Cybersécurité, une initiation: Vie privée

Bruno Martin

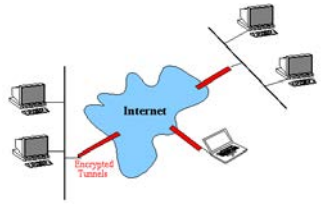
Université Côte d'Azur

M2 MIAGE SIRIS

Contenu

VPN

- Réseau privé qui utilise Internet pour connecter :
  - des pairs distants
  - des sites distants
- VPN utilise des connexions virtuelles routées au travers d'Internet vers l'entité distante



Fonctions :

- Etend la connectivité géographique
- améliore la sécurité
- réduit les coûts par rapport à un WAN dédié
- simplifie la topologie réseau

VPN – modes de fonctionnement

passent par l'interface `tun/tap` ; différence au niveau de la couche OSI

- **Mode routé (routed) :** relie des machines distantes au niveau de la couche 3 (réseau), ie au niveau d'IP via l'interface `tun`. Etablit route spécifique entre adresses réseau différentes. Pas de broadcast. Fonctionne en point à point.
- **Mode pont (bridge) :** relie des réseaux distants au niveau de la couche 2 (liaison) par protocole dédié `PPTP`, `EoIP`, `IPSec` via l'interface `tap`. ifaces VPN et LAN liées entre elles en une seule entité ; adresse VPN donnée dynamiquement au client. Routage entre réseaux fait par tables de routage au niveau du serveur VPN. Permet le broadcast et assure transparence complète.

Plan

Plan

Mécanismes de sécurité (suite)
- VPN
- Onion routing
- Firewalls
- Anti-virus
- Spam
- Détection d'intrusion

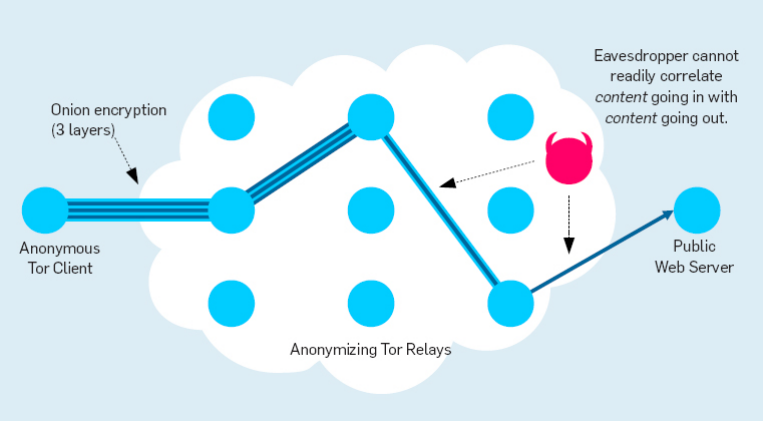
Vie Privée
- RGPD

Web tracking

Penetration testing, forensic, bug bounty

Pour terminer

Onion routing/TOR



Contrôle des frontières

- Problème principal des LAN connectés à Internet
- **solution** : utiliser des coupe-feux (firewall) avec :
  - filtres de paquets
  - proxy
  - mécanismes cryptographiques
- tout en diminuant le nombre de points d'entrée

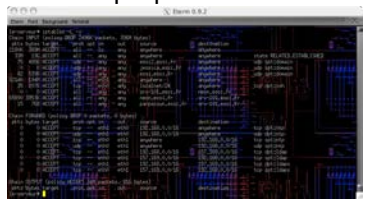
Filtres de paquets

- Fonction assurée par routeurs ou hôtes dédiés.
- Principe du contrôle :
  - redistribuer, effacer et/ou tracer chaque paquet
  - selon sur les informations d'en-tête des paquets
    - adresse source et destination
    - direction (entrée/sortie) par rapport au LAN
    - type d'application par numéro de port
  - en conjonction avec TCP/IP
  - généralement au niveau du noyau de l'OS

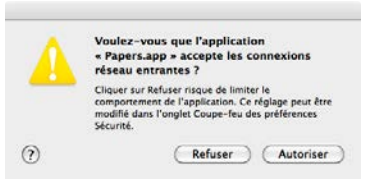
Un pare-feu désigne un logiciel et/ou un matériel (appliance), qui a pour fonction de faire respecter la politique de sécurité du réseau qui définit quelles sont les communications autorisées ou interdites.

