les contrôles de sécurité de l'entreprise et les nombreux hôtes de son infrastructure.

Il est important de ne pas se contenter d'installer un SIEM avec les configurations du constructeur et/ou par défaut, car elles sont souvent insuffisantes. Les configurations doivent être personnalisées et adaptées aux besoins des utilisateurs. De même concernant les rapports, mieux vaut créer ses propres rapports d'analyse, adaptés aux différentes menaces identifiées. À défaut, le risque est réel de ne pas pouvoir profiter des avantages d'une solution de SIEM.



3.4.5 Des investissements à bien anticiper

La collecte, le stockage et l'analyse des événements de sécurité sont des tâches qui semblent relativement simples. Cependant, leur collecte, stockage et l'exécution des rapports de conformité, l'application des correctifs et l'analyse de tous les événements de sécurité se produisant sur le réseau d'une entreprise n'est pas trivial. Taille des supports de stockage, puissance informatique pour le traitement des informations, temps d'intégration des équipements de sécurité, mise en place des alertes. . . L'investissement initial peut se compter en centaines de milliers d'euros auquel il faut ajouter le support annuel.

Intégrer, configurer et analyser les rapports nécessite la compétence d'experts. Pour cette raison, la plupart des SIEM sont gérés directement au sein d'un SOC souvent externalisé. Porteur de grandes promesses, le SIEM mal configuré peut apporter son lot de déceptions. Selon un sondage réalisé auprès de 234 entreprises (Source LeMagIT), 81 % d'utilisateurs reprochent aux SIEM de produire des rapports contenant trop de bruit de fond et pour 63% les rapports générés sont difficiles à comprendre. Faire appel à des prestataires externes disposant de l'expertise dans le domaine reste souvent la meilleure solution.

3.4.6 Un grand volume d'alertes à réguler

Les solutions SIEM s'appuient généralement sur des règles pour analyser toutes les données enregistrées. Cependant, le réseau d'une entreprise génère un nombre très important d'alertes (en moyenne 10000 par jours) qui peuvent être positives ou non. En conséquence, l'identification de potentiels attaques est compliquée par le volume de logs non pertinents.

La solution consiste à définir des règles précises (en général rédigées par un SOC) et le périmètre à surveiller que faut-il surveiller en priorité ? Le périmétrique ? L'interne ? Réseau/système/application ? Quelle technologie à prioriser ? etc.

3.4.7 Une surveillance à exercer 24h/24

Pour fonctionner correctement, les solutions SIEM nécessitent une surveillance 24h/24 et 7j/7 des journaux et des alertes. Un personnel formé ou une équipe dédiée sont requis pour consulter les journaux, effectuer des examens réguliers et extraire les rapports pertinents. Il s'agit à la fois de disposer des expertises requises, de gagner en lisibilité budgétaire et, aussi, de profiter d'engagements de services. Des conditions à réunir afin que l'investissement dans une solution SIEM marque une étape clé dans la protection de votre organisation contre les menaces avancées.

3.4.8 Analyse d'impact

Un autre problème majeur dans l'usage d'un SIEM est que l'action de comprendre l'impact réel d'une vulnérabilité ou d'une alerte IDS est généralement dévolue à un analyste cybersécurité humain, qui doit lui-même faire le lien entre toutes les informations techniques et sa connaissance de tous les services ou processus liés aux incidents de



sécurité détectés sur les composants concernées (serveurs, PC, smartphone, IoT,...) .

Le projet DRA est une étude complémentaire de CIAP qui vise à fournir une analyse de risque en temps réel, afin de déterminer automatiquement l'impact réel dû à la situation sécurité globale du système et du réseau. Pour cela une nouvelle méthodologie innovante a été développée en combinant un générateur automatique d'arbres d'attaque (attack trees/graphs) et un moteur d'analyse de risque « traditionnel » similaire à EBIOS.

Les systèmes de gestion des informations et événements de sécurité (SIEM) font régulièrement l'objet de critiques acerbes. Complexité, besoins importants en ressources de conseil externes. . . de nombreuses entreprises ont été déçues par leur expérience du SIEM pour l'implémentation de la supervision de la sécurité.

Mais la technologie n'est plus, désormais, la raison pour laquelle des entreprises peinent à réussir leurs implémentations de SIEM. Les principales plateformes de SIEM ont reçu de véritables transplantations cérébrales, se transformant en entrepôts de données taillés sur mesure pour fournir les performances et l'élasticité requises. Les connecteurs système et les agrégateurs de logs, autrefois complexes et peu fiables, sont aujourd'hui efficaces, rendant la collecte de données relativement simple.

Mais il y a une limite au SIEM, comme à toute technologie s'appuyant sur des règles : le SIEM doit savoir ce qu'il doit chercher. Aucun boîtier SIEM ne pourra identifier automatiquement, comme par magie, une attaque tirant profit d'une méthode ou d'une vulnérabilité inédite.

Le SIEM joue un rôle important dans la détection d'attaques. Mais pour qu'il puisse détecter les attaques connues et inconnues, l'entreprise qui le déploie doit construire des ensembles de règles qui lui permettront d'identifier des conditions d'attaques et des indicateurs spécifiques à son environnement. Et le tout de manière cohérente. Comment donc construire ces règles ?

3.4.9 Tout collecter

Sans disposer de suffisamment de données collectées, le SIEM n'a pas grand chose à analyser. Mais la première étape est de collecter les bonnes données. Et celles-ci sont notamment les logs des équipements réseau, de sécurité et des serveurs. Ces données sont nombreuses et faciles à obtenir. Ensuite, il faut s'intéresser aux logs de l'infrastructure applicative (bases de données, applications). Les experts du SIEM ajoutent à cela les données remontées par de nombreuses autres sources, comme celles des systèmes de gestion des identités et des accès, les flux réseau, les résultats des scans de vulnérabilités et les données de configurations.

Avec les SIEM, plus il y a de données collectées, mieux c'est. Si possible, autant tout collecter. S'il est nécessaire de définir des priorités, alors mieux vaut se concentrer sur les actifs technologiques critiques, à commencer par les équipements installés dans les environnements sensibles et ceux manipulant des données soumises à régulation, ou encore ceux touchant à la propriété intellectuelle. Le souci principal de ces outils est souvent le modèle économique et de licence. En effet, le prix est majoritairement basé



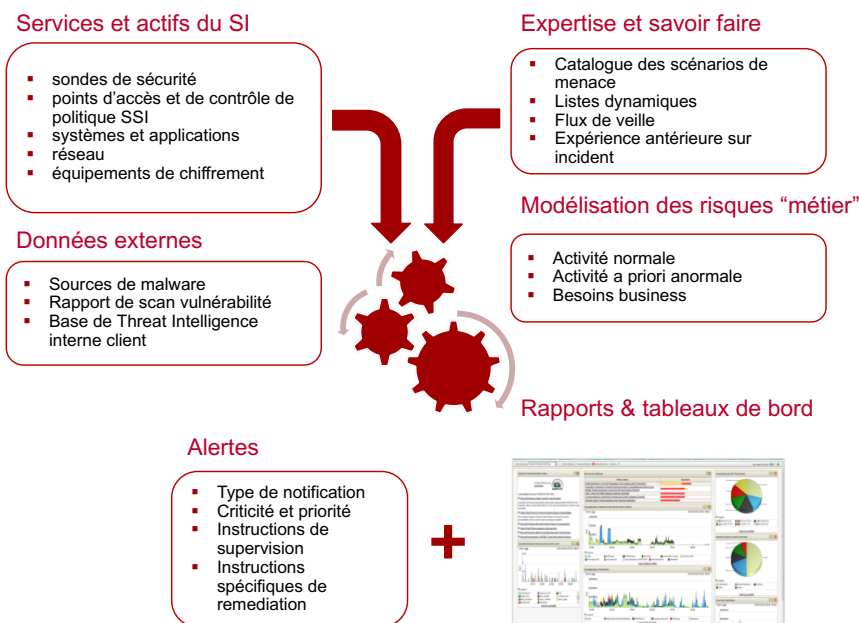


FIGURE 13. Cadre méthodologique

sur le nombre d'évènement.

3.5 Construire les règles (UseCase)

Construire une règle pour SIEM est un processus itératif. Cela signifie qu'il est relativement lent et qu'il doit être affiné, précisé au fil du temps. De nombreuses personnes sont atteintes de la « paralysie de l'analyste » en début de processus, parce qu'il existe des millions de règles pouvant être définies. Ainsi, il est conseillé de se concentrer sur les menaces les plus pressantes pour déterminer les règles à définir en premier.

