Tableaux de bord de la sécurité réseau

3e édition

Cédric Llorens

Laurent Levier

Denis Valois

Benjamin Morin

Avec la contribution de Olivier Salvatori

ISBN : 978-2-212-12821-5

EYROLLES

Table des matières

PARTIE III

Les techniques de protection du réseau

Partie V

Les techniques de contrôle de la sécurité réseau

1

Typologie des attaques réseau

Les attaques réseau sont aujourd'hui si nombreuses qu'il serait illusoire de prétendre les décrire toutes.

Il est cependant possible de dresser une typologie des faiblesses de sécurité afin de mieux appréhender ces attaques, qui ont pour point commun d'exploiter des faiblesses de sécurité.

L'objectif de ce chapitre est de présenter les faiblesses les plus couramment exploitées par les attaques et de détailler les mécanismes de ces attaques. Nous espérons de la sorte faire comprendre les dangers qui menacent les réseaux, et non de susciter des vocations de piraterie, au demeurant réprimandées par la loi.

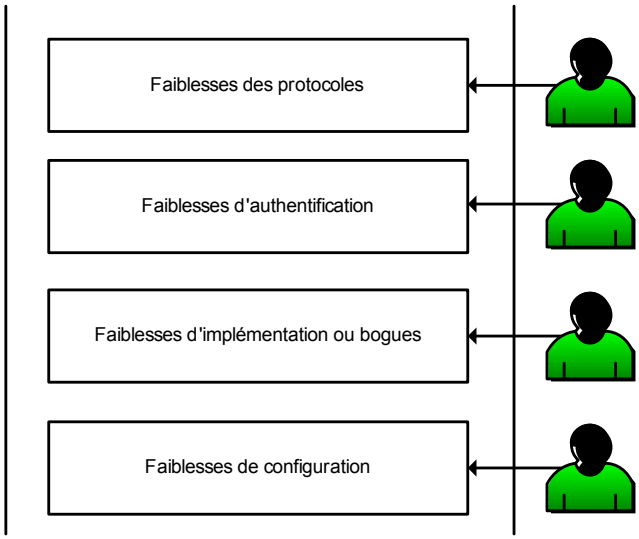
Comme tout effet a une cause, les attaques réseau s'appuient sur divers types de faiblesses, que l'on peut classifier par catégorie, comme illustré à la figure 1.2.

Les protocoles réseau sont encore jeunes, et aucun d'eux n'a été conçu pour tenir compte des problèmes de sécurité. Le protocole IP, par exemple, ne comporte pas de couche sécurité. La plupart des protocoles utilisés dans un réseau, tels SNMP (Simple Network Management Protocol) pour la supervision ou BGP (Border Gateway Protocol) pour le routage, n'implémentent pas de véritable couche de sécurité et s'exposent à diverses attaques, comme les attaques par fragmentation, déni de service, etc.

De même, les protocoles réseau n'ont prévu aucun mécanisme d'authentification véritable et subissent des attaques qui s'appuient sur ces faiblesses d'authentification, comme les attaques de type spoofing, man-in-the-middle, etc.

Figure 1.2

Typologie des faiblesses de sécurité



Les faiblesses d'implémentation ou bogues des programmes (système d'exploitation, application de routage, etc.) exposent à d'autres attaques, de loin les plus importantes en nombre. La raison à cela est que le développement des logiciels et des piles réseau se fait de plus en plus rapidement et sans règles strictes. Parmi les innombrables attaques qui utilisent de mauvaises implémentations ou des erreurs de programmation, citons les attaques de type SYN flooding et ping-of-death.

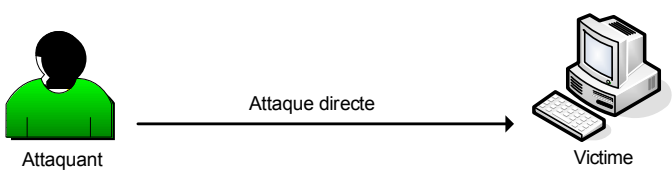
Les faiblesses de configuration des équipements réseau peuvent provenir d'une mauvaise configuration d'un pare-feu, laissant passer du trafic non autorisé par la politique de sécurité, ou d'un équipement réseau, permettant à un attaquant d'y accéder, etc.

En s'appuyant sur ces faiblesses, le pirate peut lancer un ensemble d'attaques permettant d'influencer le comportement du réseau ou de récolter des informations importantes.

Les attaques réseau peuvent être lancées directement, le pirate attaquant sa victime et exposant ainsi son identité, comme l'illustre la figure 1.3.

Figure 1.3

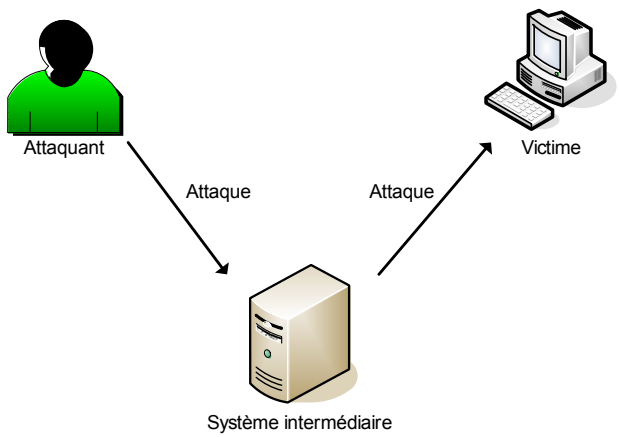
Attaque directe



Les attaques réseau peuvent aussi être lancées indirectement par l'intermédiaire d'un système rebond afin de masquer l'identité (adresse IP) du pirate et d'utiliser les ressources du système intermédiaire. Les paquets d'attaque sont dans ce cas envoyés au système intermédiaire, lequel répercute l'attaque vers le système cible, comme l'illustre la figure 1.4.

Figure 1.4

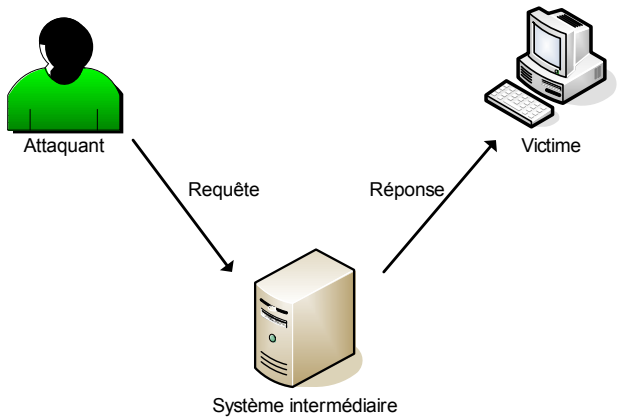
Attaque indirecte par rebond



Certaines attaques, dites indirectes par réponse, offrent au pirate les mêmes avantages que les attaques par rebond. Au lieu d'envoyer l'attaque au système intermédiaire pour qu'il la répercute, l'attaquant lui envoie une requête, et c'est la réponse à cette requête qui est envoyée au système cible, comme l'illustre la figure 1.5.

Figure 1.5

Attaque indirecte par réponse



Gardons à l'esprit qu'un réseau est la composante de plusieurs réseaux, provenant d'opérateurs différents (Internet, infrastructures publiques, etc.), *a priori* indignes de confiance.

