Faut-il croire tout ce qu'on nous dit ?

Stéphan DAVID
SSI – Sécurité des Systèmes d'information
PRH – Pôle réseaux et hébergement
DSI direction des systèmes d'information
3 avenue Vercingétorix
63033 CLERMONT-FERRAND

Résumé

Lorsque l'on parle de Cyberattaque ou de cyberdéfense, ce préfixe "Cyber" concerne l'ensemble de l'informatique connectée. La cybercriminalité est une atteinte au fonctionnement normal de nos systèmes d'information. Ces attaques, de plus en plus sophistiquées, ciblent tous les secteurs d'activités sans distinction (service bancaire, industriel, santé, service public) de taille (du grand groupe internationale, Ministère, la petite PME ou encore à des particuliers) et à des fins malicieuses.

Mots-clefs

Cybersécurité, scanner, vulnérabilité, CVE, CVSS, SIEM, SOC, IA, Machine learning...

1 Etat des lieux

Il existe différents types d'attaques informatiques parmi lesquelles on peut dégager trois grandes familles : les ransomwares, les campagnes d'attaques par hameçonnage (fishing) et les vulnérabilités.

1.1 Les ransomwares

Les ransomwares rendent inaccessibles des données, le plus souvent en les chiffrant, et demandent une rançon au propriétaire en échange de la clé de déchiffrement. On observe toutefois une évolution des pirates dans l'utilisation de ce type de menace pour aller vers le crypto-jacking (Figure 1) logiciels malveillants qui utilisent la puissance informatique des ordinateurs infectés pour miner des crypto-monnaies, et ainsi gérer des revenus de cette activité.

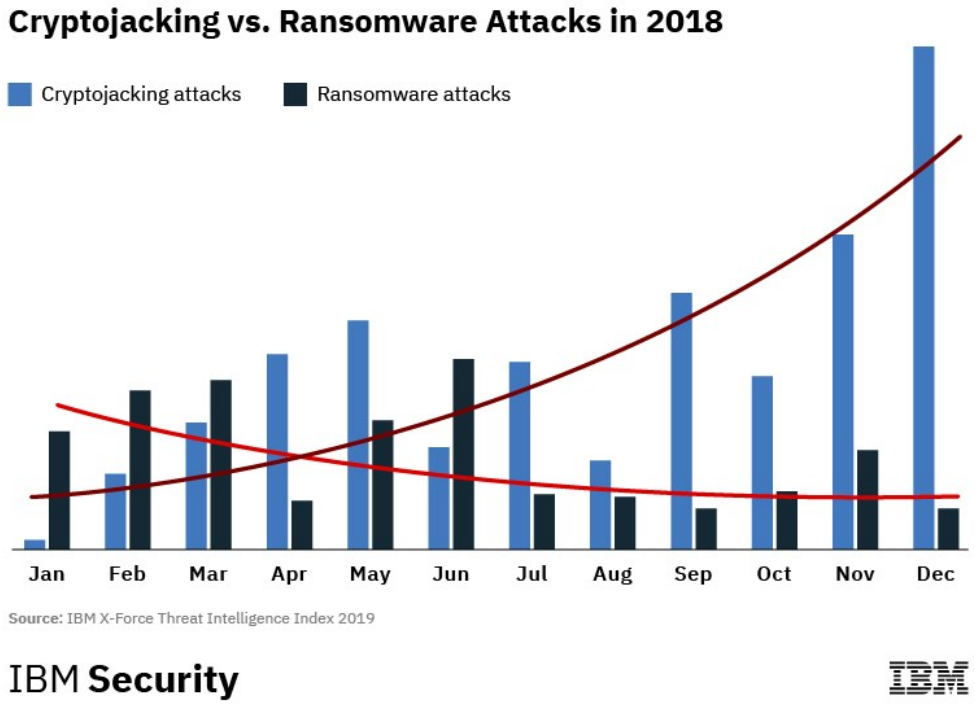

Figure 1 - comparaison entre attaque cryptojacking et ransomware en 2018

1.2 Le fishing

Le hameçonnage vise le plus souvent à usurper la crédulité des utilisateurs pour vous amener à cliquer sur une pièce jointe piégée ou un site web corrompu. Utilisant des techniques d'ingénierie sociale, cette attaque tire son efficacité de la connaissance du domaine (entreprise) ou de la personne (hobbies, famille). Le but de cette attaque est, au travers de mails ou sites d'apparence légitime, de récupérer des informations telles que vos noms d'utilisateur, mot de passe, carte de crédit. Informations directement exploitables ou revendues au marché noir sous forme de base de données. En 2019, une base nommée « Collection #1 à #5 » à été découverte regroupant 2.2 milliards d'identifiants, mails et mot de passe.

1.3 Les vulnérabilités

Les vulnérabilités qui seront plus particulièrement le sujet de cet article sont des faiblesses dans la conception d'un système d'exploitation. Ces failles permettent d'exploiter les systèmes informatiques pour en rendre le service inaccessible (on parle alors d'attaque DDoS) ou pour prendre le contrôle partiel ou total du système. Les failles récentes plus importantes , s'appellent "Spectre" et "Meltdown" et sont des vulnérabilités touchant les processeurs, ou "BlueKeep" une vulnérabilité du service RDP (Remote Desktop Protocol) des systèmes windows qui touche potentiellement plus

d’un million de machines non corrigées. Leur exploitation devient régulièrement publique, ce qui rend ces attaques plus dangereuses pour les systèmes dont les qui se sont pas corrigés. Les mises à jour régulières permettent de limiter les impacts de ce type d’attaque. Lorsque cela n’est pas possible, notamment pour les serveurs webs, des systèmes de patchs virtuels (WAF) permettent une mitigation de la vulnérabilité.