Dans le cadre du processus de modélisation, il convient de commencer par un actif important. Pour cela, il faut adopter le point de vue de l'attaquant et chercher ce que l'on pourrait vouloir voler.

Modéliser la menace. Il faut se mettre à la place de l'attaquant et imaginer comment entrer et voler les données. C'est la modélisation de l'attaque, avec énumération de chaque vecteur avec le SIEM. Et il convient de ne pas oublier l'exfiltration car sa modélisation offre une opportunité supplémentaire de détecter l'attaque avant que les données ne se soient envolées. Dans ce processus, il s'agit d'adopter des attentes réalistes car le modèle d'attaque ne peut pas par essence être parfait ni complet. Mais il convient toutefois d'engager le processus de modélisation. Et il n'y a pas de mauvais point de départ.

Affiner les règles. Il convient ensuite de lancer l'attaque contre le SI, telle que modélisée. Les outils pour cela ne manquent pas. C'est l'occasion de suivre ce que fait le SIEM. Déclenche-t-il les bonnes alertes ? Au bon moment ? L'alerte fournit-elle suffisamment



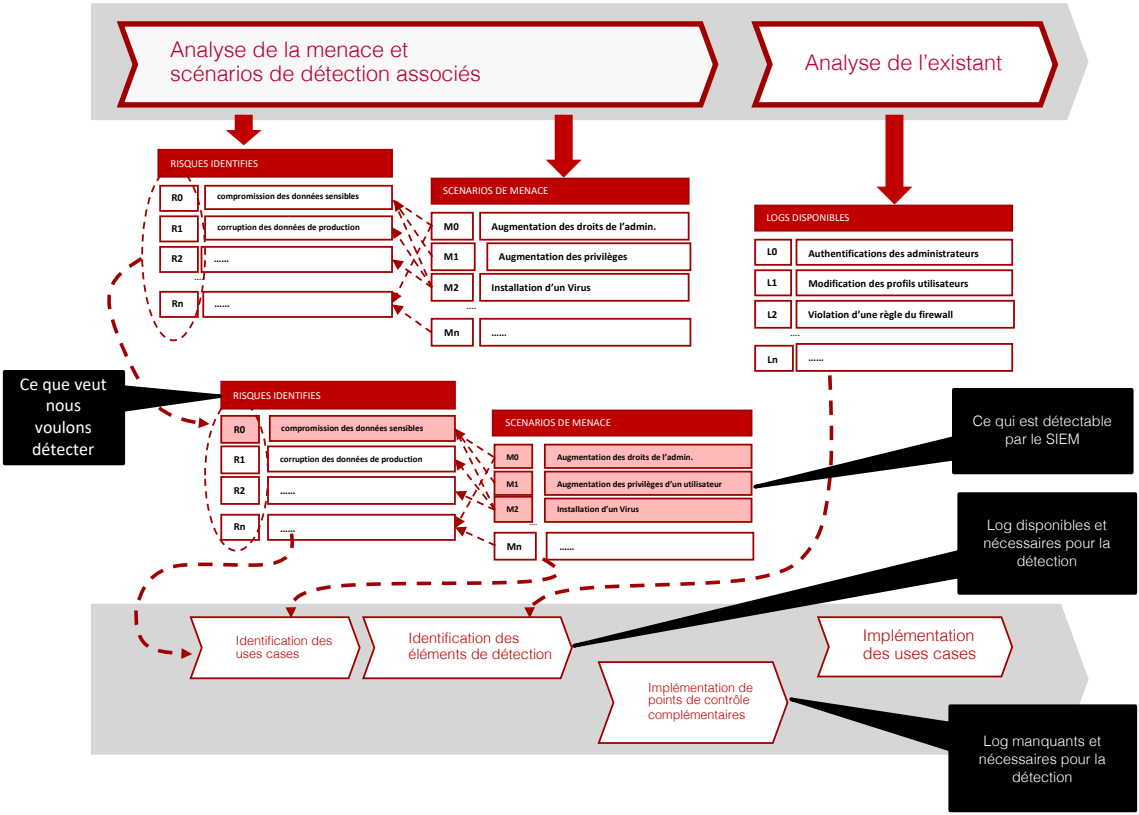


FIGURE 14. Construction des UseCase



d'informations pour assister les personnes chargées de la réaction ? Si l'alerte n'est pas adéquate, il convient de revoir le modèle et d'ajuster les règles.

Optimiser les seuils. Avec le temps, il deviendra de plus en plus clair que certaines alertes surviennent trop souvent, et d'autres pas assez. Dès lors, il convient d'ajuster finement les seuils de déclenchement. C'est toujours une question d'équilibre... un équilibre délicat.

Laver, rincer, recommencer. Une fois l'ensemble initial de règles pour ce modèle d'attaque spécifique implémenté et optimisé, il convient de passer au vecteur d'attaque suivant, et ainsi de suite, en répétant le processus en modélisant chaque menace.

Ce processus ne s'arrête jamais. Il y a constamment de nouvelles attaques à modéliser et de nouveaux indicateurs à surveiller. Il est toujours important de suivre les informations de sécurité pour savoir quelles attaques sont en vogue. Les rapports tels que celui de Mandiant sur le groupe APT1 intègrent désormais des indicateurs clairs que chaque organisation peut surveiller avec son SIEM. Armé de ces renseignements sur les menaces et d'un environnement de collecte de données complet, il n'y a plus d'excuse : il est temps de commencer à chercher les attaques avancées qui continuent d'émerger.

Mais avec le temps, il sera nécessaire d'ajouter de nouveaux types de données au SIEM, ce qui impliquera de revoir toutes les règles. Par exemple, le trafic réseau, s'il est capturé et transmis au SIEM, fournira quantité de nouvelles informations à étudier. Mais comment ce regard sur le trafic réseau sera-t-il susceptible d'affecter la manière dont certaines attaques sont traitées ? Quelles autres règles faudrait-il ajouter pour détecter l'attaque plus vite ? Ce ne sont pas des questions triviales : il convient de revoir les règles du SIEM chaque fois qu'est ajoutée une nouvelle source de données (ou retirer, le cas échéant) ; cela peut faire la différence sur la rapidité avec laquelle une attaque est détectée... si elle l'est.

Le plus important aspect de ce processus est la cohérence. Le SIEM n'est pas une technologie du type « installe et oublie ». Il requiert du temps, de l'attention, et d'être alimenté, tout au long de sa vie opérationnelle.

3.6 Quelques défis des SIEM

La problématique globale des SIEM est de corréler de l'événement, la question de fond est la collecte de ses événements. La collecte de LOG est la principale sources d'événements, toutefois, toutes les sources d'événements sont susceptibles d'enrichir la corrélation, en particulier les vulnérabilités, les IOC, les infos de end-point ... La collecte des informations d'opérations et de renseignements nécessite de la FUSION de capteurs. Cette fusion chère aux militaires est un premier pas qui prend compte aussi de l'information économique, politiques ou sociale de l'entreprise, car ces événements peuvent « matcher » avec des attaques complexes.

3.6.1 L'intelligence artificiel

L'Intelligence artificielle est devenue en quelques années un objectif marketing chez les éditeurs de cybersécurité. Avec un certains succès puisque de nombreuses entreprises



disent utiliser des solutions de sécurité basée sur de l'IA. Certaines solutions s'appuient davantage sur des moteurs de règles sophistiquées que sur de réelles fonctionnalités d'IA. Pour parler d'intelligence artificielle, il faut en effet que la technologie inclut :

- ▶ une capacité de perception de l'environnement au moyen d'un apprentissage supervisé ou non,
- ▶ une capacité d'analyse et de résolution de problème,
- ▶ une capacité de proposition d'action, voire de décision autonome.

Sur le plan théorique, les apports de l'IA en matière de cybersécurité sont donc nombreux, qu'il s'agisse de prévention, d'anticipation, de détection ou de réaction. Dans la pratique, la détection de vulnérabilités ou de menaces internes ou externes apparaît aujourd'hui l'un des usages les plus matures. Les systèmes actuels basés sur des signatures montrent encore leurs limites : nombre élevé de faux positifs, incapacité à s'adapter aux dernières menaces, notamment aux APT, lourdeur des bases de signature, ce qui a un impact sur les performances.