Nous décrivons dans ce chapitre un ensemble d'attaques classées en fonction des objectifs des pirates et reposant sur des faiblesses protocolaires, d'authentification ou d'implémentation.

Enfin, et sachant que la future version du protocole IPv4 est IPv6 (déjà en phase expérimentale, voire mise en œuvre par certains opérateurs), nous décrirons si nécessaire la projection des attaques IPv4 dans un monde IPv6.

Attaques permettant de dévoiler le réseau

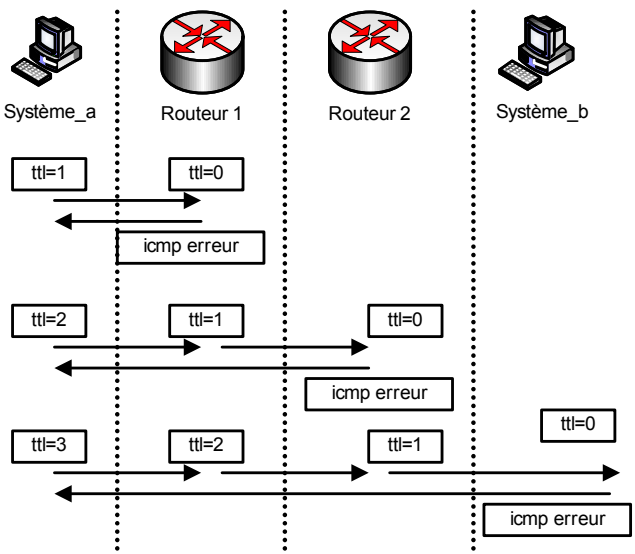
Attaque par cartographie du réseau

Les attaques visant à établir la cartographie d'un réseau ont pour but de dresser les artères de communication des futurs systèmes cibles. Elles ont recours pour cela à des outils de diagnostic tels que Traceroute, qui permet de visualiser le chemin suivi par un paquet IP d'un hôte à un autre.

Traceroute utilise l'option durée de vie, ou TTL (Time To Live) du paquet IP pour émettre un message ICMP time_exceeded (temps dépassé) pour chaque routeur qu'il traverse. Sachant que chaque routeur qui manipule un paquet décrémente le champ TTL, ce champ devient un véritable compteur de tronçon et permet de déterminer l'itinéraire précis suivi par les paquets IP vers un système cible, comme l'illustre la figure 1.6.

Figure 1.6

Fonctionnement de l'outil Traceroute



Traceroute crée un paquet avec les adresses source et destination et une valeur de durée de vie TTL initiale (nombre de passerelles traversées) égale à 1. Ce paquet s'arrête donc au premier routeur rencontré, et le routeur envoie un message d'erreur ICMP (time_exceeded). Traceroute enregistre cette information et crée un nouveau paquet avec un TTL de 2.

La traversée du premier routeur met le TTL à 1. Le paquet génère une erreur sur le deuxième routeur. Comme précédemment, le deuxième routeur envoie un message d'erreur ICMP avec son adresse, laquelle est mémorisée par Traceroute. Une fois le système cible atteint, une erreur ICMP est générée par ce système cible, et Traceroute affiche la liste des passerelles traversées ainsi que le RTT (Round Trip Time), ou temps aller-retour, pour chacune d'elles.

L'établissement de la topologie réseau n'est pas innocent et représente la première étape d'une future attaque des systèmes réseau. Dans le cas le plus fréquent, le pirate utilise plutôt la technique du balayage (scanning) pour construire l'image du réseau, car elle fournit des informations plus rapidement.

Projection dans un monde IPv6

Une adresse IPv6 est codée sur 128 bits, contre 32 bits pour une adresse IPv4. Le saut est vertigineux si l'on compare le nombre d'adresses possibles entre les deux mondes. À titre d'exemple :

- En IPv4, on trouve des tailles de sous-réseaux de l'ordre de 2^8 ou 2^{16}. Elles représentent de 256 à 65 536 adresses, qui peuvent facilement faire l'objet d'une cartographie complète par des outils du marché (ordre de grandeur de l'heure pour une cartographie complète).
- En IPv6, on trouvera des tailles de sous-réseaux de l'ordre de 2^{64}, représentant près de 180 milliards de milliards d'adresses et ne pouvant dès lors faire l'objet d'une cartographie complète qu'en millions d'années.

Bien qu'il soit difficile d'imaginer réaliser une cartographie active en IPv6 d'un, il est plus que probable que les adresses IP ne seront pas attribuées de manière aléatoire dans l'espace des 2^{64} disponibles dans un sous-réseau donné, contribuant ainsi à la limitation de l'espace à cartographier. Pour adresser cette problématique, l'IETF a normalisé la possibilité de créer des identifiants d'interface non seulement pseudo-aléatoires, mais aussi variables au cours du temps.

En tout état de cause, les cartographies actives et régulières faites en IPv4 ne pourront être portées dans un monde IPv6 sans une connaissance précise de l'espace IPv6 « réduit » à considérer.

Attaque par identification des systèmes réseau

Certaines attaques visent à identifier tous les systèmes présents dans le but de dresser les futurs moyens de pénétration du réseau ou des systèmes qui le composent.

Il existe pour cela différentes techniques de balayage des systèmes, comme l'illustre la figure 1.7.

Nous n'abordons dans ce chapitre que les techniques de base visant à découvrir les éléments du réseau. Les techniques avancées (furtives, etc.) sont abordées au chapitre 2, qui traite des attaques orientées système.

Attaque par balayage ICMP

La méthode de balayage la plus simple consiste à utiliser le protocole ICMP et sa fonction request, plus connue sous le nom de ping. Elle consiste à ce que le client envoie vers le serveur un paquet ICMP echo-request, le serveur répondant (normalement) par un paquet ICMP echo-reply, comme l'illustre la figure 1.8. Toute machine ayant une adresse IP est un serveur ICMP.

Figure 1.7

Les différents types de balayages

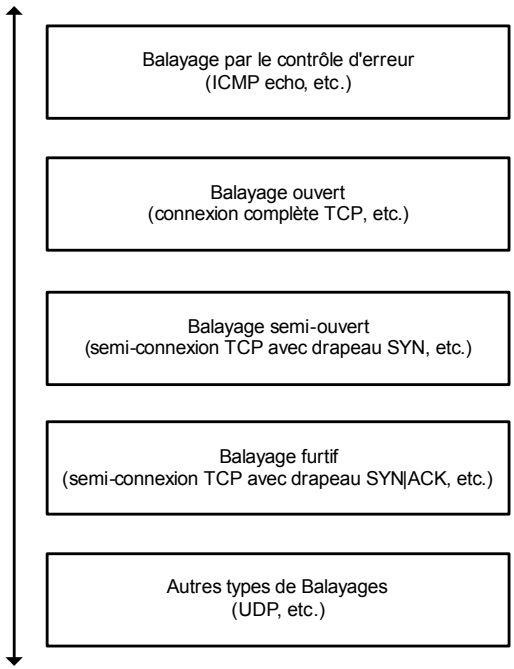
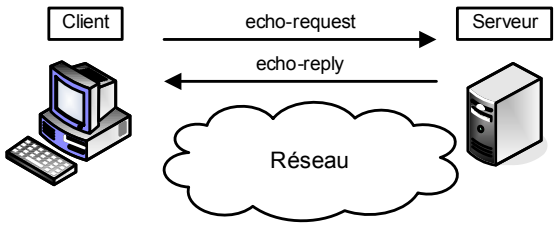


Figure 1.8

Fonctionnement de la commande ping



Il existe deux méthodes pour cartographier le réseau par cette technique :

- En balayant (scanning) le réseau et en interrogeant chaque adresse IP possible, ce qui n'est pas très discret.
- En visant une seule fois l'adresse de broadcast du réseau, ce qui fait répondre toutes les machines présentes. Une seule demande permet ainsi d'engendrer l'envoi de toutes les réponses.