2 Le coût des coups

Les coûts liés à cette cybercriminalité en France ont augmenté de 23% entre 2017 et 2018 pour atteindre 8.6 millions d’euros [1] (Figure 2)

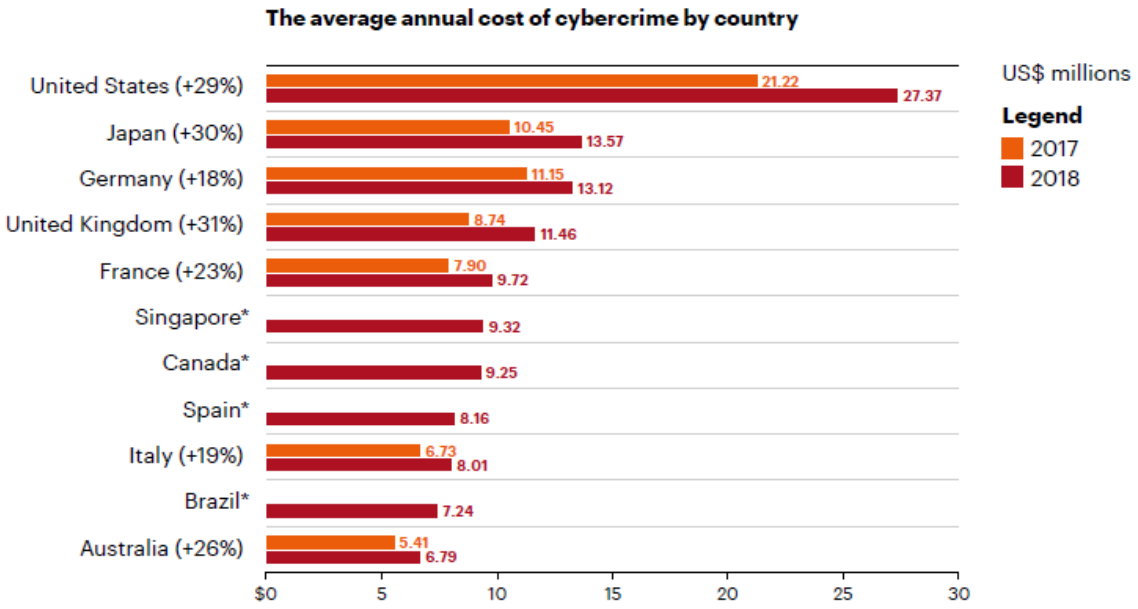

Figure 2 - Coût annuel de la cybercriminalité par pays

Cette prise de conscience est devenue un combat national traduit au fil des années, par l’application de la loi de Programmation militaire (LPM) imposée aux OIV (opérateurs d’importance vitale jugées indispensables pour la survie de la Nation), la directive Européenne NIS (Network ans Information Security), le renforcement de la cybersécurité des systèmes dits OSE (Opérateurs de Services Essentiels) qui fournissent "un service essentiel dont l’interruption aurait un impact significatif sur le fonctionnement de l’économie ou de la société", la mise en place de qualification PDIS,PASSI,PRIS pour assurer un niveau de qualité exemplaire pour les prestataires et fournisseurs de sécurité. Cette démarche vise à garantir le bon fonctionnement de notre société publique ou privé. Elle est dirigée par l’ANSSI (Agence Nationale de la Sécurité

des Systèmes d'Information) dont le budget 2019 atteint 100 millions d'euros pour la cybersécurité et un effectif d'environ 570 personnes..

Dans un monde presque totalement numérique et pour contrer les attaques de plus en plus sophistiquées et de plus en plus sévères, la détection des menaces est donc un objectif et un enjeu stratégique majeur. Une multitude de solutions de sécurité proposent des outils de détection de vulnérabilités. Nous proposons de faire un retour d'expérience sur la mise en place de ces outils sur les datacenter du Ministère de l'Education Nationale.

3 Un outil pour quoi faire ?

Nommé communément scanner de vulnérabilité, leurs fonctionnalités principales sont de connaître la surface d'exposition de votre système d'information en précisant les cibles que vous souhaitez superviser et d'en faire un état des lieux. Il s'agit aussi de vérifier les vulnérabilités de vos services, dans certains cas, la reconnaissance de vos sites web et, pour les plus évolués d'entre eux, de vérifier si vous êtes conforme à des règles existantes comme PSSIE (Politique de Sécurité des Systèmes d'Information de l'Etat), PCI-DSS (Payment Card Industry Data Security Standard) désignant les normes de sécurité des données applicables à l'industrie des cartes de paiement ou des règles établies par vos soins.

4 Principe de fonctionnement

Le principe de fonctionnement de ces outils est souvent identique avec une phase dite de "découverte" des assets (subnet, adresse IP, url site web) répondant à une requête ICMP, TCP, ARP afin de détecter les adresses actives sur le réseau. Ensuite, le scanner détermine les ports TCP et UDP (selon la demande de l'utilisateur) accessibles sur la cible active ainsi que le service associé (FTP, SSH, HTTPS, …) et sa version (Apache HTTP 2.4.8). Cette découverte du service et sa version utilise généralement la bannière présentée par le service ou les entêtes (fingerprint).

Une base de données des vulnérabilités est ensuite utilisée afin de déterminer ou non si le service découvert est vulnérable. Si la base de données la plus utilisée est publiée et maintenue par le MITRE, organisme à but non lucratif soutenu par le département de la Sécurité intérieure des États-Unis (https://www.cve.mitre.org), certains éditeurs proposent leur propre base comme par exemple RHSA pour RedHat. Normalisé par un identifiant unique sous la forme CVE-AAAA-NNNN (AAAA est l'année de publication et NNNN un numéro d'identifiant unique), les vulnérabilités sont

ainsi répertoriées, avec une description du risque, des systèmes affectés, des solutions possibles de remédiation ainsi que des liens pour avoir plus d'informations.