Fiche TECHNO : SIEM et IA

L'évolution des SIEM sont orientée par le traitement de masse d'évènement. L'IA du BigData offre des possibilités nouvelles. L'IA dans la détection d'attaque est un bon sujet pour une fiche TECHNO .

3.6.2 Quelques SIEMs

On peut citer ainsi quelques SIEM non pas pour en faire un publicité particuliers mais simplement pour donner quelques indications sur la provenance ...

Le Gartner positionne régulièrement des produits et services dans son magic Quadrant. en 2019, Splunk, IBM QRadar et LogRhythm NextGen SIEM sont toujours bien positionnés. Dell Technologies (RSA NetWitness), Exabeam (Security Management Platform), McAfee (Enterprise Security Manager) et Securonix complètent le carré des Leaders. Toutefois des entreprises comme Microsoft challenge ces acteurs.

3.7 Threat Hunting

3.7.1 la chasse aux menaces

La chasse aux menaces dormantes ou aux compromissions (Threat Hunting) fait partie intégrante de la gestion de la menace. C'est souvent les équipes « Hunters » assure le maintien du contact entre la défense et les attaquants lors d'une attaque en cours. Souvent issus d'équipe de compétences liés à réponse à incident.

La chasse aux menaces est une tactique permettant de connaître avec plus d'acuité l'environnement de la menace et donc le degré de risque de cyber-attaques auquel est soumise une entreprise.



La terminologie threat hunting regroupe plusieurs types d'action et la définition de n'est pas totalement stabilisée. Globalement on y trouve deux grandes classes de « *threat hunting* » :

- ▶ Celles travaillant autour de l'environnement, de la surface d'attaque et qui orientent ses actions sur des méthodes de « recherche » permettant de débusquer des menaces latentes ou des menaces dormantes, de les réveiller, de les suivre, de les comprendre pour établir le contact avec l'attaquant.
- ▶ Et une autre plus active ou proactive dont l'objectif est de rester, conserver le contact avec l'attaquant lors d'une réaction à une alerte.

3.7.2 Etablir le contact

Quand on parle d'établir contact, nous parlons d'aller au contact au sens martial du terme. c'est dire en direct de suivre, caractériser la sources de la menace et jouer avec elle.

La méthode de « hunter » consiste en premier à dresser un portrait global de la surface d'attaque, tout en identifiant les attaquants potentiels, leurs motifs et leurs façons de faire. Plus précisément, le « threat hunting » consiste en une analyse détaillée de différents éléments :

- ▶ la position de l'entreprise, notoriété, popularité sur internet, en analysant en particulier, les médias traditionnels et les médias sociaux ;
- ▶ l'environnement économique de l'entreprise dont ses fournisseurs, ses clients, ses partenaires, ses employés ;
- ▶ le corpus technologique et physique de l'entreprise, dont les architectures techniques et les mécanismes informatiques avec l'environnement économique ainsi l'environnement sécuritaire de ses relations.

Sur la base de cette analyse globale, des SPOF (*Single Point Of Failure*) peuvent être trouvés.

Grace à la visualisation globale des liens il est possible de comprendre où, comment, pourquoi et potentiellement par qui (activistes, anciens employés, fournisseurs, etc.) la prochaine attaque pourrait être perpétrée. Les « threat hunters », ne sont pas simplement en attente de répondre aux alertes du système de défense, ils cherchent activement des menaces dans leurs propres réseaux afin de prévenir ou de minimiser les dommages. Cette méthode s'avère l'une des plus proactives.

3.8 la détection de menace sur des terminaux

La mise en place de mesures de remontées de LOGs pour les terminaux fait partie généralement des politiques de cyberdéfense, et les journaux d'évènement au sein des systèmes d'exploitation ou des applications sont utilisables pour les SIEM. La sécurité du



poste de travail est un axe à part entière des politiques de sécurité. Dans cette sécurité, on y trouve évidemment la sécurité périmétrique avec les *Firewalls* personnels, les anti-virus. La détection de menace au plus prêt du terminal et donc au plus prêt de l'utilisateur, permet d'être, bien des fois, plus pertinent pour contenir des attaques ou des déviances. Cette détection peut être réalisée par un EDR (*Endpoint Detection and Response*) qui est caractérisée par ses capacités de détection, d'investigation et aussi de remédiation.

Etre au plus prêt de l'utilisateur, donc du terminal, est toujours un axe majeur à pour un responsable sécurité.

Chaque jour des milliers de terminaux (*Endpoints*) sont compromis par des attaques ciblées. Parmi ces attaques, les menaces avancées et persistantes (APT : Advanced Persistent Threats) compromettent des machines et avec les risques associés.

Des techniques d'attaques permettent de by-passer les solutions locales d'antivirus et de firewall personnel ou d'utiliser plus simplement la crédulité des utilisateurs pour infecter des machines ou tout simplement utiliser les faiblesses de configuration du systèmes pour ex-filtrer des données.

Dans ce contexte l'EDR est un outil de sécurité complémentaire aux outils de sécurité "personnels" avec lequel il travaille afin de bloquer les menaces (connues ou nouvelles (zero-days)). C'est un outil qui se place au niveau du système d'exploitation en complément du niveau réseau ... Un EDR fait de l'analyse comportementale, et permet de monitorer les actions de l'utilisateur ou des applications au niveau terminal et de réagir.

3.8.1 Détection

L'EDR est capable de surveiller l'exploitation de failles de sécurités en surveillant les appels noyaux et les différents services habituellement ciblés notamment chez Windows. Cette capacité de surveillance et la corrélation d'évènements lui permettent de reconnaître des méthodes et habitudes qu'ont les hackers et dont il est plus difficile de se prémunir.

L'analyse comportementale (UBA - User Behavior Analytic) est un autre point supporté par les EDR et qui permet de reconnaître des comportements déviant d'un cadre normatif, ou technique souvent après une phase d'apprentissage. Grâce à ces analyses, l'EDR peut émettre des alertes vérifiable qui renforceront l'apprentissage.

L'intérêt de cette technique est qu'elle permet de stopper un attaquant dans son élan : si un pdf PDF contient un script qui ouvre powershell et ouvre une connexion sur un port classique d'un serveur extérieur au SI alors cette suite d'action sera considérée comme anormale et va être bloquée par l'EDR. Cette visibilité est une grande force de l'EDR car elle permet une remédiation à la source de l'infection.

3.8.2 Investigation

Comme mentionné plus haut, l'EDR permet d'observer des suites d'actions dont le résultat est des plus douteux. Cette visibilité sur les processus est d'une très grande aide pour faire de l'investigation : les actions sont corrélées et remontées dans une plateforme



centralisée qui permet d'étendre l'apprentissage observé d'un poste à tous les autres.

Grâce à ces plateformes, un *Security Operation Center* (SOC) est capable de savoir immédiatement combien de postes sont touchés et d'en remonter la piste : c'est un outil d'investigation qui peut s'interfacer à un *Security Information and Event Management* (SIEM) en offrant de la visibilité sur les terminaux. D'autant plus que votre SOC pourra se servir de l'EDR pour récupérer des artefacts de l'attaques à distance.

3.8.3 Remédiation

En ce qui concerne la remédiation, l'EDR a des capacités similaires à celles d'un antivirus de nouvelle génération et peut notamment bloquer, supprimer et mettre en quarantaine des fichiers. Les équipes de sécurité pourront aussi s'appuyer sur l'outil pour faire du nettoyage de clé de registre par exemple, ou patcher la mémoire en direct pour contrôler un malware. De plus, certains EDR permettent aux analystes du SOC d'Orange Cyberdefense de prendre la main à distance sur un terminal qui nécessiterait une investigation plus poussée encore.

L'EDR n'est généralement pas une solution stand-alone. C'est un complément qui s'intègre à d'autres outils de sécurité locale, de SIEM ou de sécurité réseau. Il permet d'étendre la visibilité sécurité du SI jusqu'au terminaux et permet d'améliorer la pertinence et la précision de détection des scénarios.