Cependant, du fait de l'accroissement constant de l'insécurité, nombre d'administrateurs de pare-feu ont pris l'initiative de ne pas laisser passer les réponses à de telles demandes.

Projection dans un monde IPv6

Le protocole ICMP, qui était confiné à la remontée d'erreurs, de tests et de découverte en IPV4, est fortement étendu en IPv6, puisqu'il intègre notamment les fonctionnalités suivantes :

- découverte des voisins ;
- découverte des routeurs ;
- remontée des erreurs et ping ;
- découverte des préfixes utilisés sur le réseau (pour s'autoconfigurer) et des adresses dupliquées ;
- identification des groupes multicast en intégrant les fonctionnalités du protocole d'accès multicast.

ICMPv6 embarque ainsi des fonctionnalités plus riches, mais offre en contrepartie des possibilités plus avancées pour mener des attaques.

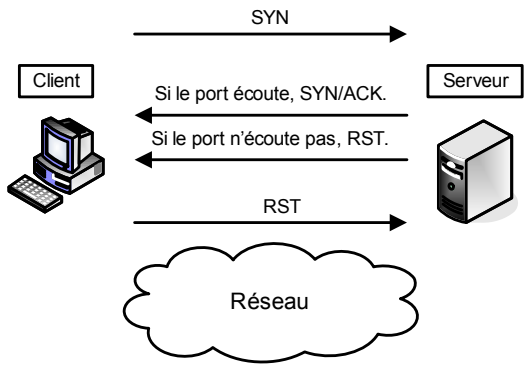
Attaque par balayage TCP

C'est en partant du principe que le flux réseau toujours accessible au pirate est celui qui est destiné à être accessible au public que la technique du balayage TCP a été inventée.

Similaire au balayage ICMP, sa spécificité est de s'appuyer sur le protocole TCP. Le client envoie un paquet SYN vers un port réseau particulier de l'adresse IP du serveur. Si le port est en écoute, un paquet SYN/ACK est reçu en retour. Sinon, la réception d'un paquet RST signifie qu'il n'y a pas de service en écoute sur le port. Le client envoie en réponse un paquet RST pour terminer la connexion, comme l'illustre la figure 1.9.

Figure 1.9

Le balayage TCP



Si aucune réponse n'est reçue en retour, c'est qu'il existe un équipement filtrant entre le serveur et le client ou qu'il n'y a aucune machine derrière l'adresse IP visée.

Cette technique est cependant si peu discrète, que des variantes ont été élaborées pour améliorer le balayage en jouant sur le principe de fonctionnement de la pile TCP/IP.

Attaque par identification des routeurs

Certaines techniques permettent de découvrir plus particulièrement les équipements assurant des fonctions de routage. L'écoute d'un réseau, par exemple, peut permettre d'analyser les trames échangées, de capturer les mises à jour des tables de routage et d'identifier les routeurs participant au routage du réseau.

Il est également possible de lancer des requêtes spécifiques afin de forcer ces mêmes routeurs à répondre. Par exemple, des requêtes peuvent s'appuyer sur une demande ICMP de découverte de routeur (ICMP router discovery) ou des requêtes de routage (OSPF, BGP, etc.).

Un pirate peut aussi envoyer des requêtes IRDP (ICMP Router Discovery Protocol), également appelées sollicitations de routeur (router solicitations), vers l'adresse de broadcast afin de connaître la route par défaut du réseau.

Attaque par traversée des équipements filtrants

Lorsqu'un pirate désire établir la cartographie d'un réseau, il rencontre généralement sur son chemin un équipement filtrant. Celui-ci peut être un routeur avec des règles de filtrage ou un pare-feu.

Dans les deux cas, des techniques permettent de traverser les filtres de cet équipement, par l'exploitation d'un bogue, par exemple, ou d'une faiblesse de configuration.

Attaque par modification du port source

Lorsqu'un pare-feu n'est qu'un simple routeur utilisant des listes de contrôle d'accès (ACL) ou un pare-feu qui ne peut détecter qu'un flux correspond au trafic retour d'une session sortante déjà initiée (le pare-feu est alors dit « stateful »), il est possible de passer outre les règles de filtrage appliquées en usurpant (spoofing) le port source du paquet émis (source porting).

Comme l'illustre la figure 1.10, le pare-feu a pour mission d'autoriser les flux sortants pour n'importe quel port source TCP associé au serveur SMTP situé sur le réseau de l'entreprise, à condition que ces flux visent n'importe quelle machine sur Internet sur le port destination 25/TCP (le port utilisé par le service SMTP). Il s'agit d'une règle typique pour le trafic SMTP permettant aux serveurs de messagerie d'envoyer des messages électroniques vers l'extérieur. Un pirate peut donc accéder aux ports TCP du serveur SMTP situé dans le réseau de l'entreprise en attaquant avec le port source 25/TCP. Il peut atteindre, par exemple, le port Telnet (23/TCP) du serveur distant.

Ce type d'attaque est rendu possible par l'absence de contrôle par l'équipement filtrant d'un ensemble de caractéristiques associées au paquet IP. Aucune vérification des bits SYN et ACK n'étant effectuée, le fait qu'un paquet SYN sans ACK arrive depuis Internet ne perturbe pas l'équipement filtrant, qui est pourtant configuré pour n'accepter que les retours de sessions sortantes. Il n'y a pas non plus de maintien dynamique des tables de trafic ayant transité par l'équipement filtrant. Celui-ci ne fait donc pas la différence entre une réponse à un trafic sortant et un trafic entrant initié de l'extérieur.

Si l'équipement filtrant appliquait ces contrôles, il ne serait pas vulnérable à ce type d'attaque.

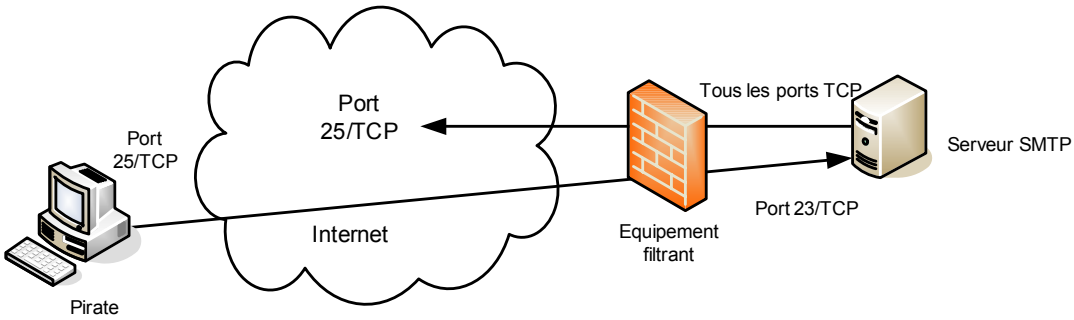


Figure 1.10

Traversée d'un pare-feu en fixant le port source

Attaque par fragmentation des paquets IP

Deux techniques permettent de jouer sur la fragmentation des paquets : celle dite par Tiny Fragments et celle par Fragment Overlapping.