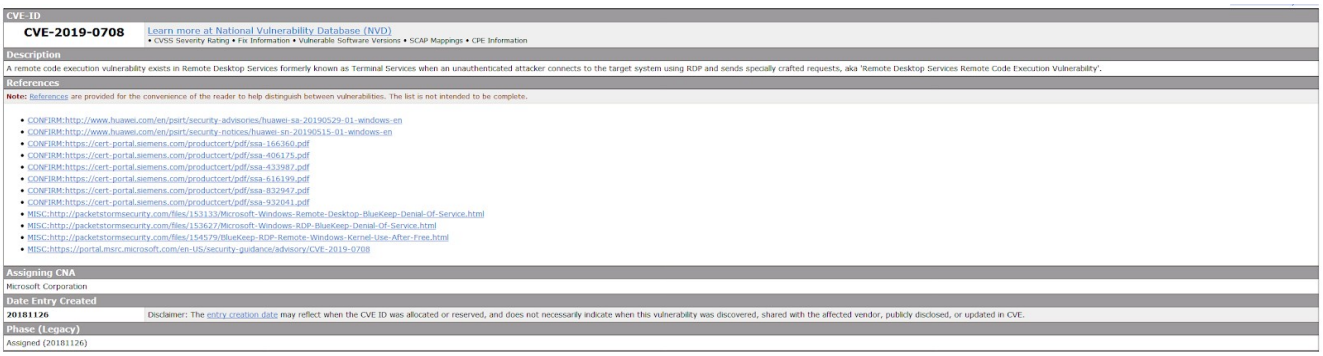

Figure 3 CVE-2019-0708 | Remote Desktop Services Remote Execution Vulnerability

Un autre principe de fonctionnement consiste à utiliser les outils réalisés pour exploiter certaines vulnérabilités afin de vérifier si oui ou non elle fonctionne sur l'adresse active auditée. Parmi les plus utilisés, Metasploit est un framework incontournable écrit en ruby et permettant de développer et d'exécuter les exploits (scripts permettant d'exploiter à son profit une vulnérabilité) contre une cible distante. Exploit-db (figure 4) est un site web qui centralise également un ensemble d'exploits qui peuvent récupérer, modifier et exécuter librement.

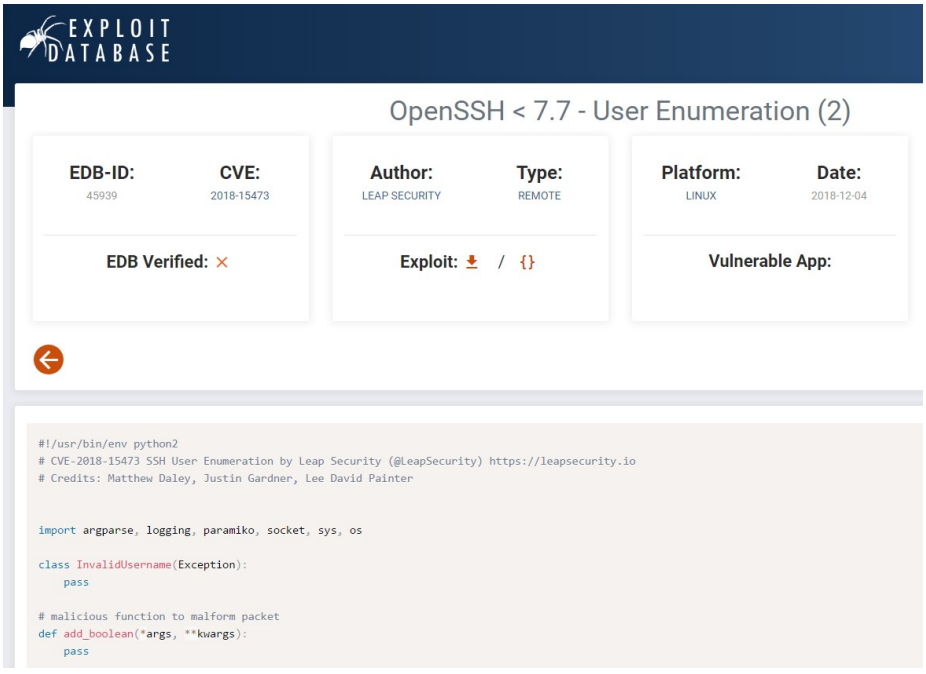

Figure 4 – script permettant l'exploitation de la CVE-15473

Si la détection des vulnérabilités sur les services web et versions détectées lors d'un scan sont relativement faciles à déterminer (Apache, IIS, Nginx...), les sites web nécessitent une attention plus importante. En effet, des erreurs de conception (code) ou de configuration dans l'application peuvent permettre à un attaquant de se connecter sans autorisation ou d'accéder à des parties cachées ou protégées. Des modules sont donc spécialisés dans ce type de détection en testant toutes les entrées possibles du site web audité en envoyant un nombre élevé de requêtes. L'OWASP (Open Web Application Security) recense les types de vulnérabilité des sites web et propose un outil de scan des sites web.

La détection de mauvaises configurations est aussi une des fonctionnalités importantes des scanners de vulnérabilité. Si le service est à jour et non vulnérable, une erreur de configuration peut permettre à un attaquant d'utiliser des identifiants par défaut pour se connecter ou d'accéder à des fichiers confidentiels non protégés. Plus difficile à déterminer par un scanner nous verrons des exemples de faux positifs ou d'oubli plus loin dans cet article. Citons par exemple, comme détection de mauvaise configuration, des cookies non sécurisés permettant le rejeux pour se connecter, les transferts de zone autorisés pour un serveur DNS, des configurations par défaut laissées sur des équipements ou sites Internet.

Pour terminer cette liste, la possibilité de se connecter sur la cible avec un compte afin d'effectuer des "scans authentifiés" permet de déterminer de manière exhaustive les programmes installés et leur vulnérabilités potentielles. Ceci permettra d'évaluer la surface d'attaque par rebond ou les possibilités d'élévation de privilège en cas de perte ou de vol de compte.

4.1 Classement des vulnérabilités

Les vulnérabilités sont alors classées par criticité :

- Critiques : prise de contrôle ou exécution de commande à distance, simple à mettre en œuvre.