Différents types de pare-feu

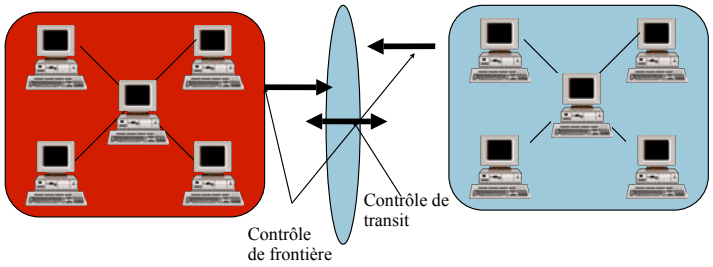
Filtre de paquets



Filtrage applicatif



Filtrage de paquets

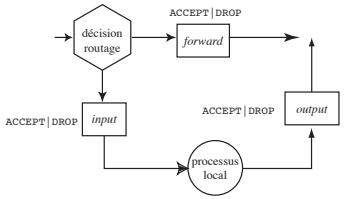


Différents selon l'OS

- **linux** : IPtables/netfilter ... nftables Firewall de paquets disponible sous linux. Composé de règles de filtrage
  - chaîne : liste ordonnée de règles
  - chaque règle exprime une condition
  - si règle i ne s'applique pas, consulter règle $i + 1$
  - une fois épuisé l'ensemble des règles, appliquer la politique par défaut de la chaîne (`ACCEPT`,`DROP`)
- **BSD** : 3 mécanismes différents :
  - **IPFILTER** : commande `ipf`
  - **IPFIREWALL** : commande `ipfw`
  - **PacketFilter** : commande `pf`
- **Windows** : Windows defender (ou de tierce partie)
- **MacOS** : 2 mécanismes intégrés, un provenant de BSD et un firewall applicatif.

IPtables – chaînes prédéfinies

`iptables` filtre les paquets qui traversent une machine au moyen des chaînes : `INPUT`, `OUTPUT` et `FORWARD`.



Limites d'un pare-feu

- toutes les communications doivent passer par le pare-feu
- le pare-feu doit être convenablement configuré
- éviter le contournement (modem, gsm...)
- éviter l'utilisation de clés usb, ordinateurs portables
- tenir un journal (logs)
- détecter les anomalies et/ou les intrusions

IPtables – routage

Un paquet arrivant sur une NIC, le noyau examine sa destination, par le routage.

- Destiné à la machine, il traverse la chaîne `INPUT`. S'il est autorisé à poursuivre son chemin (par un `ACCEPT`), il est traité par le processus local auquel il est destiné. Si la décision est `DROP`, le paquet est supprimé.
- Destiné à une autre interface, le paquet traverse la chaîne `FORWARD` et, s'il est accepté, il poursuit son chemin. Si le `forwarding` n'est pas activé ou si on ne sait pas comment transmettre ce paquet, le paquet est supprimé.
- Un processus local exécuté par la machine peut également envoyer des paquets traités par la chaîne `OUTPUT`.

Qu'est-ce qu'un virus ?

Un automate autoréplicatif à la base non malicieux, mais aujourd'hui souvent additionné de code malveillant, conçu pour se propager à d'autres ordinateurs en s'insérant dans des logiciels légitimes, appelés « hôtes ». Il perturbe plus ou moins gravement le fonctionnement de l'ordinateur infecté. Il peut se répandre par tout moyen d'échange de données numériques comme les mails les réseaux informatiques et les cédéroms, les clefs USB, les disques durs, etc.

Programmes souvent assimilés aux virus : le cheval de troie (qui hébergent souvent des ransmoware), les vers (qui n'infectent pas les fichiers mais se propagent aux autres ordinateurs)

Programme anti-virus

Protège l'ordinateur contre les virus et programmes assimilés.
Ces programmes contiennent un catalogue de miliers de virus connus qu'ils peuvent détecter et éradiquer.
Ils détectent les modifications standard apportées au fichiers par les virus pour rechercher de nouveaux virus.
On trouve des programmes anti-virus gratuits comme Avast ou ClamAV, proposés par l'OS comme Windows defender ou proposés comme applications tierces payantes par Symantec, McAfee, Norton,...