Attaque par Tiny Fragments

L'attaque par Tiny Fragments consiste à fragmenter sur deux paquets IP une demande de connexion TCP ou d'autres demandes sur une machine cible tout en traversant et en déjouant (par le mécanisme de fragmentation) un filtrage IP.

Le premier paquet IP contient des données telles que les huit premiers octets de l'en-tête TCP, c'est-à-dire les ports source et destination et le numéro de séquence. Le second paquet contient la demande de connexion TCP effective (flag SYN à 1 et flag ACK à 0).

Les premiers filtres IP appliquaient la même règle de filtrage à tous les fragments d'un paquet. Le premier fragment n'indiquant aucune demande de connexion explicite, le filtrage le laissait passer, de même que tous les fragments associés, sans davantage de contrôle sur les autres fragments. Lors de la défragmentation au niveau IP de la machine cible, le paquet de demande de connexion était reconstitué et passé à la couche TCP. La connexion s'établissait alors malgré le filtre IP, comme l'illustre la figure 1.11.

Sur la figure, la demande de connexion est fragmentée en deux paquets IP contenant les fragments 0 et 1, chacun d'eux passant le système de filtrage et étant à nouveau assemblé par le système cible reconstituant la demande de connexion TCP.

Les filtres IP actuels prennent en compte les paquets fragmentés et inspectent tous les fragments de la même manière afin de se prémunir de ce type d'attaque.

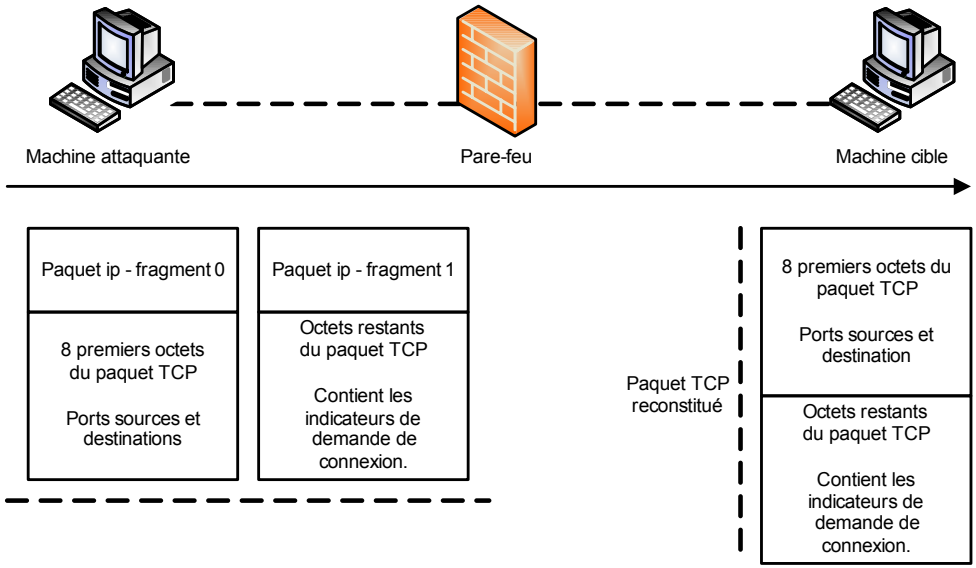


Figure 1.11

L'attaque par Tiny Fragments

Attaque par Fragment Overlapping

L'attaque par Fragment Overlapping consiste à fragmenter deux paquets IP au moyen de l'option Overlapping pour faire une demande de connexion TCP ou une autre demande sur une machine cible tout en traversant un filtrage IP.

Le premier paquet IP contient les données de l'en-tête TCP avec les indicateurs à 0. Le second paquet contient les données de l'en-tête TCP avec la demande de connexion TCP (flag SYN à 1 et flag ACK à 0).

La figure 1.12 illustre cette attaque.

Sur la figure, la demande de connexion est fragmentée en deux paquets IP contenant les fragments 0 et 1, chacun d'eux passant le système de filtrage et étant à nouveau assemblé par le système cible reconstituant un mauvais paquet TCP dû au chevauchement (overlapping) des fragments 0 et 1.

Projection dans un monde IPv6

En IPv6, seule la machine émettrice peut fragmenter les paquets, permettant par conséquence aux sauts IP de ne pas avoir à analyser, et donc de ne pas être vulnérables aux attaques par fragmentation. Cependant, il reste à la charge du destinataire du trafic de se prémunir contre ses attaques (il doit défragmenter les paquets).

Pour protéger les destinataires IPv6 de ce type d'attaque, les équipements de sécurité de protection du périmètre doivent être en mesure de contrôler la fragmentation des paquets en transit.

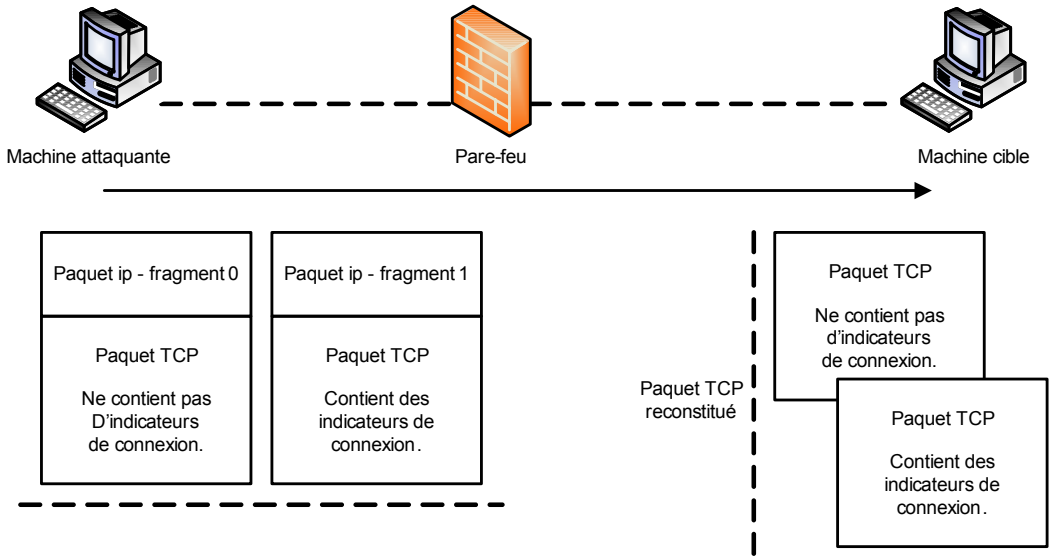


Figure 1.12

L'attaque par Fragment Overlapping

Attaques permettant d'écouter le trafic réseau

Cette technique est généralement utilisée par les pirates pour capturer les mots de passe. Lorsqu'on se connecte à un réseau qui utilise le mode broadcast, toutes les données en transit arrivent à toutes les cartes réseau connectées à ce réseau. En temps normal, seules les trames destinées à la machine sont lues, les autres étant ignorées.