- Majeures : vulnérabilités permettant une prise de contrôle ou une exécution de commande à distance complexe à mettre en œuvre.

-Moyennes : vulnérabilités ayant des impacts limités nécessitant des conditions initiales non triviales.

- Mineures : vulnérabilités ayant des impacts faibles ou nuls à moins d'être combinées à d'autres vulnérabilités plus importantes.

4.2 High-score

Chaque vulnérabilité dispose d'un score CVSS (Common Vulnerability Scoring System) de 0 à 10 prenant en compte le vecteur d'attaque (réseau, local, physique), la complexité de l'attaque (exploit existant, complexe à mettre en œuvre), les privilèges requis (aucun, avoir un compte), l'interaction (amener un utilisateur légitime à effectuer des opérations), le périmètre (rebond possible ou non), et les impacts sur la confidentialité/disponibilité/intégrité ou non.

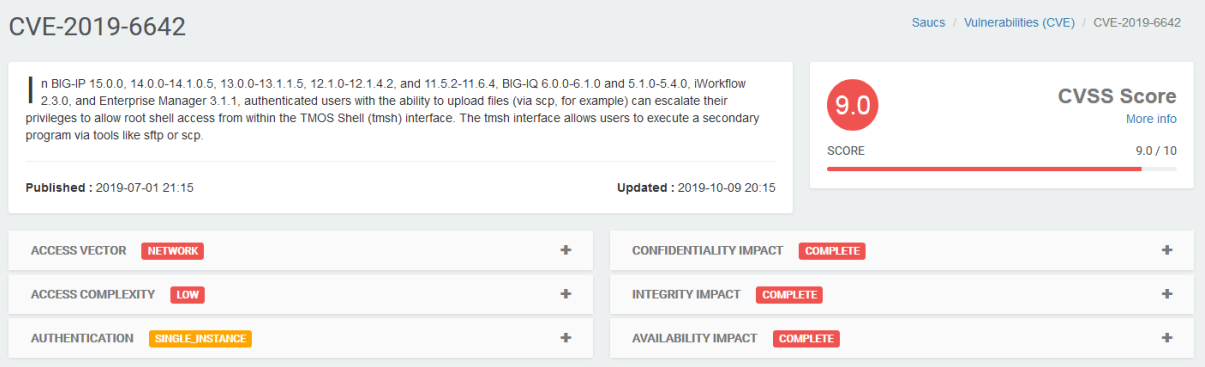

Figure 5 – score CVSS de la vulnérabilité CVE-2019-6642

Des rapports de vulnérabilités vont alors permettre aussi bien aux équipes opérationnelles de connaître leur surface d'attaque, leurs failles et comment y remédier, qu'aux COMEX, CODIR, COPIL, COMOP et autres comités exécutifs de présenter les résultats en matière de sécurité opérationnelle. L'utilisation régulière de scans, programmés dans le temps va également permettre de connaître l'évolution dans le temps de son Système d'Information, de challenger les équipes et de détecter rapidement de nouvelles vulnérabilités.

5 Anatomie d'une attaque

Le principe général d'une attaque informatique est de cumuler les vulnérabilités pour arriver à obtenir les privilèges les plus hauts de la machine et sa latéralisation pour compromettre d'autres systèmes avoisinants. Par exemple l'utilisation d'une vulnérabilité de drupal (CVE-2019-6340 medium) ou de wordpress permet d' obtenir un compte permettant de se connecter au CMS, de déposer un shell à partir de ce compte (image piégée), de se connecter à distance à la machine (reverse shell) d'utiliser des failles de noyau (dirtycow, Meltdown, spectre, zombieload, sudo) pour obtenir les privilèges hauts de la machine et de scanner le réseau alentour à la machine pour latéraliser son attaque.