Bonne pratiques

En plus d'avoir un bon logiciel antivirus, il faut ajouter de bonnes pratiques :

- sauvegardes fréquentes : en cas d'attaque fructueuse, récupérer son travail
- installer des logiciels d'origine : ne pas faire l'économie du prix d'un logiciel et être sûr de sa source
- scanner son ordinateur fréquemment, surtout après le passage d'un réparateur ou d'un consultant
- ne jamais ouvrir les pièces jointes à un mail ou un lien sur un réseau social (mail, .doc, image, ...)
- vérifier tout ce qui est connecté à un système (physiquement ou logiquement)

Cibles favorites

Qu'est-ce que le spam ?

Du mail non sollicité, généralement pour vendre un produit, parfois sans ciblage, parfois avec (voir vie privée). Le spam généralise les courriers (papier) publicitaires, surtout à cause de son coût quasi nul pour inonder les destinataires.
Le spam a un coût :

- pour les destinataires : trier les bons mails des mauvais
- pour les fournisseurs d'accès : les spams sont transportés, acheminés et traités comme des mails "légitimes" avant d'être filtrés in fine avant leur (non)-distribution.

Une estimation : 75% des mails acheminés sont du spam.
Le spam présente aussi des risques car ils acheminent souvent des virus ou des liens vers des logiciels malicieux.

Catégories de spams

- **Publicitaires :** la plupart du spam provient de companies inconnues qui proposent des produits qui ne vous concernent pas.
- **Phishing :** plus dangereux, le phishing vous invite à divulguer des informations privées (numéro de carte de crédit, informations de connexion à des réseaux sociaux, des sites,..). Cela peut être ravageur, même avec des utilisateurs avertis
- **scams :** promettent de vous offrir quelque chose (de l'argent, un prix, une reconnaissance). Vous pensez qu'on peut gagner quelque chose avec Internet ?

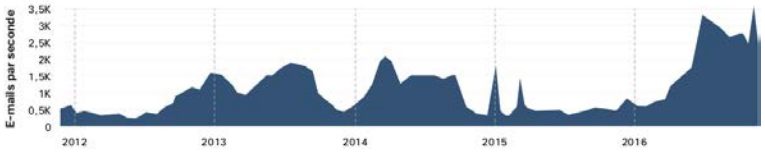
Lutte contre le spam

3 grandes catégories de lociels :

- **filtre anti-spam :** fonctionne en interaction avec le serveur mail de l'entreprise. A l'arrivée les mails sont filtrés, ceux légitimes délivrés et ceux douteux rangés dans un répertoire spécifique ou effacés
- **boitiers anti-spam :** proposés par des constructeurs, sous la forme d'un boitier. Aucune configuration, tout est fait par le fournisseur qui met à jour des listes noires, grises, d'IP douteuses, filtre avec des règles plus complexes. Solution onéreuse
- **service dématérialisé dans le cloud :** Le mail est récupéré par le service dédié, filtré et redistribué à l'utilisateur. Fonctionne selon le principe de l'abonnement.

Spam en chiffres

Figure 15 Volume total de spams



Source : CBL

Figure 16 Taille de la liste de blocage SCBL (SpamCop Blocking List)

Détection d'intrusion

Un système de détection d'intrusion (ou IDS : Intrusion Detection System) est destiné à repérer des activités anormales ou suspectes sur la cible analysée (un réseau ou un hôte). Il permet d'acquérir une connaissance sur les intrusions réussies ou échouées.
Classés en plusieurs catégories :

- **NIDS :** Network based IDS qui surveillent l'activité au niveau du réseau (exemple : `snort`)
- **HIDS :** Host based IDS qui surveillent l'activité au niveau des hôtes
- **IDS hybrides :** qui combinent les 2 précédents

Contenu

Classification Privacy (Solove, 2006)

- collecte d'information
  - surveillance
  - interrogation
- traitement de l'information
  - agrégation
  - identification
  - insécurité
  - usage détourné
  - exclusion
- dissémination de l'information
  - perte de confidentialité
  - divulgation
  - exposition
  - accessibilité accrue
  - blackmail
  - appropriation
  - distortion
- invasion
  - intrusion
  - interférence décisionnelle

Définition Privacy

Concept abstrait et subjectif qui dépend de la discipline d'étude, de normes sociales et des attentes sociétales ainsi que du contexte.

- Point de vue légal : "le droit à la tranquillité" ou the right to be let alone ou le droit de chaque personne de décider quelle information personnelle peut être divulguée aux autres et dans quelles circonstances
- Point de vue psychologique : "la liberté de construire son identité propre dans un environnement contraint" car la construction de son identité dépend de la vision d'autrui ; cf. réseaux sociaux, profilage,....