Attaque par sniffing

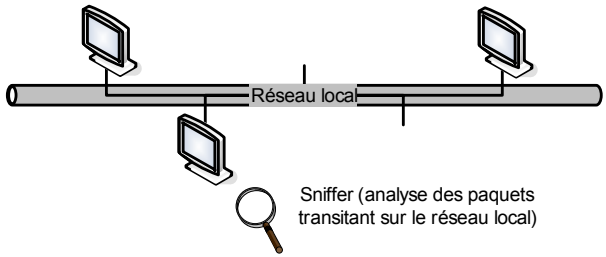
Grâce à une table d'écoute (sniffer), il est possible d'intercepter les trames reçues par la carte réseau d'un système pirate et qui ne lui sont pas destinées, comme l'illustre la figure 1.13.

Le système pirate se situe sur le réseau local et capture tous les paquets réseau transitant sur ce réseau afin d'obtenir des mots de passe, etc. Il n'est pas nécessaire que le sniffer possède une adresse IP sur le réseau qu'il écoute. Une interface réseau active sans adresse IP suffit. L'écoute est alors totalement indécelable au niveau ARP.

Grâce à des outils tels qu'Ethereal ou WinDump/TCPDump, le sniffer peut analyser tous les paquets IP ainsi que les protocoles contenus dans les données du paquet. Par exemple, un sniffer peut analyser un paquet Ethernet susceptible de contenir un paquet IP, qui lui-même pourrait contenir un paquet de type TCP, lequel à son tour pourrait contenir un paquet HTTP renfermant des données HTML.

Figure 1.13

Écoute sur un réseau local



Si une personne établit une session authentifiée sur un flux réseau non chiffré (Telnet, X11, etc.), son mot de passe transite en clair sur le réseau. De même, il est possible de connaître à tout moment les personnes connectées au réseau, les sessions de routage en cours, etc., par une analyse des paquets qui transitent sur le réseau et qui contiennent toutes les informations nécessaires à cette analyse.

Dans un réseau commuté, il n'est théoriquement pas possible d'écouter le réseau, car le commutateur envoie à chaque machine uniquement les paquets de données qui lui sont destinés. Mais comme tout équipement réseau, les commutateurs ont leurs faiblesses. Ainsi, un client qui enverrait des paquets usurpant l'adresse MAC du serveur qu'il désire écouter pourrait recevoir ces données. Selon les marques et les modèles de commutateur, le comportement diffère totalement. Cela échoue souvent, mais il arrive que cela marche. Dans certains cas, le commutateur panique et se place en déni de service.

Attaque de commutateur

Le commutateur (switch) a pour fonction de permettre la cohabitation de différents sous-réseaux physiques, qui ne communiquent pas nécessairement entre eux, sur le même équipement.

Pour atteindre cet objectif, le principe du VLAN (Virtual LAN) a été développé. À la base, un port du commutateur est assigné à un VLAN particulier, et seuls les ports du même VLAN peuvent s'échanger de l'information. Dans le but d'améliorer le confort pour l'administrateur et la qualité de service (redondance, etc.), des fonctionnalités supplémentaires ont vu le jour, avec leurs faiblesses. Ainsi, une attaque ARP spoofing peut permettre à une machine de recevoir des données qu'elle n'est pas censée recevoir.

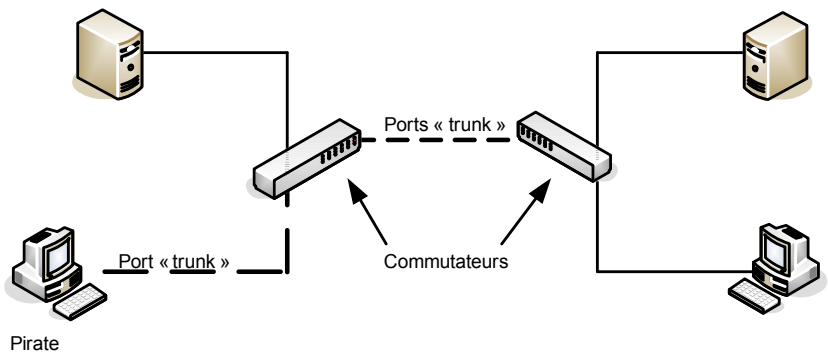
Le protocole IEEE 802.1q a pour fonction principale de permettre à des commutateurs de s'échanger des données entre des VLAN partagés par plusieurs commutateurs. Certaines faiblesses de ce protocole sont cependant exploitables par quiconque est susceptible d'initier et de générer du trafic 802.1q avec le commutateur (ce qui constitue techniquement une faiblesse de configuration).

Par exemple, la technique dite du saut de VLAN (VLAN hopping) consiste pour le pirate à envoyer vers son port des paquets 802.1q ou ISL (Inter Switch Link) afin qu'il devienne un port « trunk », port utilisé par les commutateurs pour partager des VLAN. C'est ce qu'illustre la figure 1.14.

Si l'attaque réussit, le port par lequel le pirate est attaché au commutateur devient un port « trunk ». À ce titre, il reçoit une copie de tous les paquets en transit sur tous les VLAN du commutateur.

Figure 1.14

L'attaque VLAN Hopping



Attaques permettant d'utiliser des accès distants Wi-Fi

La technologie sans fil Wi-Fi (IEEE 802.11) s'appuie sur les ondes hertziennes pour établir les communications entre les équipements. Il suffit de se trouver dans la zone de couverture des émetteurs pour écouter les données. Compte tenu du risque intrinsèque d'une telle méthode de communication, des protocoles ont été développés afin de pallier cette insécurité. Ainsi, le protocole WEP (Wired Equivalent Privacy) est censé améliorer la confidentialité des flux réseau échangés.

WEP est un protocole de sécurité défini dans le standard IEEE 802.11b. Il est chargé d'assurer un niveau de sécurité équivalent à celui des réseaux filaires en chiffrant les données transitant sur les ondes radio afin de réduire le risque d'écoute.

WEP chiffre chaque trame 802.11 échangée entre l'émetteur et le récepteur (point d'accès ou client) en s'appuyant sur l'algorithme de chiffrement symétrique en continu RC4 et sur un secret partagé entre les deux parties pour générer une clé de chiffrement en continu. Pour construire la clé de chiffrement, WEP calcule une « graine » (*seed*) correspondant à la concaténation de la clé secrète fournie par l'émetteur et d'un vecteur d'initialisation, ou IV (Initialization Vector), généré aléatoirement sur 24 bits.

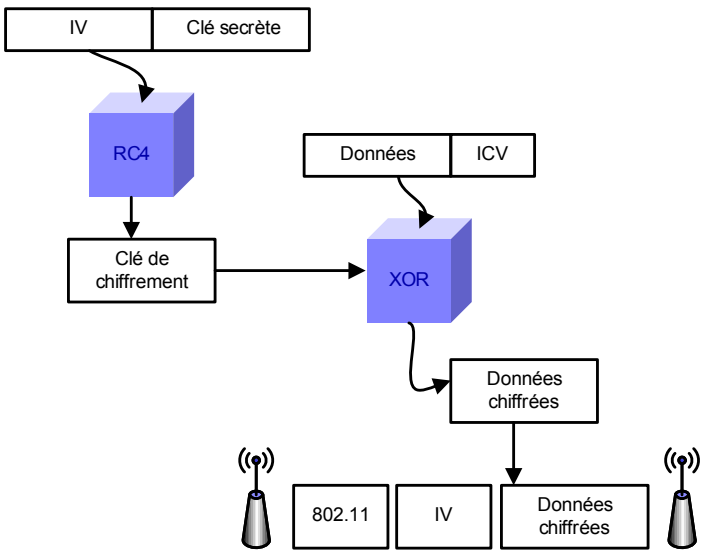
Un calcul d'intégrité, utilisant un algorithme CRC 32 et appelé ICV (Integrity Check Value), est également effectué sur les données et concaténé avec celles-ci.