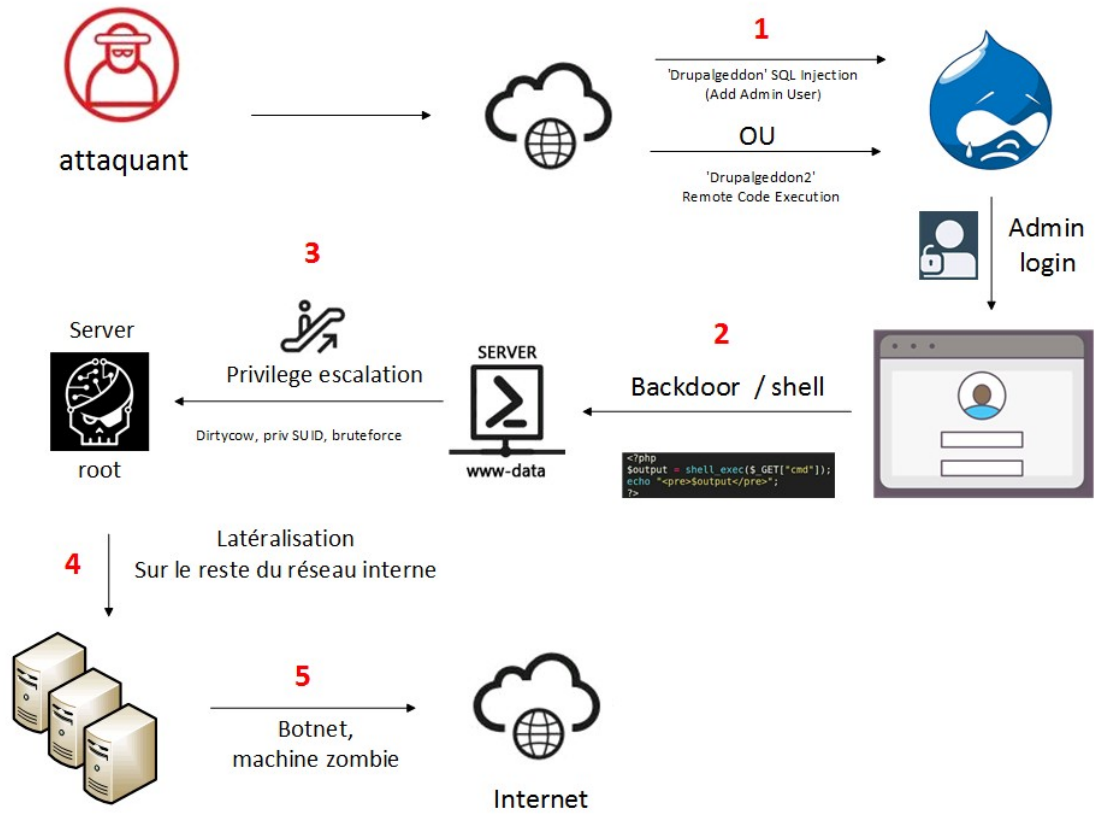

Figure 6 – anatomie d'une attaque

On voit donc que la détection en matière de sécurité opérationnelle est un ensemble complexe de type d'attaques et de contexte associés. Une multitude d'outils existe pour vérifier ses systèmes d'information, souvent libres, ils couvrent l'ensemble des domaines de la détection.

6 Quel outil choisir ?

Parmi les principaux outils, de manière non exhaustive:

- **Nmap** est spécialiste de la détection et de l'identification des services et des ports ouverts auquel on peut aujourd'hui adjoindre des scripts de détection de vulnérabilité.

- **Dirbuster** permet de reconstruire l'arborescence complète d'un site Web et ainsi vérifier l'accessibilité à certains dossiers ou données.

- **Nikto, Wapiti, Owasp-Zap** pour auditer les serveurs web

- Des outils spécialisés dans la détection web ou plus spécialistes de CMS (Content Management System) comme **wpscan** par exemple pour les sites Wordpress.

Des scripts d'exploitation d'une vulnérabilité sont également diffusés chaque jour, écrits en Perl, en Php, en Ruby, des plateformes comme exploit-db ou le framework **Metasploit** accélèrent leur diffusion. L'attaque est alors bien plus dangereuse car des néophytes, communément appelés "script kiddie" dépourvus de compétence en détection, peuvent infiltrer des systèmes en utilisant ces plateformes sans même en connaître les impacts et les conséquences. Il n'est pas rare de retrouver dans nos logs des attaques faites "à la maison" par ces "pirates amateurs". Incontournable également le logiciel **BurpSuite** est une application java spécialisée dans la détection et les tests d'intrusion sur des applications web. Il nécessite une bonne connaissance des architectures web pour en tirer le meilleur. Cette liste pourrait à elle seule faire l'objet d'une étude tant elle est croissante et évolutive.

Si l'exploitation de ces outils reste parfois triviale, une bonne connaissance de ce que l'on recherche est nécessaire. Très souvent chronophage, ces outils permettent de réaliser des audits à la demande mais rarement d'industrialiser leur utilisation sur un ensemble de serveurs importants comme un datacenter et de manière régulière. Une des bonnes pratiques consiste à cartographier et détecter régulièrement les vulnérabilités dans un système d'information , car le nombre et la criticité des vulnérabilités ainsi que leur exploitation est en augmentation dans le temps. Avec 16556 CVE en 2018 contre 14714 CVE en 2017, l'année 2019 compte déjà 11913 vulnérabilités par mois (source : cvedetails).

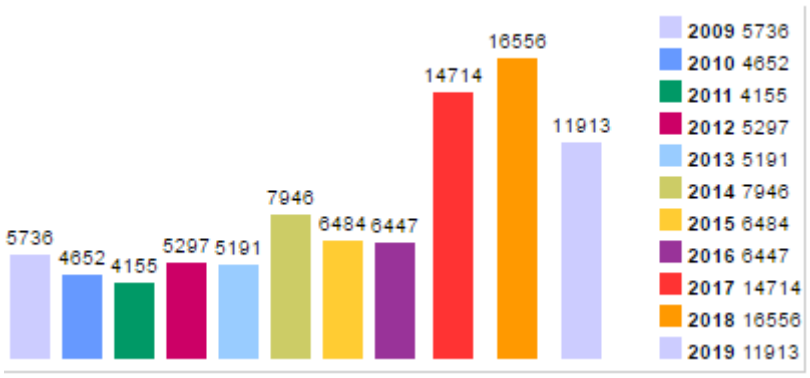

Figure 7 – Nombre de CVE de 2009 à 2019 (partiel)

7 Scanner de vulnérabilité

Il faut donc des plateformes regroupant les fonctions de ces principaux outils de façon à permettre à des utilisateurs qui ne sont pas nécessairement spécialistes du domaine de lancer de manière simple des détections sur un périmètre identifié et de manière récurrente afin de connaître l'état de leur système à un moment donné. Ces plateformes sont appelées scanners de vulnérabilités. Il en existe des OpenSource

comme Openvas ou sous licence (généralement à l'IP) comme Nessus (Tenable), Qualys VM (Qualys), InsightVM (Rapid7), Radar (F-secure) pour ne citer que les principales (source: gartner. Reviews market vulnerability-assessment).

Renater (Réseau de télécommunications français reliant les différents établissements enseignement recherche) propose un outil de détection des vulnérabilités SCan'ER depuis 2015 qui permet de cartographier son réseau (public) afin de vérifier les services ouverts et leurs potentielles vulnérabilités. L'utilisateur est libre de choisir de scanner sa cible tous les jours tous les 2 jours ou tous les 3 jours. Il n'y a pas de personnalisation sur le type de scan. Simple d'utilisation, le scanner Renater permet ainsi d'établir une cartographie de la menace sur son réseau et d'en assurer la surveillance.

Si l'outil permet d'avoir un véritable regard sur son système d'information, il faut cependant savoir en faire une bonne lecture.

Quel que soit l'outil, avec un classement et une notation des vulnérabilités, il est courant de ne s'attacher qu'aux vulnérabilités critiques et majeures.