Protection des données

- 1995 : Directive Européenne sur la protection des données
- 2016 : RGPD (règlement général sur la protection des données)
- s'applique aux "données personnelles" : toute information relative à un individu (pas d'application aux activités de sécurité nationale ou juridiques)
- Règlement sur la protection des personnes physiques relatif au traitement des données personnelles et au mouvement de ces données.

RGPD : les principes

- loyauté : ce qui est traité doit correspondre à ce qui a été décrit
- transparence : les personnes ont le droit d'obtenir les infos nécessaires pour assurer un traitement loyal
- limitation des finalités : les données ne peuvent être obtenues que pour des finalités déterminées, explicites et légitimes
- minimisation des données : la quantité minimale de données doit être recueillie et conservée pour un traitement spécifique.
- exactitude : les détenteurs de données doivent créer des processus de rectification et suppression dans les BD.
- droit d'accès et de rectification :
- limitation de la conservation
- intégrité et confidentialité
- responsabilité : Le responsable du traitement des données doit être capable de démontrer sa conformité avec la totalité des autres principes

Monde offline $\rightarrow$ monde online

- information difficile à collecter, mémoriser, chercher et accéder
  - conversation F2F
  - documents papier
  - paiement liquide
  - filature
  - trouver vos relations
  - recherche dans dictionnaires
- difficile de copier, diffuser, facile à détruire
- difficile à agréger, profiler et inférer
- oubli dans le temps
- ...

- information facile à collecter, mémoriser, chercher et accéder
  - messagerie instantanée
  - mails
  - fichiers numériques
  - paiement par carte
  - géolocalisation
  - amis en ligne
  - requêtes google,. . .
- facile à copier, diffuser, dur à détruire
- facile à agréger, profiler et inférer
- information jamais perdue
- ...

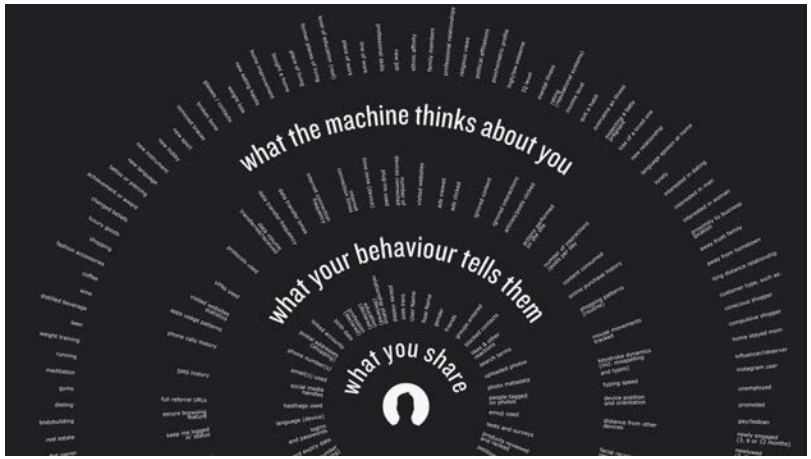
Protection des données : introduction

La CNIL en collaboration avec un Youtuber a présenté une introduction au RGPG : RGPD en video réponse aux questions en Video

Rien à cacher ?

- l'argument "je n'ai rien à cacher" repose sur l'hypothèse sous-jacente que la vie privée est de cacher des actions négatives
- Ce qui rend une société agréable à vivre est sa capacité à protéger les individus de l'intrusion des autres dans la vie privée. Une société sans protection de la vie privée deviendrait vite étouffante.
- Ecart entre secret et privé :
  - vos actions quotidiennes, vos amis ou relations, vos propos dans une conversation, vos hobbies,...
  - tout ce qui précède n'est peut-être pas secret mais vous n'avez pas forcément envie de rendre ces activités publiques ou accessibles à d'autres qui peuvent les analyser et en tirer des conclusions