La graine est ensuite utilisée par l'algorithme RC4 pour générer en continu une clé de chiffrement aléatoire. Le chiffrement des données se fait alors par un XOR (OU exclusif logique) bit à bit entre cette clé de chiffrement et les données concaténées avec l'ICV, formant en sortie les données chiffrées.

La figure 1.15 illustre ce processus de chiffrement WEP.

La clé secrète partagée peut être d'une longueur de 40 ou 64 bits, certaines versions offrant même des clés allant jusqu'à 128 bits (104 bits réels).

Figure 1.15
Chiffrement WEP



Attaque FMS (Fluhrer, Mantin, Shamir) sur RC4

RC4 est connu depuis des années pour être vulnérable à des attaques de type « texte déchiffré connu » *(known plain text attack)*. Ce type d'attaque consiste à deviner la clé secrète en s'appuyant sur la connaissance de tout ou partie des données de la version déchiffrée. La technique d'attaque FMS a démontré qu'il fallait environ 1 000 000 paquets pour casser une clé de 128 bits et 300 000 pour une clé de 64 bits.

La figure 1.16 illustre le processus d'attaque de la clé secrète lors d'une demande de connexion à un point d'accès sur lequel WEP n'est pas activé.

Le pirate a enregistré l'échange challenge/réponse de l'utilisateur, et il sait que la réponse contiendra la version chiffrée avec la clé secrète. Il connaît également ce challenge puisqu'il a transité en clair lors de l'établissement de la session de l'utilisateur. Le pirate peut donc se procurer la clé secrète en pratiquant une attaque de type « texte déchiffré connu » sur la réponse de l'utilisateur. Une fois le challenge obtenu dans le message de réponse, la clé secrète est trouvée.

Il existe d'autres méthodes pour casser une clé WEP ou permettre de déchiffrer un paquet chiffré avec une clé WEP sans avoir connaissance de la clé (vulnérabilité du contrôle de conformité).

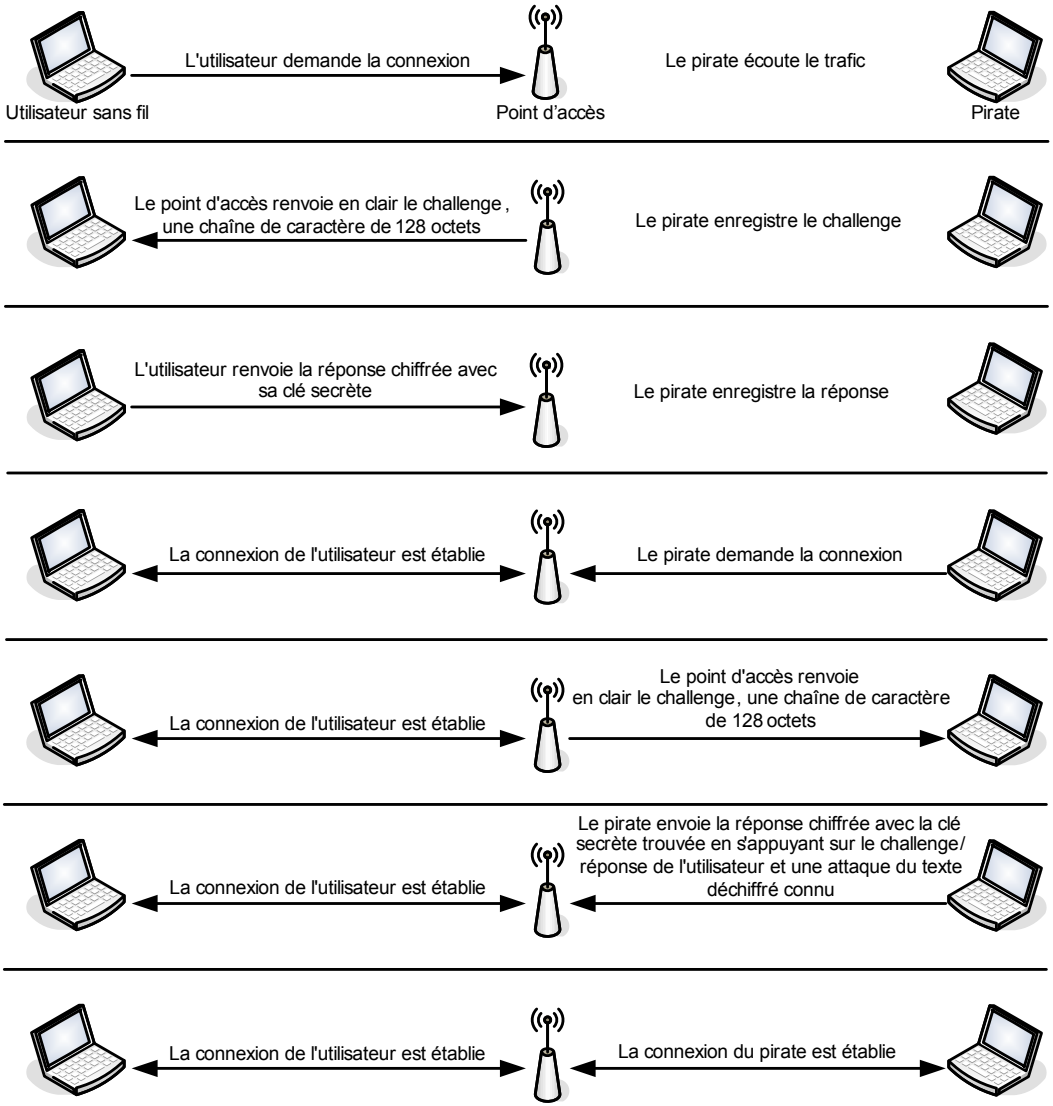


Figure 1.16

Attaque de la clé secrète sans utilisation de WEP

Attaque par modification de paquet

WEP utilise un checksum pour s'assurer de l'intégrité d'un paquet. Cependant, WEP utilisant une fonction linéaire pour calculer ce checksum, il est possible de modifier le contenu d'un paquet (et de son checksum) sans aucune détection possible de la part du récepteur.

Cette attaque est également connue sous le nom de Bit Flipping Attack, une variante consistant simplement à déplacer les bits.

Attaque par envoi de paquet ou par répétition

Nous avons vu précédemment qu'une partie de la clé secrète reposait sur le vecteur d'initialisation généré aléatoirement. Il est cependant possible de réutiliser un vecteur d'initialisation, sans que cela soit considéré comme un comportement anormal.

Grâce à cette particularité, il est possible pour un pirate d'envoyer des paquets avec un vieux vecteur d'initialisation, considéré comme obsolète, dans la communication entre un client et un point d'accès en espérant qu'il soit de nouveau utilisé par cette communication (replay attack).

Attaque par redirection d'adresse IP

Cette attaque nécessite que le point d'accès permette l'accès au réseau Internet, ce qui est fréquemment le cas. Elle suppose en outre que le pirate contrôle un ordinateur sur Internet.