3.8.4 Quelques EDR de références

Microsoft



Les fonctionnalités de détection de point de terminaison et de réponse de Microsoft Defender - PACM permettent de détecter des attaques avancées qui sont en temps réel et exploitables. Les analystes de la sécurité peuvent hiérarchiser efficacement les alertes, obtenir une visibilité sur l'ensemble des violations et prendre des mesures pour remédier aux menaces. En cas de détection d'une menace, des alertes sont créées dans le système pour qu'un analyste examine. Les alertes associées aux mêmes techniques d'attaque ou affectées au même agresseur sont agrégées dans une entité appelée incident. L'agrégation des alertes de cette manière permet aux analystes de rechercher et de répondre à des menaces collectivement. Microsoft Defender ATP collecte en continu le comportement du terminal via de la télémétrie cyber. Cela inclut les informations de processus, les activités réseau, les composants optiques intégrés au noyau et au gestionnaire de mémoire, les activités de connexion utilisateur, les modifications de la base de Registre et du système de fichiers, etc. Les fonctionnalités de réponse vous permettent d'apporter une correction rapide aux menaces en agissant sur les entités affectées.

🔧 Classe : **EDR**, Site de référence : Microsoft ¹¹ 🌐 Éditeur : **Microsoft** 🌐 <https://www.microsoft.com/fr-fr/microsoft-365/windows/microsoft-defender-atp>

3.9 Caractérisation de la menace

Nous avons évoqué les mécanismes de surveillance et de détection constitués en particulier d'outils comme le SIEM, ou l'EDR. Il est toutefois important de disposer de compétences et de processus structurés pour caractériser la menace. Cette caractérisation permet de lever les doutes, de caractériser les éléments d'une attaque dont les impacts mais surtout les raisons et les intentions des attaquant. Ces éléments sont évidemment primordiaux pour définir une réponse adaptée.

3.10 De l'usage d'un CSIRT

Nous verrons l'usage d'un CSIRT dans le chapitre sur la réponse à incident toutefois nous pouvons donner ici les fonctions ou compétences clefs pour effectuer cette caractérisation :

- ▶ Rétro-conception : afin d'analyser un code malveillant ;
- ▶ Liaisons police/gendarmerie : afin de lancer au plus tôt les investigations via les services de l'état ;

¹¹. <https://www.microsoft.com/fr-fr/microsoft-365/windows/microsoft-defender-atp>



- ▶ Base « enrichie » d'IP & Url malveillantes : afin de caractériser non seulement la provenance, mais aussi l'intentionnalité de l'attaque ;
- ▶ Etudes de cyber-intelligence sur le secteur d'activité de l'entreprise : afin de corrélérer une agression avec des risques dans un environnement politique ou économique particulier.

4. (SOC) Security Operation Center

Le SOC (Security Operation Center) est au coeur du système de « Veille Alerte et réponse ». C'est une tour de contrôle sécurité de l'espace Cyber.

Il est constitué généralement d'une équipe d'analystes, et d'outils permettant de surveiller l'environnement. Cette surveillance s'effectue sous la forme de l'exploitation de différents outillages (SIEM, EDR....).

Il intègre l'ensemble des fonctions liées à la menace :

- ▶ Veille sur la menace
- ▶ Détection d'évènements à risques et gestion de ceux ci
- ▶ Détection d'attaques ou de comportement critiques
- ▶ Réaction aux incidents et remédiation

Malheureusement, dans encore beaucoup de cas, les équipes SOC et les équipes liées à la gestion des vulnérabilités sont cloisonnées, ce qui ne couvre pas de manière intégrée l'ensemble des fonctions de cyberdéfense d'entreprise.

On peut aussi Intégrer dans le SOC des fonctions de *Threat Hunting* .

Les grands principes de réussite d'un SOC sont :

- ▶ Une veille Cyber efficace et à large ouverture en terme de menaces ;
- ▶ Une capacité à identifier d'une cartographie détaillée des ressources de l'infrastructure et de correctement identifier les menaces avec des analyses des risques ;
- ▶ Réaliser une collecte des évènements de sécurité, pour nourrir une corrélation temps réel ;
- ▶ Contextualisation et amélioration continue afin de limiter le nombre de faux-positifs ;
- ▶ Faciliter la communication entre les niveaux opérationnels
 - Niveau 1 : **réception** des alertes en temps réel ;
 - Niveau 2 : **corrélation** et analyse multi-alertes pour déclenchement de l'incident ;
 - Niveau 3 : **investigations** poussées, forensics et découverte des indicateurs de compromission.

4.0.1 Le SOC de demain

On peut par ailleurs s'interroger sur le fait qu'un tel système peut et doit opérer d'autres missions que les missions de sécurité pures. Si la supervision des réseaux a été longtemps au outils au services des techniciens, la supervision de l'environnement digital c'est à dire l'environnement informationnel de l'entreprise est un axe fondamental. Le SOC peut devenir **Cybersecurity Operational Center** (CSOC) opérant le suivi des risques digitaux au sens large, incluant les réseaux sociaux et leur



cohorte de fausses informations et d'information pouvant être des indicateurs de crise à venir pour l'entreprise.

Aujourd'hui les entreprises organisent donc leur environnement de gestion de la menace dans la sécurité informatique avec ce centre opérationnel de sécurité CSOC. Ses principales missions se structurent donc autour de :

- ▶ la supervision de la sécurité ;
- ▶ du management du risque ;
- ▶ de l'analyse des menaces ;
- ▶ de l'audit ;
- ▶ de l'investigation numérique ;
- ▶ de la prévention.

Chacune de ces missions nécessite des spécialistes, des outils et une gouvernance. Cependant au centre de ces activités le CSOC s'appuie entre autre sur des données de référence

- ▶ Les journaux d'activité ;
- ▶ des scénarios menaces ; mis à jour avec des outils de surveillance d'"émergence de la menace, en plus de scénarios spécifiques à l'entreprise ;
- ▶ des bases de vulnérabilités et des accès à des services permettant ;
- ▶ des outils de surveillance des fuites (Data Leak Detection) dans le Darkweb par exemple ;
- ▶ des outils de remédiation, de notarisation (enregistrement des données à valeur probante).

Le SIEM et l'EDR répondent à une partie des besoins des CSOC et doit pour être efficaces s'intégrer à des services connexes. Le SIEM existe depuis les années 90, sous la forme de SIM et de SEM. Le rôle du SIM est de centraliser tous les logs dans une seule base de données pour permettre des analyses et de l'archivage. Celui du SEM est de surveiller en temps réel les événements sur le système d'information, les corrélater et alerter si les conditions sont remplies. Ces deux outils de sécurité ont fini par fusionner dans un seul produit, permettant de lier les événements aux informations, mais il y a encore des convergences à venir.

4.0.2 Evaluation d'un SOC

L'efficacité d'un SOC peut être évaluée. A l'image d'équipe de Pentest qui testent la résistance d'un système, des équipes de tests de SOC peuvent être déployées pour auditer le niveau d'efficacité d'un SOC. Les équipes qui testent des OSC sont nommées des **Purples Team**.

un CSOC est efficace s'il arrive à détecter avec pertinence les attaques en cours, Cependant, il est important, de comprendre qu'un CSOC ne détectera jamais des attaques dont le scénario n'a pas été pensé / « programmé ». On trouvera dans une publication du CLUSIF [↗](https://clusif.fr/publications/reussir-deploiement-dun-soc/)¹², les critères pour réussir le déploiement d'un SOC.

4.0.3 Les outils connexes d'un SOC

Au delà des SIEM, il semble important d'ajouter à l'outillage d'un SOC un ensemble de systèmes permettant de mesurer et d'évaluer l'impact des attaques. Un travail intéressant autour de la notion

12. <https://clusif.fr/publications/reussir-deploiement-dun-soc/>



d'Echelle de RICHTER (Voir un article du FIC 2014 [🔗](#)¹³) d'une attaque afin de définir des indicateurs « de cotation ».

- ▶ **l'origine** de l'attaque qui mesure la puissance potentielle de la source de menace : du hacker de base à la menace étatique ;
- ▶ Le type de **cible** qui mesure la précision de la diffusion de la menace : de la cible au hasard à la menace ciblée ;
- ▶ Le **vecteur** d'attaque qui mesure le niveau de sophistication de la menace : du malware « sur étagère » à l'APT élaborée ;
- ▶ Le **préjudice** qui mesure l'impact subi par la cible : d'une perte faible à une mise en péril de la résilience même de l'organisme ;
- ▶ La **visibilité** de la menace qui mesure de nombreux éléments comme la motivation ou durée de l'attaque : d'un DDOS immédiatement constaté à une attaque invisible ;
- ▶ La **persistance** qui mesure la fréquence de l'attaque sur sa cible : d'une fréquence forte de type robotisée (Bots) à une fréquence unitaire visant un but précis, ou la furtivité.