Niveaux de vie privée



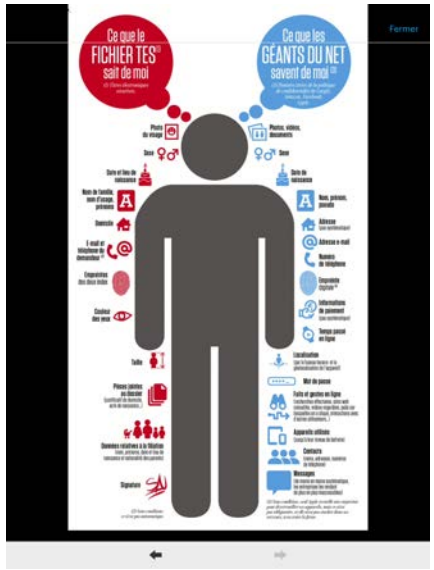
Vu sur capital

Mieux vaut lire les conditions d'utilisation avant de s'inscrire sur Tinder. Le géant américain Match Group qui possède plus de 130 sites de rencontre à travers le monde comme Meetic ou OkCupid collecte la majorité des données de ses utilisateurs. Rien y échappe : position géographique, adresses mail, profession, loisirs, photos, mensurations et mêmes les préférences sexuelles.

Surtout, le groupe n'hésite pas à revendre ces données à qui le souhaite. L'ONG Tactical Tech et la chercheuse Joana Moli ont pu facilement acquérir ces informations concernant des millions de personnes pour environ 135 euros sur le site US Date. Un achat tout à fait légal. L'utilisateur donne son accord pour cette revente de données lorsqu'il accepte les conditions générales d'utilisation pendant son inscription. Il ne peut donc pas attaquer Match Group sur la vente de ces données.

Cette pratique est courante dans l'industrie selon Tactical Tech. L'ONG craint surtout que la récolte de ces informations soit utilisée à mauvaise escient par d'autres groupes privés ou administratifs. "L'utilisateur pourrait se voir imposer des restrictions sur son assurance maladie, ses demandes de crédits, son accès à l'éducation et bien plus encore. Exploiter ce genre de données très intimes peut causer des dégâts catastrophiques sur la vie des personnes concernées. Surtout si leurs profils est rendu public", conclut Tactical Tech.

Fichier TES, validé 18/10/18

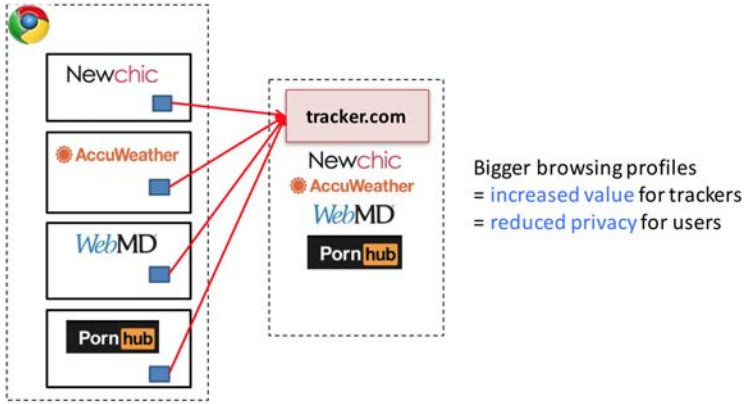


Vie privée et technologie

- ligne rouge : nos actions et interactions sont de plus en plus tracées par la technologie
  - on laisse des traces partout
  - ces traces sont combinées, agrégées et analysées pour déduire plus d'informations sur nous et de prendre des décisions qui auront un impact
  - nous n'avons pas de contrôle sur nos informations ou sur les déductions qui en découlent
- l'information n'est jamais oubliée
  - mais peut être parfois utilisée hors contexte

Plan

Web tracking (N. Bielova, INRIA)



Exemple

Navigation sur un site de presse :

Avec 21 traceurs !
D'après Springer : "le contenu journalistique est seulement un prétexte pour que le lecteur regarde les publicités".

Pourquoi s'inquiéter ?

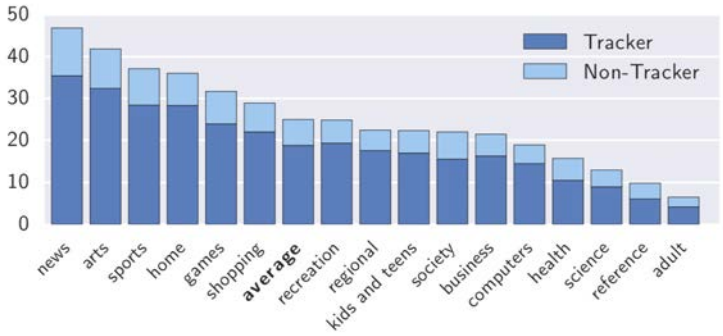
- Collecte de nos données sans qu'on le sache
  - sur des sites sensibles
  - mémoire de nos habitudes de navigation sur le web, de nos préférences, envies, humeurs
- utilisation de nos données
  - ciblage publicitaire
  - manipulation

Cambridge Analytica

Comment ça marche ?