La séquence des événements est la suivante :

1. Le pirate modifie l'intégrité d'un paquet en remplaçant l'adresse IP destination par l'adresse de l'équipement qu'il contrôle. Il s'appuie pour cela sur un paquet capturé et la méthode dite du bit flipping.
2. Il garde une copie du paquet chiffré.
3. Le paquet est déchiffré par le point d'accès puis envoyé en clair sur le réseau vers l'adresse IP destination (donc l'ordinateur sous contrôle du pirate), laquelle reçoit la version en clair du paquet de données.
4. Le pirate récupère cette version en clair.

Le pirate possédant la version chiffrée et déchiffrée du paquet, il peut commencer une attaque de type « texte déchiffré connu » pour trouver la clé WEP.

Attaques permettant d'interférer avec une session réseau

La plupart des protocoles réseau n'ayant prévu aucun mécanisme d'authentification véritable, ils subissent des attaques qui s'appuient sur ces faiblesses d'authentification, au premier rang desquelles les attaques ARP spoofing et man-in-the-middle.

Attaque ARP spoofing

Comme son nom l'indique, l'attaque ARP spoofing s'appuie sur le protocole ARP (Address Resolution Protocol), qui implémente le mécanisme de résolution d'une adresse IP (32 bits) en une adresse MAC (48 bits) pour rediriger le trafic réseau de un ou plusieurs systèmes vers le système pirate.

Lorsqu'un système désire communiquer avec ses voisins sur un même réseau (incluant la passerelle d'accès à d'autres réseaux), des messages ARP sont envoyés afin de connaître l'adresse MAC des systèmes voisins et d'établir ainsi une communication avec un système donné.

Sachant que chaque système possède localement une table de correspondance entre les adresses IP et MAC des systèmes voisins, la faiblesse d'authentification du protocole ARP permet à un système pirate d'envoyer des paquets ARP réponse au système cible indiquant que la nouvelle adresse MAC correspondant à l'adresse IP d'une passerelle est la sienne.

Le système du pirate reçoit donc tout le trafic à destination de la passerelle. Il lui suffit d'écouter ou de modifier passivement le trafic et de router ensuite les paquets vers leur véritable destination, comme l'illustre la figure 1.17.

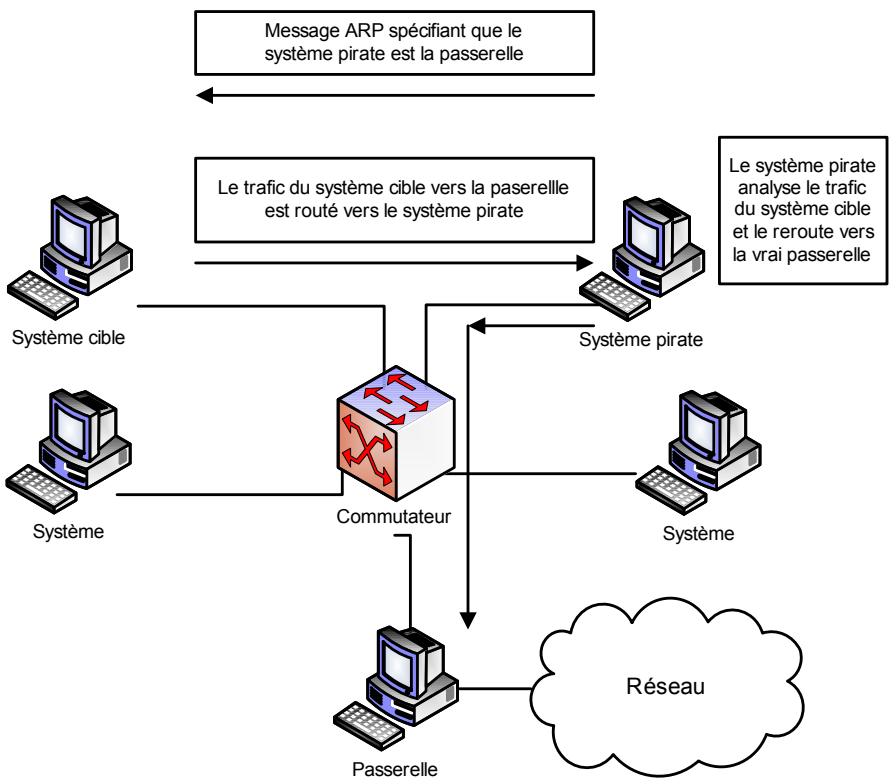


Figure 1.17

L'attaque ARP spoofing

Projection dans un monde IPv6

Bien que le protocole ARP ait complètement disparu dans la migration vers IPv6, le principe reste le même. Là où ARP faisait ses annonces en broadcast, ICMPv6 les fait en multicast.

Sachant qu'un système IPv6 possède une table de ses voisins (Neighbor Cache) qui est mise à jour par le biais de messages ICMPv6 de type NS (Neighbor Sollicitation) et NA (Neighbor Advertisement), les attaques visant ARP sont toujours possibles en IPv6.

Il est toujours possible d'envoyer de faux messages NS et NA pour corrompre les tables de voisinage d'un système IPv6 et ainsi attirer le trafic vers son adresse MAC.

Attaque IP spoofing

Puisqu'un paquet IP n'est qu'une suite d'octets construite par un système d'exploitation s'exécutant sur un système hardware, cette suite d'octets peut être forgée et envoyée sur le réseau sans contrôle préalable de ce dernier.

La plupart des moyens d'authentification s'appuyant de nos jours sur les adresses IP, ce moyen faible d'authentification peut entraîner de graves problèmes de sécurité si l'authentification ne recourt qu'à ce mécanisme. Si un système peut donner des privilèges particuliers à un ensemble d'adresses IP sources, un paquet IP forgé avec une telle adresse IP est reçu par ce système avec les privilèges associés.

L'attaque IP spoofing consiste à se faire passer pour un autre système en falsifiant son adresse IP. Le pirate commence par choisir le système qu'il veut attaquer. Après avoir obtenu le maximum de détails sur ce système cible, il détermine les systèmes ou adresses IP autorisés à se connecter au système cible. Le pirate procède ensuite aux étapes illustrées à la figure 1.18 pour mener à bien son attaque sur le serveur cible en utilisant l'adresse IP de la machine A.

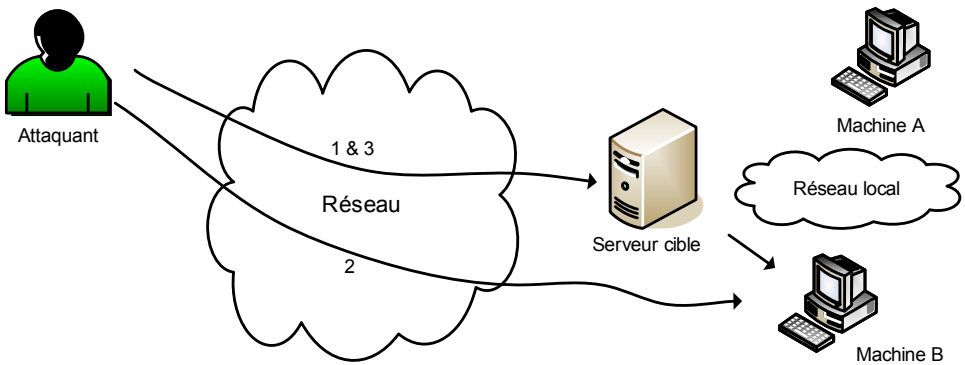


Figure 1.18

L'attaque IP spoofing

L'attaque se déroule de la façon suivante :

1. Le pirate essaye de prévoir le numéro de séquence des paquets du serveur cible en envoyant plusieurs paquets et en analysant l'algorithme d'incrémentation de ce numéro.
2. Le pirate rend inopérante la machine A autorisée à accéder au serveur cible, de façon à s'assurer qu'elle ne répond pas au serveur cible.
3. Le pirate falsifie son adresse IP en la remplaçant par celle de la machine invalidée et envoie une demande de connexion au serveur cible.
4. Le serveur envoie une trame SYN|ACK à la machine qu'il pense être l'émettrice.
5. Celle-ci ne pouvant répondre, le pirate acquitte cette connexion par une trame ACK, avec le numéro de séquence prévu. Il établit de la sorte en toute impunité la connexion avec le serveur cible.