7.1 L'aiguille dans la botte d'aiguille

Avec souvent beaucoup de vulnérabilités à corriger, il est parfois difficile de faire le tri. Pourtant, nous avons déjà trouvé des informations confidentielles comme des informations de connexion ou examen dans un pdf, un dossier, un fichier texte accessible sans restriction, remonté par le scanner comme mineure ou simple "information".

On peut le voir dans la Figure 8 ou nous avons trouvé des identifiants et mot de passe dans le dossier script. Il faut donc parcourir cet ensemble de dossier afin de garantir qu'il ne diffuse aucune information sensible. Et c'est souvent avec le métier et les développeurs que l'on peut assurer ce travail.

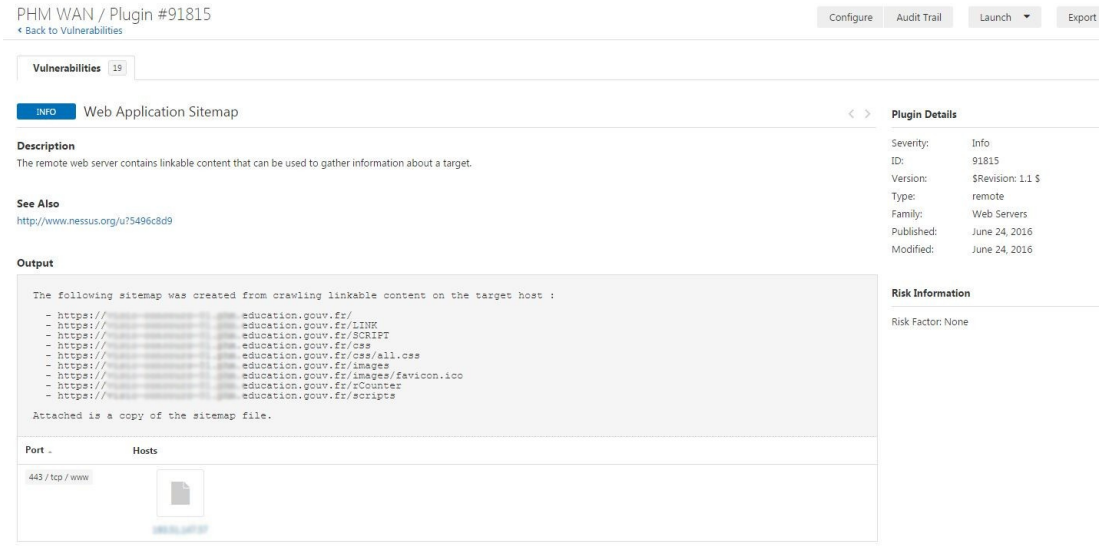

Figure 8 – des dossiers accessibles signalés comme « information »

Figure 8b – informations contenues dans un fichier pdf considéré comme « normal »

Une vulnérabilité peut être considérée comme médium alors qu'un script d'exploitation permet de l'exploiter de manière extrêmement simple et donc de corrompre un système d'information. Il n'est pas rare de trouver une vulnérabilité critique ou majeure qui nécessite un contexte trop spécifique et ne correspondant pas au

votre. Sans les négliger, il faut être en mesure de prioriser la résolution des failles en fonction de leur criticité du point de vue risque et de leur exploitabilité.

Si ces outils permettent pour certains de tester les formulaires web avec des injections basiques, il est rare qu'ils soient en mesure de personnaliser ces attaques. Un auditeur humain saura quant à lui reconnaître une variable dans laquelle il pourra injecter un code ou une redirection en adaptant, en customisant son contenu et en rejouant son attaque en fonction des retours du serveur alors que le scanner se limitera à quelques inputs.

Entre les faux positifs/négatifs ou les vulnérabilités manquées, il est impossible de garantir la santé de son SI uniquement sur la base de ces analyses automatisées. L'utilisation d'un seul outil est même parfois insuffisant. Souvent redondants, les résultats obtenus en utilisant deux outils peuvent être différents. Certains scanners web sont spécialisés sur la détection XSS, d'autres sur les injections SQL ou encore mieux adaptés aux systèmes basés sur JAVA. Il faut se pencher sur les résultats avec un œil d'expert afin de garantir une meilleur visibilité sur son SI.

De la même manière, un scan non authentifié ou authentifié pourra retourner des résultats différents et même suivant le rôle du compte utilisé pour le scan authentifié.

Un scan lent ou actif fera également réagir vos pare-feux, votre système de préventions des intrusions (IPS) ou vos pare-feux applicatifs web (WAF) de manière différentes. Il conviendra alors d'identifier la possibilité de contourner ces solutions. Il est donc important lorsque l'on utilise ces outils de pouvoir personnaliser leur mode de détection et leur degré d'intensité (nombre de requêtes simultanées). Il est nécessaire de connaître un minimum son infrastructure pour savoir si l'on doit lancer un scan web, sur des actifs réseaux, sur des services ssh. Les modules utilisés pour la détection seront alors différents.

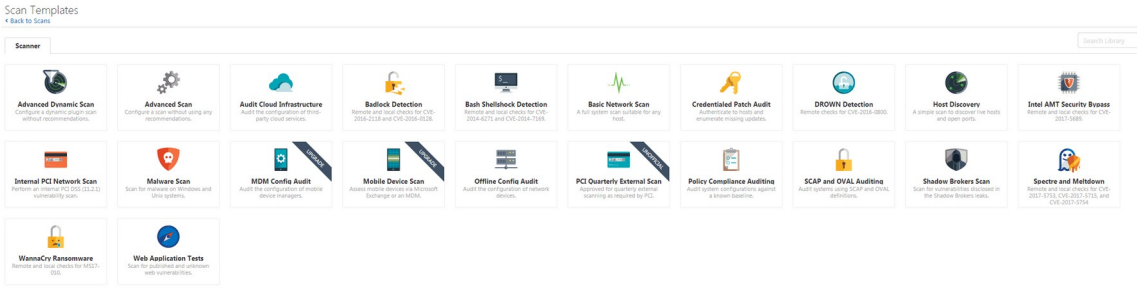

Figure 9 - Template de scan à choisir selon son architecture, le type de scan recherché.