- Par les cookies
- par d'autres mécanismes de stockage et les cookies zombies
- par des mesures à large échelle
- avec des publicités ciblées et du cookie synching
- par l'empreinte de votre navigateur

Combien de traceurs par site ?



Authentification par les cookies

1. l'utilisateur s'authentifie par un login/password sur un site
2. le site lui envoie une cookie d'identification
3. le cookie est inclus dans les échanges pour identifier la connexion de l'utilisateur
4. à la deconnexion de l'utilisateur, le serveur termine la session et invalide le cookie

Le cookie d'identification est une chaîne de caractères aléatoire et unique, pe : D6F8B2BE3ED3040D9A3C10

Usurpation d'identité par les cookies

Un attaquant vole le cookie de l'utilisateur et se connecte en même temps.... Que se passe-t-il ?
Il se fait passer pour l'utilisateur légitime et il usurpe l'identité de la victime et peut accéder à son compte.
C'est possible

- en écoutant le traffic `http` et en interceptant le cookie
- en dérobant le cookie sur le poste de travail (vulnérabilité du système, par ingénierie sociale, via une faille du serveur)

Termes

- **Penetration test :** méthode d'évaluation de la sécurité d'un hôte ou réseau en simulant une attaque. Pour ça :
  - rechercher les points d'accès
  - rechercher les vulnérabilités

  selon différentes approches :
  - **Black box (covert)** : pas de connaissance de l'infrastructure ; effacer ses traces
  - **White box (overt)** : infrastructure connue, avec RSSI
  - et les variantes entre les deux (grey box).
- **Forensic** : Terme adapté de l'anglais «computer forensics», l'expression « investigation numérique » représente l'utilisation de techniques spécialisées dans la collecte, l'identification, la description, la sécurisation, l'extraction, l'authentification, l'analyse, l'interprétation et l'explication de l'information numérique.

Plan

Phases du PenTest

- Pre-engagement interaction : négociation avec client : "contrat"
- Intelligence gathering : récupération de toutes les infos possibles sur le client (réseaux sociaux, scan, footprint,...)
- Threat modeling : utiliser les infos de l'IG pour identifier les vulnérabilités, choisir les attaques en fonction des buts cherchés
- Vulnerability analysis : trouver les attaques possibles en fonction de l'analyse des ports et des vulnérabilités,...
- Exploitation : réalisation d'exploits
- Post exploitation : attaques en whitehat
- Reporting : rapporter le détail des opérations menées

Intelligence gathering

- un bon hacker (bidouilleur) programme un outil pour scanner (explorer) le réseau
- il le publie sur Internet
- un script kiddie (novice) l'utilise pour trouver des systèmes vulnérables ou des points d'accès

Vulnérabilités

- le novice utilise ensuite une liste d'IP vulnérables pour accéder au système
- selon les faiblesses, il peut éventuellement créer/utiliser un compte ou un accès privilégié
- il l'utilise pour acquérir de nouveaux privilèges et pour pirater de nouveaux systèmes connectés à sa 1re victime
- exemple : se faire passer pour une machine du réseau attaqué (avec `wireshark` ou `ettercap`)

Scanner de ports : `nmap`

Un port scanner peut parcourir une grande plage d'adresses IP et retourner les ports ouverts (donc les services accessibles) ainsi que les version d'OS