4.1 Les outillages d'un SOC

Au delà des SIEM, des sondes, des EDR, l'orchestration est au coeur de l'efficacité des fonctions d'un CSOC.

L'orchestration et l'automatisation de la sécurité permet de réduire les délais de réponse, de limiter l'exposition aux attaques et offrir une cohérence des processus cyberdéfense. Ces outils d'automatisation et d'orchestration, appelés SOAR, sont conçus pour améliorer la productivité et l'efficacité de centres des opérations de sécurité et des analystes.

Ces outils automatisent les tâches de routines chronophages, ils aident à coordonner les cycles de vie de réponse aux incidents et de gestion des incidents. Outils de cohérence, ils permettent d'assurer reproductibilité de la discipline des opérations de cybersécurité et permet de réduire le temps nécessaire pour détecter et traiter les incidents.

On y trouve par exemple dans ces outils de « *Security Orchestration, Automation, and Response* » (SOAR) :

- ▶ l'introduction de sources de menaces de manière automatique au base SIEM (abonnement de threat-intelligence) ;
- ▶ la production de règles sur la base de déviations relevées ;
- ▶ le pilotage automatique des composants de sécurité (modification de règles, passage en mode dégradé ...) ;
- ▶ l'exécution de tâche de conservation de traces légales (notarisation) ;
- ▶ la gestion automatisée de « patches » critiques (intégration au DEVSECOPS) ...

4.1.1 DEVSECOPS

Dans de nombreuses entreprises les processus liés à la sécurité sont encore isolés et confiés à une équipe spécifique sans s'intégrer totalement avec les chaines de développement ou dans les équipes opérationnelles et encore moins dans les chaines intégrées DEVOPS. Si une approche

13. <https://observatoire-fic.com/prendre-la-mesure-des-cyberattaques-peut-on-definir-une-echelle-de-richter-dans-le-cyb>



DevOps efficace garantit des cycles de développement rapides et fréquents, les équipes sécurité doivent de plus en plus s'intégrer dans les processus DEVOPS tant pour y apporter un volet *security by design*, mais surtout intégrer des mécanismes de sécurité opérationnelle. Ce sont ces dynamiques de sécurité tant de conception, que de sécurité opérationnelle que nous appelons « DEVSECOPS ».

Fiche TECHNO : DEVSECOPS

Le domaine du DEVSECOPS est un sujet à part entière de la gestion de la sécurité dans les chaînes de prise en compte de la sécurité en DEVOPS, et l'utilisation des techniques DEVOPS dans les chaînes de gestion du dynamique du changement dans la détection de menace et la réponse sur incident. Un sujet intéressant pour une fiche TECHNO .

4.2 l'efficacité du CSOC

Pour mesurer l'efficacité d'un CSOC, il existe plusieurs moyens de mesures :

- ▶ La **couverture fonctionnelle et technique** du CSOC pour estimer l'efficacité de l'articulation entre les stratégies de cyberdéfense, de cyber-protection et les stratégies de surveillance détection,
- ▶ La **performance de la détection** pour évaluer l'efficacité des règles de corrélation en place, basée sur les indicateurs de services (Nombre de détections, nombre d'évènement ...);
- ▶ La **maturité du service CSOC**, mesurée sur le niveau d'organisation des services (ITIL par exemple), les coûts, les compétences, les services connexes ...

4.3 Leak : surveiller les fuites

J'ai ajouté un chapitre spécial sur les fuites de données pour deux grandes raisons :

- ▶ La détection des fuites de données peuvent simplement se révéler par l'apparition de tout ou partie de ces données dans le Darkweb.
- ▶ Les fuites de données étant souvent des fuites de données de type « données personnelles », elles impliquent le déroulement de processus de déclaration au titre de la GDPR.

Je ne rentrerai pas ici dans la présentation du RGPD avec son cortège d'exigence et d'organisation à mettre en place (Liste de traitement, déclaration, nomination de responsable, etc). Je ne vous propose que de regarder rapidement, la partie détection et partie réponse à incident.

Le terme « fuite de données », ou « data breach » en anglais, est utilisé pour toute situation impliquant la perte, la modification injustifiée ou la publication par accident, par malveillance, de données considérées ou marquées comme confidentielles.

Il est important dans la mise en place de scénario dans les SIEM, et dans le traitement de SOC que l'évènement de fuites de données personnelles puissent être traité avec un mécanisme précis et documenté, car ces évènements sont très contraint par la réglementation. A titre de remarques, les évènements touchant la fuite de données liées à la protection du secret de défense (Secret Défense) puisse aussi être traité dans un processus particulier car les ces fuites peuvent aussi faire l'objet de procédure au pénal.

Le GDPR prévoit que le responsable du traitement des données à caractère personnel signale au plus tôt les fuites de données pouvant constituer une atteinte à la vie privée des personnes



concernées. Cette information à la CNIL et aux personnes concernées en cas d'impact important sur ces personnes.

La méthodologie est assez simple pour peu que le constat de l'incident puisse être fait le plus vite possible. Cela peut se faire sur la base d'évènement provenant des équipements de sécurité (via un SIEM par exemple) ou par l'utilisation de services de veille, ou simplement par l'avertissement d'un tiers qui découvre cette fuite.

- ▶ Détection,
- ▶ Enrayer la fuite, limiter l'impact,
- ▶ Analyser les sources de menaces,
- ▶ réagir de manière juridique.

Deuxièmement, vous devez entreprendre dès que possible les démarches pour enrayer l'incident ou en limiter l'impact. Tous les collaborateurs doivent respecter plusieurs règles. S'ils trouvent des informations à un endroit inapproprié, ils doivent les supprimer ou en informer un responsable. Il peut s'agir de supports physiques, mais aussi de fichiers sur le réseau. Ils doivent également donner l'alerte s'ils rencontrent des étrangers non accompagnés dans une zone sécurisée. Et ainsi de suite. Si des alarmes indiquent un piratage ou une infection des systèmes, les gestionnaires de ces systèmes devront les examiner au plus vite et peut-être les désactiver de manière préventive.

En cas de doute, il est préférable d'arrêter un traitement ou d'empêcher le transport des données traitées jusqu'à ce que vous sachiez clairement s'il y a effectivement un problème, et dans quelle mesure les données traitées sont encore correctes. Cela permet souvent d'éviter qu'un incident ne se transforme en fuite de données. Tant que des données traitées à mauvais escient ne sont pas diffusées ou rendues publiques, il n'y a pas d'infraction, et donc pas d'impact. Au sens strict, il n'est pas encore question d'une fuite de données.

Ensuite, et éventuellement en parallèle, vous pouvez lancer une analyse des faits. D'une part, il faut établir la cause du problème. Vous pourrez ensuite réfléchir aux améliorations dans l'organisation, les systèmes ou les applications, et dans le mode de travail de vos collaborateurs, pour éviter que l'incident ne se reproduise. D'autre part, il faut examiner l'impact réel ou éventuel de l'incident. Y a-t-il des risques pour la confidentialité et l'intégrité des données ? S'agit-il (en partie) de données à caractère personnel ? Quelles peuvent-être les conséquences de cette infraction ? Dans de nombreux cas, il vous faudra du temps pour savoir quelle quantité de données a été impactée et combien de personnes sont concernées. Souvent, vous ne saurez pas non plus d'emblée s'il y a véritablement un risque d'impact, ni quelle peut être l'ampleur des dommages.

Ce n'est que lorsque vous aurez une réponse à toutes ces questions qu'il vous sera possible de faire le bon choix quant à la nécessité de signaler la fuite de données à la Commission de protection de la Vie Privée ou aux personnes concernées. Le quand et le comment de ce signalement seront abordés dans le prochain article.

5. Technologies et Organismes connexes

5.1 Stratégie de cyberdéfense et de surveillance

Ce chapitre en construction donne quelques éléments complémentaires sur les stratégies et techniques de cyberdéfense d'entreprise.



5.1.1 Technique de déception (Deceptive Defense)

D'une protection statique en profondeur encore trop souvent périmétrique, vers une détection résiliante pour désormais envisager des logiques de contre-attaques dynamiques, les stratégies de cybersécurité évoluent. L'échelle de temps entre l'occurrence d'une attaque, sa détection et son élimination est un marquant significatif du risque : de plusieurs mois à quelques jours, de jours à quelques heures, l'enjeu est de nos jours d'agir en temps réel contre l'attaquant. Le leurrage numérique (deceptive security) est au cœur de nouvelle stratégie de cybersécurité. Il s'agit de retourner la dynamiques de l'attaquant en cherchant à le tromper pour tenter de le démasquer, ou de le le dissuader de continuer de part le risque d'être découvert qu'il prend. Le leurrage numérique peut relever d'une forme de dissuasion cyber.