Rares sont les outils qui permettent de vérifier les configurations. Par exemple si vous déployez un serveur Openldap dans sa dernière version, le scanner va bien vérifier la mise en oeuvre du protocole TLS, les potentielles vulnérabilités mais ne vous informera pas quevous n'avez pas supprimé les accès anonymes (NULL BIND).

Plus rares encore sont les outils permettant de valider la conformité à sa politique de sécurité. Notre ssh est à jour, bien configuré mais il est ouvert sur Internet or notre PSSI

me l’interdit. Il est également fondamental de connaître ses ressources afin d’identifier ce que l’on doit protéger. Maîtriser ou Périr ! Souvent complexes les études de risque ont tendance à ne pas être systématiques et rarement reconduites dans le temps.

Il n’en reste pas moins qu’en corrigeant 20 % des failles de sécurité web on couvre 80 % du risque (principe de Pareto). Ces outils sont donc indispensables dans la gestion des vulnérabilités et du cycle de vie du SI.

Obtenir rapidement des rapports sur la santé de ses systèmes, avoir des remédiations à proposer aux équipes d’exploitation, pouvoir effectuer des tests réguliers voilà donc l’objectif de ces outils de gestion de vulnérabilités. Se sentir en sécurité simplement avec les résultats de ces outils serait une erreur mais il est difficile et très chronophage de passer sur chaque serveur plusieurs outils de détection et de consacrer du temps à la recherche manuelle de possibles entrées permettant la compromission. Là encore l’intérêt de classifier ses environnements par criticité prend tout son sens. C’est donc un équilibre à trouver entre outil et méthode manuelle mais la ressource disponible pour ce travail n’existe pas toujours.

Aujourd’hui un scanner de vulnérabilité permet de contrôler les assets des deux datacenters du Ministère (environ 5000 IP), avec deux sondes (moteurs) de scans regroupant les résultats sur une console centrale. Parallèlement, nous utilisons toujours le scanner Renater pour la partie externe (croisement des résultats avec le scanner du Ministère). La partie exposée sur Internet et la partie interne totalement ouverte sur le réseau RACINE (Réseau d'Accès et de Consultation des INtranets de l'Education) sont contrôlées chaque semaine. La partie interne cloisonnée (serveurs d’applications, base de données, ressources transverses et d’infrastructure) est scannée de manière hebdomadaire ou mensuelle selon son exposition et sa criticité.

Avant chaque MEP (mise en production) d’un nouveau système d’information dans un des Datacenter, un audit complet est réalisé à partir des outils cités auparavant et un rapport complet est livré à l’exploitant. L’arrivée de système OSE (Opérateurs de Services Essentiels) oblige ce renforcement de sécurité dans la détection et le maintien en condition opérationnelle de ces systèmes. Parallèlement à ces audits en cas de vulnérabilité critique, une étude d’impact et un contrôle manuel sur les serveurs concernés est effectué. Dernier cas d’usage: CVE-2019-0708 du 14 mai 2019 vulnérabilité dans RDP, Blue Keep qui a fait l’objet d’une étude d’impact et d’une vérification complète hors scanner.

8 Objectif SOC

Le scanner de vulnérabilité ne doit néanmoins être qu'un des éléments de la sécurité opérationnelle à mettre en place. D'autres signaux doivent être introduits pour garantir une visibilité globale :

- des signaux internes tels que les logs, et tout le travail de corrélations issus d'un SIEM.

- des signaux externes tels que les indicateurs de compromission (IOC), CERT (Computer Emergency Response Team), Ip-file-mail réputation (avec système de notation), shodan, breachaware, Openbugbounty, mettre en place des honeypots (voir ce très bel article de Fabrice Prigent "10 ans de protection par Honeypot", Université Toulouse JRES 2017).

Si ces signaux passent par l'implémentation d'outils de SIEM, de gestion d'IOC, de gestion d'incident, de Bugbounty; l'apport contextuel de l'intelligence humaine et de la compétence dans la compilation de cet ensemble est fondamental. Là encore, on parle bien de ressources et pas d'outils. Ils sont gestionnaires de projet, auditeurs, consultants, analystes, experts en réponse à incident, juristes, et bien d'autres profils encore à assurer les métiers de la sécurité numérique. L'organisation de ces forces vives et logistiques est également le point central du bon fonctionnement.

La sécurité ne doit pas « être un échec » pour reprendre l'expression de Nicolas RUFF mais ne doit ni ne peut se résumer à « faire des chèques » pour acquérir des outils et se sentir protégé.

9 Le remède magique

Et pourtant depuis quelques années, l'arrivée de l'intelligence artificielle nous promet une solution miracle à tous nos cyber-maux. Le machine-learning (apprentissage statistique) devient un outil indispensable de détection d'anomalies soit par apprentissage supervisé, c'est à dire apprendre ce qui est normal et ce qui ne l'est pas, soit par comparaison avec ce qui est habituel.

Si le concept n'est pas nouveau, c'est l'arrivée du Big-data avec les nouvelles capacités technologiques de stockage et d'exécution qui va permettre de garantir un traitement sur les données et les métadonnées récoltées, la ou les capacités d'analyse humaines seraient impossibles.

En traitant une grande masse d'informations toujours plus volumineuse, le traitement automatique d'indices de validation d'un critère sera d'autant plus efficace face à l'expert que le volume d'indices est grand.

Surfant sur cette vague, le marché cyber propose toute une gamme de produits intégrant analyse comportementale, détection intelligente des menaces ou analyse intelligente de la sécurité.

On le voit, c’est donc en collectant toujours plus d’informations que les systèmes de machine learning apprennent à établir des modèles prédictifs. Mais cette intelligence est utilisée aussi bien dans un objectif de cyberdéfense que de cyberattaque.

Les pirates informatiques vont « noyer » leur présence en lançant des attaques génératrices de bruit afin d'affaiblir les systèmes de défense automatisés. En devenant une exception statistique, l’attaque a un bien plus de chance de succès en passant sous les radars du modèle statistique et de ses prédictions.