Scanner de vulnérabilités – Nessus

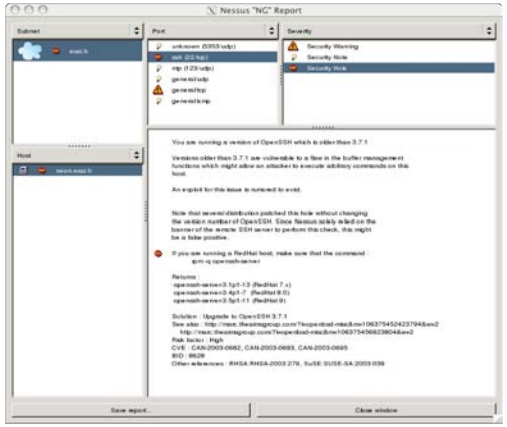
The "Nessus" Project aims to provide to the internet community a free, powerful, up-to-date and easy to use remote security scanner, software which will audit remotely a given network and determine whether someone may break into it, or misuse it. Nessus does not take anything for granted. That is, it will not consider that a given service is running on a fixed port - a web server on port 1234, will be detected it and its security tested. Nessus is fast, reliable and has a modular architecture that allows you to fit it to your needs. Nessus works on Unix-like systems (MacOS X, FreeBSD, Linux, Solaris and more) and a Windows version called NeWT is available.

Nessus : cibler

Nessus - OpenVAS

Nessus est devenu payant (cher) en 2005.
OpenVAS représente la branche « libre » de nessus. Contenait en 2011 plus de 23 000 tests de vulnérabilité, reliés à la base “Common Vulnerabilities and Exposures” CVE qu’on peut interroger
Il est possible d’ajouter des plugins dans le langage NASL, comme dans nessus.
`http://www.openvas.org/`

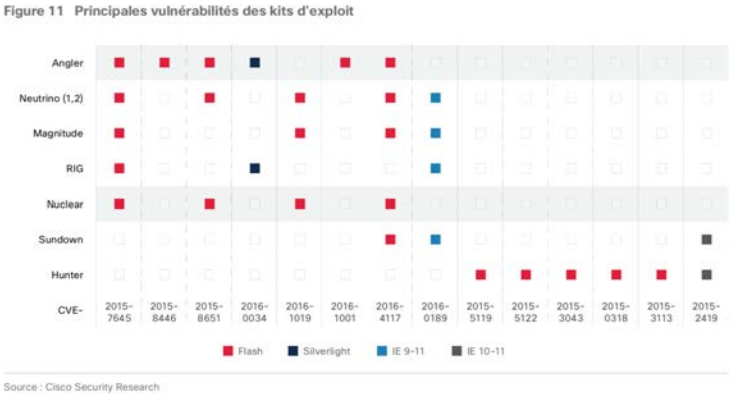
Nessus : scanner



Attaquer

- généralement au moyen de `rootkits`
- un `rootkit` est un terme qui décrit un ensemble de scripts et d’exécutables qui permettent à un pirate de cacher ses agissements et d’obtenir un accès privilégié au système :
  - modifie les `logs`
  - modifie les outils système pour rendre la détection du piratage difficile
  - crée une trappe d’accès cachée
  - utilise le système comme point d’entrée sur les autres hôtes du LAN

Kits d'exploits



Voir rapport CISCO

Bug bounty

Un bug bounty est un programme proposé par de nombreux sites web (depuis 1995 avec Netscape) et développeurs de logiciel qui permet à des personnes de recevoir reconnaissance et récompense après avoir reporté des bugs, surtout ceux concernant des exploits et des vulnérabilités.
Il y a bien sûr des règles à respecter et chaque Bug Bounty doit énoncer clairement les limites que le hacker ou l'expert ne doit pas franchir, mais en général, comme ça se passe sur des services en production, il vaut mieux éviter de tout casser si on veut sa récompense ;-).
En 2015, M. Litchfield dit avoir gagné plus de 300000$ en trouvant des failles.
Source korben

Framework Metasploit

Metasploit (écrit en `ruby`) est un framework qui permet :

- de collecter le résultat des différents scanners (port, vulnérabilité,. . .)
- d'automatiser (et de rejouer) des attaques contre des vulnérabilités identifiées (et d'en ajouter)

outil très puissant mais compliqué à utiliser (gui : armitage)

Plan

Les deux piliers de la cybersécurité

- **Prévenir** : éviter de se faire attaquer (au moyen de firewall, programmes antivirus, penser à faire les mises à jour, vérifier les mails entrants)
- **Reprise d'activité** : Il n'est pas possible de prévenir tous les risques. Des attaques fructueuses peuvent toujours être menées. Il faut prévoir des sauvegardes et un plan de reprise d'activité

Les 10 règles de la sécurité

- Sécuriser les points faibles
- Opérer en profondeur
- Bien gérer les cas d'erreur
- Principe du strict minimum
- Cloisonner
- Rester simple
- Encourager le secret
- Il est difficile de garder un secret
- Rester méfiant
- Utiliser les ressources de la communauté