5.1.2 Honeypots

Les pièges « honeypots » sont un leurre pour les attaquants, en imitant une ressource de calcul réel (par exemple, un service, une application, un système ou des données). Toute entité entrée en connexion à un « honeypot » est alors considérée comme suspecte, et son activité est surveillée pour détecter une malveillance. L'arsenal du leurrage défensif est historiquement basé sur les pots de miel (honeypots). Ceux-ci reposent sur l'analyse statique d'écart de composants par rapport à un comportement connu et sain. Ces technologies se heurte à deux problèmes : le passage à l'échelle pour couvrir la diversité et la complexité des systèmes numériques, et la génération excessive de faux positifs. Les honeypots évoluent pour devenir des pièges actifs qui sont disséminés dans l'environnement réel pour mieux cerner les stratégies de l'attaquant. Les architectures de déploiement des leurres se spécialisent selon le domaine d'application (systèmes d'information, systèmes industriels, finance, médical...) ou en fonction des composants ciblés par les attaques (serveurs, pare-feux, antivirus...) ou encore par rapport à la charge offensive (malwares...). Les leurres tendent à générer de vrais positifs en temps réel. Leur efficacité repose sur deux propriétés, l'une inhérente aux composants de sécurité, la non-compromission, et l'autre caractéristique de l'attaque : la furtivité. Cette nouvelle génération de leurres numériques enrichit les stratégies d'investigations au sein des centres opérationnels de sécurité (SOC). Ainsi, des logiques de raisonnements déductifs (déterministes) ou inductifs (hypothétiques) se confrontent pour caractériser finement le mode opératoire des attaquants en le resserrant si possible jusqu'à l'attribution de l'attaque. Le caractère actif de ces nouvelles technologies de leurre soulève toutefois de nombreuses interrogations sur le plan éthique et réglementaires (respect de la vie privée en particulier).

Le leurrage numérique devient une composante essentielle d'une lutte informatique défensive et contribue de plus en plus au processus des scénarios de réponses, ripostes et d'escalade.



Fiche TECHNO : « Deceptive Defense » en cyberdéfense

Les techniques de déception en cyberdéfense sont en pleine évolution. C'est un sujet parfait pour une fiche TECHNO avec les différents thèmes :

- ▶ le leurrage numérique : honeypots, leurres, pièges
- ▶ les architectures de déploiement de leurres selon les domaines d'application
- ▶ la spécialisation de leurres pour les services, pour la sécurité, contre les malwares...
- ▶ les propriétés du leurrage numérique : non-compromission, furtivité...
- ▶ l'apport à l'investigation numérique : raisonnements déductifs/inductifs, caractérisation des attaques, attribution...
- ▶ la contribution à la lutte informatique défensive : scénarios de ripostes et d'escalade.
- ▶ le positionnement du leurrage dans les modèles d'attaques (MITRE ATT@CK,...), par rapport à la caractérisation des attaques (CAPEC...) et plus généralement son apport à la connaissance du risque cyber (cyber threat intelligence – CTI)
- ▶ le leurrage et la réglementation (NIS, RGPD,...).

5.2 Quelques techniques d'attaques

A titre d'illustration je vous propose d'illustrer les typologies d'attaque en analysant l'architecture classique d'un système dit de BOTNET, mécanisme de commande et de contrôle de codes malveillants pilotés/télécommandés.

6. Botnet, des codes malveillants organisés

6.1 Les attaques

Liste des attaques utilisables par les botnets :

- ▶ Dénî de service distribué (DDoS ¹⁴)
- ▶ DDoS contre paiement
- ▶ Récupération des identifiants
- ▶ Exposition médiatique ou démonstration de force
- ▶ Dissimulation d'une autre attaque
- ▶ Création d'un avantage concurrentiel
- ▶ Censure par attaque de serveurs
- ▶ Vengeance par cryptolocker
- ▶ Infrastructure d'anonymisation
- ▶ Recherche de vulnérabilités
- ▶ Infrastructure d'anonymisation des communications
- ▶ Contournement de mesures de limitation ou blocage

¹⁴. Distributed Denial of Service



- ▶ Envoi de pourriels
- ▶ Diffusion de codes malveillants
- ▶ Exécution de codes malveillants sur les machines-zombies
- ▶ Hébergement de codes malveillants
- ▶ Fraude aux clics
- ▶ Compromission d'accès
- ▶ Brute force hors-ligne
- ▶ Brute force direct
- ▶ Cryptominage

6.2 Cycle de vie d'une attaque :

Pour la compréhension, il est nécessaire de comprendre les différentes étapes depuis l'infection jusqu'au fonctionnement complet du botnet.

6.2.1 Infection de la machine

Cette première étape a généralement pour but de télécharger la charge virale sur un serveur. Elle peut être initiée via les vecteurs suivants :

- ▶ Par spam(existence de spambot)
- ▶ Exploitation de faille lié à la navigation sur un site web(malvertising ¹⁵, waterholing ¹⁶)
- ▶ P2P
- ▶ Spear phishing ¹⁷
- ▶ SMS, MMS
- ▶ Bluetooth
- ▶ TDS ¹⁸
- ▶ Exploit kits ¹⁹

6.2.2 Activation

Après téléchargement, l'installation du malware peut établir un premier contact avec le botnet (serveur dédié, servant-bot) ayant une fonctionnalité de C&C. Le téléchargement de rootkit d'installation ou de DLL complémentaire finalise la mise en place du botnet sur la machine infectée.

6.2.3 Mise à jour

Les échanges permettent l'ajout de fonctionnalité, de configurations afin que le botnet puisse identifier et s'adapter à son environnement. Il peut, par exemple, vouloir modifier son hash ²⁰ afin de conserver une certaine furtivité pour la continuité de l'attaque.

15. exploitation de pop-up publicitaires

16. ciblage de sites web fréquentés

17. hameçonnage ciblé pour récupérer données ou identifiants

18. Traffic Distribution Service, outil et service de redirection de trafic

19. plate-forme d'exploitation supporté par un site web permettant de tester une liste d'exploits

20. signature numérique, ici on parle de signature virale



6.2.4 Auto-protection

La persistance et la dissimulation sont les facteurs clés de cette étape. L'installation de rootkit de protection, la modification du système, etc permettent de masquer l'action du botnet.

6.2.5 Propagation

Cette phase d'extension est à la fois locale par du scan et distante par diffusion virale (mail avec lien ou pièce jointe).

6.2.6 Phase opérationnelle

Cette dernière phase vise à accomplir les actions souhaitées de l'attaquant. Déclenchées, synchronisées ou persistantes ces attaques s'adaptent aux cibles désignées. Ordonné par le C&C elles peuvent être activées ou mises en sommeil afin de ne pas attirer l'attention.

6.3 Définition

Le terme **botnet**, contraction de l'anglais **robot+net**, se définit par l'ensemble des programmes, machines, serveurs connectés à internet ayant un ou plusieurs processus commun de communication. Placé sous le contrôle d'un opérateur humain, appelé botmaster, le botnet recrute des machines en exploitant les vulnérabilités, failles, infections afin d'étendre son réseau à travers l'utilisation de canaux de Command and Control(C&C).

Avec l'IoT²¹, et ses appareils connectés, le réseau s'étend de plus en plus au sein de notre société. L'actuelle faiblesse en terme de sécurité lié aux objets connectés représente une menace majeure et croissante dans notre environnement.

6.4 Historique

Le concept, inventé en 1988 à l'université de Oulu en Finlande, fut développé à l'origine pour gérer les services associés au protocole IRC²².

Le premier bot "<GM"> assistait ainsi l'utilisateur dans la gestion des connections IRC. Cette gestion automatisée, permettant via un accès à distance, de contrôler et de réaliser des opérations a très vite montré un haut pouvoir malveillant.

En Mai 1999, Pretty Park, un malware de forme trojan horse se propageant sur le net permettait de voler les mots de passe.

Les premières dérives furent notamment l'affrontement de botnet IRC (Eggdrop en décembre 1993, puis GTbot en avril 1998).