Plus on recueille d'informations, meilleures sont les chances de succès. C'est pourquoi les pirates collectent de grandes quantités de données pour améliorer les techniques d'ingénierie sociale et ainsi obtenir de meilleurs résultats dans l’obtention d’un accès non autorisé.

En utilisant les techniques de machine learning, les attaques se personnalisent et s’adaptent en fonction de la victime. Avec des malwares automatisés comme par Stuxnet ou Wannacry, on constate que les attaquants ont adopté l’utilisation du machine learning.

Les prévisions de menaces de McAfee 2018 indiquent que le développement de l'apprentissage automatique va se transformer en une course aux armements entre les défenseurs et les attaquants. Les défenseurs apprenant des attaquants, les attaquants assimilant les techniques de défense, une sorte d’AlphaZero en cybersécurité (l’apprentissage ne se fait pas à partir d’une base de données alimentées par des humains mais en apprenant simplement les règles du jeu et des algorithmes. A partir de là, le système a appris "seul" et en jouant "contre lui-même").

En développant des programmes malveillants sophistiqués pouvant échapper à la détection, même les meilleurs outils de machine learning échouent en raison du manque de maturité ou de précision du logiciel d'apprentissage automatique.

Sans aucune prétention, nous souhaitons simplement livrer ici notre retour d'expérience sur des outils intégrant le machine learning.

Notre première expérimentation concerne une plateforme de détection des anomalies par analyse du trafic réseau et d’apprentissage automatique. On nous promet de détecter toute anomalie en moins de 3 mois. Le résultat s’est avéré décevant avec deux alertes remontées pour une mise à jour et une synchronisation de serveur considérée comme fuite de données. Notre premier constat est que même avec un échantillon important de données, l’apprentissage du modèle doit être long et récursif.

Notre deuxième expérimentation avait pour thème l’accompagnement automatisé par apprentissage dans des règles de corrélation sur un SIEM. Le système a pu déterminer des règles habituelles de connexion des utilisateurs à leur ressource. Il nous permet de déterminer que Mme MICHU se connecte généralement la nuit depuis son domicile . Cela va nousaider à définir une règle générale sur ses habitudes de connexion et à vérifier les comportements anormaux. Dans la pratique, il s’agit d’automatiser les principales tâches d’investigation sur les incidents, jusqu’à la production de recommandations et/ou de remédiations. L’objectif a été rempli avec des recommandations qui ont permis d’augmenter le niveau de sécurité en remontant des

incidents liés à une mauvaise configuration en comparaison avec d'autres configurations existantes de même type. L'outil va potentiellement jusqu'à permettre de pousser automatiquement des règles sur vos pare-feu, load-balancer, proxy. Cependant, il faut savoir identifier les faux positifs qui pourraient entraîner un blocage total du système d'information en automatisant la remédiation aux incidents. Nous nous sommes contentés, dans le cadre de cette expérimentation, d'utiliser ces remontées pour qualifier et ensuite pour effectuer des actions manuelles.

Au vu des ces coûts très élevés, le machine learning n'est donc pas une solution miracle. Il doit venir en aide à l'expert, pour des tâches particulièrement longues ou difficiles. Dépourvu de créativité et d'intuition, il est une aide très précieuse pour accélérer la détection des menaces et assurer une meilleure protection contre les menaces avancées.

Avec peut-être un manque de maturité mais en constante évolution, les années à venir nous donneront certainement des résultats bien plus probants. Il ne faut cependant pas oublier que le maillon faible en matière de sécurité des données restera le facteur humain et que le retour sur investissement (bénéfice/coût) risquera donc d'être assez long.

10 Conclusion

L'augmentation du nombre d'attaques, ainsi que le nombre d'attaquants est une réalité, toutes les statistiques le montre. D'après les spécialistes, les outils utilisés à des fins malhonnêtes sont de plus en plus élaborés, spécifiques, faciles à utiliser et pour des coûts moindre. De l'autre côté si la cybersécurité commence à être vue comme un véritable enjeu, mettre en place une stratégie de cybersécurité complète et efficace reste difficile, chronophage et nécessite d'avoir des ressources pour le faire.

Pour garantir une sécurité optimale de son système d'information, connaître sa santé et la maintenir à un niveau acceptable est donc essentiel. Mettre en place un premier « socle » d'outils de sécurité pour cartographier, connaître ses faiblesses, commet y remédier tout en gardant un regard critique sur les résultats obtenus est une base incontournable.

Les questions ne sont pas : « suis-je bien protégé ? Mon système est-il compromis ? » car il est est difficile d'affirmer que son système est sécurisé à 100 %.

En revanche il est capital de pouvoir satisfaire à ces deux conditions :

- avoir des systèmes à jour pour pouvoir reconstruire ses serveurs en cas de perte totale. Nous avons déjà pu constater des reconstructions complexes voire impossibles parce que les système d'exploitation faisant tourner les applications obsolètes étaient introuvables en l'état.

- des sauvegardes régulières, validées et testées (restauration).

La résilience de son système à jour, sauvegardé et vérifié restera la seule garantie de faire face à une crise majeure afin de répondre à cette question : « en combien de temps puis-je remettre un système d'information fonctionnel ? »

Bibliographie - Liste de références

Internet Security Threat ReportVolume 24 | February 2019

accenture the cost of cybercrime

https://www.cvedetails.com/browse-by-date.php

https://www.gartner.com/reviews/market/vulnerability-assessment

http://www.clusib.be/wp/wp-content/uploads/2012/11/Risquesinformatiques_fr.pdf.pdf

https://conf-ng.jres.org/2017/document_revision_2613.html?download

https://www.ssi.gouv.fr/uploads/2016/05/liste_metiers_ssi_v4_secnumedu_anssi.pdf

https://www.sstic.org/2009/presentation/Pourquoi_la_securite_est_un_echec/