6.5 Motivations liées à la menace botnet

- ▶ L'aspect lucratif représente l'intérêt majeur pour l'utilisateur de botnet. L'automatisation d'une tâche contrôlée à distance permettant de rapporter facilement des revenus (revente d'information, fraude au clic, spam), surtout si celle-ci est réalisée de manière anonyme(réseau TOR²³).
- ▶ La motivation idéologique, comme par exemple, lors du conflit entre la Georgie et la Russie en 2008 ou de nombreux sites étatiques faisait l'objet de cyber-attaques massives paralysaient les infrastructures.

21. Internet of Things

22. Internet Relay Chat, un protocole de communication textuel

23. The Onion Routing, un réseau d'anonymisation



- ▶ La motivation personnelle, à travers la vengeance ou le chantage, est également une finalité grâce notamment au caractère anonyme de l'attaque.

6.6 Les type de menaces

Les botnets représentent les outils de diffusion des attaques. Cet outil permet aux cybercriminels de disposer d'un grand nombre de services développés dans un environnement collaboratif. Vendus sur le web, ils instrumentalisent l'attaque quelque soit le but recherché.

Liste des menaces possibles :

- ▶ Relayer du spam pour du commerce illégal ou pour de la manipulation d'information (par exemple des cours de bourse)
- ▶ Réaliser des opérations d'hameçonnage
- ▶ Identifier et infecter d'autres machines par diffusion de virus et de programmes malveillants (malwares)
- ▶ Participer à des attaques groupées de déni de service (DDoS)
- ▶ Générer de façon abusive des clics sur un lien publicitaire au sein d'une page web (fraude au clic)
- ▶ Capturer de l'information sur les machines compromises (vol puis revente d'information) ;
- ▶ Exploiter la puissance de calcul des machines ou effectuer des opérations de calcul distribué notamment pour cassage de mots de passe
- ▶ Voler des sessions utilisateurs par credential stuffing ;
- ▶ Mener des opérations de commerce illicite en gérant l'accès à des sites de ventes de produits interdits ou de contrefaçons via des techniques de fast flux, simple ou double-flux ou RockPhish
- ▶ Miner des cryptomonnaies, telles que le bitcoin¹.



6.7 Architecture aléatoire

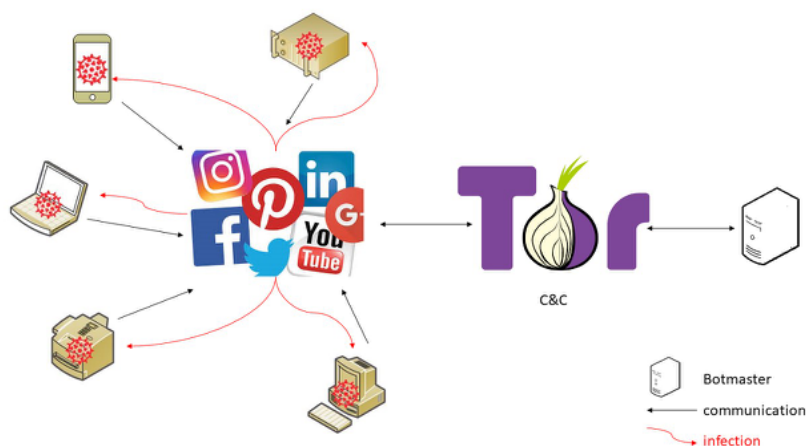


FIGURE 15. Architecture aléatoire

6.7.1 Définition

Ce concept représente une variante de l'architecture centralisée et peut se retrouver dans une variété de malware connue sous le nom de RATs²⁴. Ces chevaux de Troie fonctionnent sur un principe de client/serveur parfois mis en place avec des techniques de social engineering (exemple : un fichier d'installation récupéré sur un site douteux). Ils exécutent la partie client à l'insu de l'utilisateur pour se connecter au serveur.

Cette architecture tire profit de l'exploitation de plate-formes existantes supportant le protocole HTTP comme Facebook, Twitter, Yahoo, Evernote, Google, etc. et de réseau permettant de camoufler les échanges comme TOR, Hornet, etc.

6.7.2 Liste des protocoles utilisés par le botnet

- ▶ HTTP
- ▶ protocole propriétaire (exemple XMPP²⁵)
- ▶ IRC

6.7.3 Avantages

- ▶ Architecture déjà existante
- ▶ Réseau important en terme d'utilisateurs
- ▶ Utilisation des canaux existants (exemple : messagerie instantanée)
- ▶ utilisation des fonctionnalités du réseau (exemple : Le réseau anonymisation TOR)
- ▶ Difficulté de démanteler son propre réseau

24. Remote Access Trojan

25. Extensible Messaging and Presence Protocol de MSN Messenger



6.7.4 Inconvénients

- ▶ Vulnérabilité du botnet face aux mécanismes de défense
- ▶ Connexion en permanence
- ▶ bloqué par le filtrage de liens de la plate-forme

6.7.5 exemples

6.8 Architecture centralisée

6.8.1 Définition

Un ou plusieurs nœud de communication permettent aux bots d'échanger des données via un canal de communication. Les nœuds représentent un serveur ou serveur relais avec comme fonction le C&C.

6.8.2 Liste des protocoles utilisés par le botnet

- ▶ IRC²⁶
- ▶ HTTP²⁷
- ▶ IRC modifié

6.8.3 Avantages

- ▶ Architecture centralisée
- ▶ Simplicité de mise en oeuvre (mIRC, ...)
- ▶ Utilisation des canaux IRC (topics, messages) pour l'envoi des commandes vers les botsPerformance (non gourmand en bande passante)
- ▶ Connexions régulières entre les bots et le C&C (non-permanente)
- ▶ Recherche des ordres dans des forums, avec des mots clés ou même dans des images (stéganographie)

6.8.4 Inconvénients

- ▶ Vulnérabilité du botnet (serveur central)
- ▶ Connexion en permanence
- ▶ Facile à détecter (filtrage du flux IRC)

6.9 Architecture décentralisée

6.9.1 Définition

En utilisant des réseaux peer-to-peer, on s'affranchit d'un point central de communication. Chaque bot, selon ses caractéristiques, apporte les ressources pour élaborer le système de C&C. Il existe plusieurs typologies de réseau overlay²⁸ :

26. Internet Relay Chat

27. HyperText Transfer Protocol

28. réseau logique de recouvrement



- ▶ **overlay P2P non-structuré** : les topologies sont aléatoires (loi de puissance, aléatoire uniforme,...)
- ▶ **overlay P2P par super-pairs** : tous les pairs du réseau ne sont pas égaux, certains d'entre eux étant automatiquement sélectionnés pour servir temporairement le rôle de serveur pour les recherches ou le contrôle du réseau (comme FastTrack ou Gnutella)
- ▶ **overlay P2P structuré** : une cartographie établissant le lien entre le contenu et son emplacement ; ce type de réseau implémente en général – mais pas systématiquement une table de hachage distribuée (DHT) ; on retrouve dans cette catégorie les protocoles P2P Chord, Tapestry et Kademlia (utilisé par le logiciel eMule).

6.9.2 Liste des protocoles utilisés par le botnet

- ▶ TCP/IP
- ▶ UDP

6.9.3 Avantages

- ▶ Architecture décentralisée
- ▶ Indépendant de l'architecture DNS
- ▶ difficile à repérer
- ▶ Connexions régulières entre les bots et le C&C (non permanente)
- ▶ Le botmaster donne les informations comme un bot faisant partie du réseau
- ▶ L'information transite de voisin en voisin
- ▶ très difficile à neutraliser

6.9.4 Inconvénients

- ▶ Pas de vision globale du réseau par un bot
- ▶ Connexion en permanence
- ▶ Facile à détecter (filtrage du flux IRC)

6.10 Architecture hybride

6.10.1 Définition

Une architecture hybride intègre une solution de repli, comme à l'aide d'un DGA²⁹ ou de plusieurs niveaux successifs entre le P2P et le C&C. Cette hiérarchie permet de masquer une partie des adresses IP utilisées afin de rendre plus complexe l'analyse du botnet. L'association de bot clients et de bots servent³⁰ met en évidence l'organisation du réseau par le botmaster. Ces niveaux intermédiaires permettent de solidifier l'architecture du réseau.

29. Domain Generation Algorithm

30. serv-**eur** et cli-**ent**



6.10.2 Liste des protocoles utilisés par le botnet

Reprenant les protocoles présentés dans les architectures précédentes, L'efficacité, la furtivité et la complexité des méthodes de communication ont pour objectif de nuire aux efforts de démantèlement. Ainsi, les coopérations internationales entre acteurs institutionnels et privés sont généralement nécessaires pour permettre de démanteler les botnets les plus sophistiqués.

6.10.3 Avantages

- ▶ Nombre important de domaines
- ▶ masquage de l'adresse IP

6.10.4 Inconvénients

- ▶ tributaire de la bande passante

6.10.5 exemples

6.11 L'analyse

6.11.1 L'analyse statique

Réalisée en sandbox, sur une machine virtuelle ou une machine dédiée avec des outils pré-installés comme InetSim, FakeNet ou Mozzle, elle débute par une analyse statique pour identifier les éléments et la composition du malware. L'examen du code et des fonctions appelées permettent d'évaluer les capacités du botnet.

6.11.2 L'analyse dynamique

Cette étape est relativement utile pour la compréhension de la menace car elle présente la machine infectée sous plusieurs états à l'aide de snapshots³¹. Ces captures instantanées situent l'avancement de l'infection lors de l'attaque. Il est souvent nécessaire en présence d'algorithme chiffré d'utiliser cette méthode pour désobfusquer le code et comprendre la structure du malware.

6.12 La défense et le blocage

Au même titre que la protection contre les malwares, les recommandations en terme de SSIREcommandations ANSSI, applicables localement, doivent s'inscrire dans nos habitudes et augmentent ainsi la probabilité de bloquer l'étape initiale de l'infection (spam, navigation non-sécurisée, etc.).

Les mises à jour logicielles et système sont essentielles pour bloquer l'exploitation de CVE³² Common Vulnerabilities and Exposures.

Au niveau du FAI, les notifications en cas de connexions malveillantes et la surveillance des adresses IP sont un frein à l'extension du botnet.

Enfin la détection d'un appel de fonctions anormal par l'antivirus, l'autorisation et l'identification des flux sortants par le firewall permettent le blocage de l'activité malveillante. Les fonctionnalités

31. copie des données/modifications apportées à un système

32. Common Vulnerabilities and Exposures



recherchées de l'antivirus dans ce cadre sont un firewall bidirectionnel, une protection contre le phishing, la vérification de la certification, la lutte contre le tracking, la vérification du téléchargement, le blocage des pop-ups et pages WEB malveillantes, etc.

6.13 Le démantèlement

Les méthodes de détection impliquent parfois des actions offensives visant à entraver le développement du botnet. Il est cependant nécessaire d'avoir un support juridique et judiciaire pour mener à bien des actions adaptées au type et à la taille du botnet. Celles-ci sont généralement menées en coopération avec les industriels (Microsoft, Level 3, Cisco, etc.) et la communauté scientifique.

6.14 La détection

Recherche d'anomalie, comparaison de signatures, pots-de-miel, toutes ces techniques basées sur l'activité du réseau reposent sur l'inspection des flux et des paquets.

6.14.1 La détection passive

L'identification et l'analyse passive des flux (adresses IP, port source et destination, étiquette MPLS³³, etc.) permettent de classer les protocoles suspects et les serveurs de C&C.

BotFinder, par exemple, permet de décomposer un flux (durée moyenne des connexions, nombre d'octets transférés, etc.) et de le comparer à l'activité normale du réseau.

L'observation des DNS³⁴ permet aussi d'identifier les domaines malveillants afin de caractériser le botnet suspecté. BotGad, un système d'exploitation permettant d'analyser le trafic DNS sur un réseau local, utilise un algorithme basé sur l'apprentissage afin de définir la stratégie de groupe du botnet.

EXPOSURE, un autre système d'exploitation déployé au sein de l'ISP³⁵, permet d'analyser à large échelle mais sur une durée de plusieurs mois le trafic DNS. Produisant une liste de domaines malveillants, il permet, par exemple, d'identifier un volume conséquent de requêtes pour un même domaine.

Enfin le recours aux pots-de-miel et l'analyse des journaux d'activité sont les éléments de base d'une recherche d'activité liée aux botnets. Suivant cette idée, le SIEM³⁶, une solution de gestion de la sécurité, représente un outil précieux et novateur afin d'optimiser la veille du trafic et d'automatiser les processus de sécurité en cas de comportement anormal.

6.14.2 La détection active

Différentes techniques existent comme le sinkholing³⁷ redirigeant le trafic vers des serveurs afin de simuler le comportement du C&C et de diminuer la puissance du réseau du botnet.

L'infiltration, fonction de l'architecture du botnet, consiste à simuler le comportement d'un botnet contrôlé à l'aide de drones IRC ou de script afin de capturer le trafic et de remonter jusqu'au botmaster. Le projet Pebbletrace reprend cette idée d'identification du botmaster en piégeant les équipements infectés avec une charge défensive afin de retourner le trafic contre le botnet.

Les botnets font maintenant partie d'une économie souterraine sous forme de services payants. Le harcèlement, le vol, les attaques par déni de service, etc figurent comme des produits de vente

33. MultiProtocol Label Switching

34. Domain Name Server

35. Internet Service Provider

36. Security Information and Event Management

37. également appelé serveur gouffre, gouffre Internet ou Blackhole DNS



accessibles, moyennant finances, pour n'importe quel criminel.

D'après l'ENISA³⁸ les prix varient suivant la fiabilité, la durée et le type de service requis. Par exemple une heure de DDoS est disponible pour 38\$. Les différentes architectures permettent de mieux comprendre l'organisation du botnet et l'importance du C&C.

Les notions de veille technologique et de partage d'informations sur la menace sont essentielles du fait de l'implication du botnet dans les réseaux publics et privés.

Dans le cadre du renseignement lié aux menaces, les constructeurs de smartphone fournissent les informations (recherches, cibles des attaques, menaces associés aux mobiles, vulnérabilités des objets IoT) issues de leurs Threat Intelligence Center³⁹.

Selon Nokia (Nokia's Threat Intelligence Report, l'activité des botnet sur l'IoT représente 78% des événements de détection de malware en 2018.

La menace omniprésente de ces objets connectés (santé, domotique, médias, électroménager,...) n'est aujourd'hui pas assez prise en considération par notre société de consommation. Le manque de sécurité accrue de ces appareils fait apparaître ces objets connectés comme des acteurs potentiels constituant le réseau d'un botnet.

L'arrivée prochaine de la 5G⁴⁰ sera, dans ce domaine, un vecteur majeur pour la diffusion de l'infection du malware et le nombre d'attaques associés au botnet (exemple attaques DDoS).

38. European Union Agency for Cybersecurity, anciennement European Network and Information Security Agency

39. Centre de renseignements liés aux menaces cyber

40. prévisions en moyenne 100Mbit/S en download et 50 Mbit/s en upload



7. Contributions

7.1 Comment contribuer

Les notes et les présentations sont réalisées sous \LaTeX .

Vous pouvez contribuer au projet des notes de cours CNAM SEC101 (CYBERDEF101). Les contributions peuvent se faire sous deux formes :

- ▶ Corriger, amender, améliorer les notes publiées. Chaque semestre et année des modifications et évolutions sont apportées pour tenir compte des corrections de fond et de formes.
- ▶ Ajouter, compléter, modifier des parties de notes sur la base de votre lecture du cours et de votre expertise dans chacun des domaines évoqués.

Les fichiers sources sont publiés sur GITHUB dans l'espace : (edufaction/CYBERDEF) ⁴¹. Le fichier Tex/Contribute/Contribs.tex contient la liste des personnes ayant contribué à ces notes. Le guide de contribution est disponible sur le GITHUB. Vous pouvez consulter le document **SEC101-CO-Contrib.doc.pdf** pour les détails de contributions.

7.2 Les contributeurs/auteurs du cours

Les auteurs des contributions sont :

7.2.1 Années 2020

- ▶ **David BATANY** (Contributeur LATEX) : BOTNET
- ▶ **Charly Hernandez** : User and Entity Behavior analytics, UEBA
- ▶ **Florian PINCEMIN (Orange)** : SIEM en quelques mots

7.2.2 Années 2019

- ▶ **François REGIS** (Orange) : CyberHunting

7.2.3 Années 2018

- ▶ **Julia HEINZ** (Tyvazoo.com) : ISO dans la gouvernance de la cybersécurité

41. <https://github.com/edufaction/CYBERDEF>



Table des matières