Cette attaque est assez difficile à effectuer, car elle se réalise en aveugle, le pirate ne recevant pas les données transmises par le serveur. Il doit donc maîtriser parfaitement les protocoles pour savoir ce qu'attend le serveur à tout moment. D'autres techniques plus évoluées permettent de contourner ce problème, comme les attaques dites man-in-the-middle (l'homme au milieu) ou les attaques de routage.

Attaque man-in-the-middle

L'attaque man-in-the-middle consiste à faire passer les échanges réseau entre deux systèmes par le biais d'un troisième, sous le contrôle du pirate. Ce dernier peut transformer à sa guise les données à la volée, tout en masquant à chaque acteur de l'échange la réalité de son interlocuteur.

Pour mettre en œuvre l'échange réseau approprié, il faut soit que la machine du pirate se trouve physiquement sur le chemin réseau emprunté par les flux de données, soit que le pirate réussisse à modifier le chemin réseau afin que sa machine devienne un des points de passage (nous détaillons plus loin ce type d'attaque).

Au final, l'échange se présente sous l'une des trois formes suivantes :

- Relais transparent. La machine du pirate transforme les données à la volée. Elle veut rester la plus transparente possible et se comporte comme un routeur, conservant toutes les caractéristiques des paquets dont elle assure le transit, à l'exception du contenu. En termes d'adresses IP, A et B sont réellement en relation l'une avec l'autre *(voir figure 1.19).*

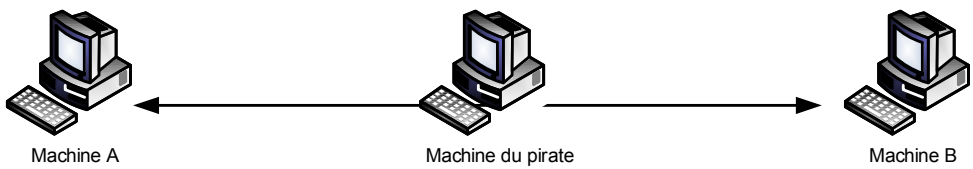


Figure 1.19

Machine du pirate en tant que relais transparent

- Relais applicatif. La machine du pirate assure l'échange entre les deux machines A et B. A parle avec la machine du pirate, laquelle parle avec B. A et B n'échangent jamais de données directement. Cette méthode est nécessaire pour les attaques vers SSL, par exemple *(voir figure 1.20)*.

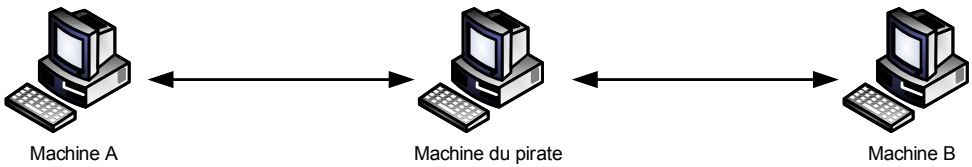


Figure 1.20

Machine du pirate en tant que relais applicatif

- Hijacking. La machine du pirate utilise la session engagée entre les deux machines A et B afin que ce soit elle (la machine du pirate) qui soit en session avec la machine B. A perd la session avec B, et la machine du pirate continue la session engagée par A sur B *(voir figure 1.21)*.

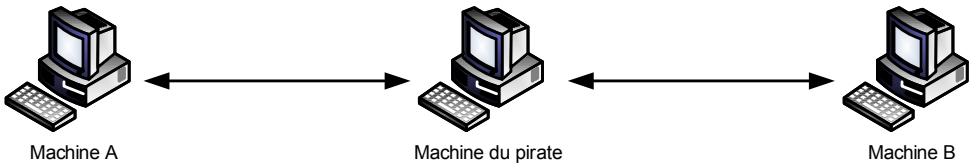


Figure 1.21

Machine du pirate en tant que hijacker

Le détournement (hijacking) des sessions TCP permet de rediriger un flux TCP en outrepassant les authentifications nécessaires à l'établissement des sessions (Telnet, FTP, etc.). Cette attaque porte de manière plus spécifique sur l'analyse des numéros de séquences et des numéros d'acquittements relatifs aux paquets TCP.

La première étape consiste à écouter le trafic réseau entre deux systèmes et à analyser les numéros de séquences et d'acquittements, ainsi que les indicateurs TCP, à l'aide d'un sniffer tel que tcpdump, par exemple.

Les traces fournies par tcpdump sont de la forme :

```
src > dst : flags data-seqno ack window urgent
```

- src et dst sont les adresses IP source et destination avec les ports associés.
- flags est la combinaison des indicateurs TCP S (SYN), F (FIN), P (PUSH), etc.
- data-seqno est constitué de numéros de séquences séparés par le caractère « : ». Les numéros de séquences sont utilisés par TCP pour ordonner les données reçues.

- ack est le numéro de séquence destiné à informer l'expéditeur de la bonne réception des données.
- window est la taille du tampon TCP de réception.
- urgent indique si le drapeau URGENT est positionné.

Les traces tcpdump relatives à l'établissement d'une connexion rlogin entre le système A (système_a) et le système B (système_b) sont de la forme suivante :

```
système_a.1023 > système_b.login : S 768512 :768512(0) win 4096
système_b.login > système_a.1023 : S 947648 :947648(0) ack 768513 win 4096
système_a.1023 > système_b.login : . ack 1 win 4096
système_a.1023 > système_b.login : P 1:2(1) ack 1 win 4096
système_b.login > système_a.1023 : . ack 2 win 4096
système_a.1023 > système_b.login : P 2:21(19) ack 1 win 4096
système_b.login > système_a.1023 : P 1:2(1) ack 21 win 4077
système_b.login > système_a.1023 : P 2:3(1) ack 21 win 4077 urg 1
système_b.login > système_a.1023: P 3:4(1) ack 21 win 4077 urg 1
```

La première ligne indique que système_a initié l'envoi de paquets TCP à partir du port 1023 vers système_b sur le port du rlogin. Le *S* indique que l'indicateur TCP *SYN* est positionné et que le numéro de séquence est égal à 768512 et qu'il n'y a pas de données.

système_b répond par un SYN + ACK (ligne 2), et système_a renvoie un ACK pour confirmer la connexion (ligne 3). Les autres lignes montrent les échanges de messages entre système_a et système_b avec l'émission de données.

L'attaque par hijacking d'une session TCP crée un état de désynchronisation de chaque côté de la connexion TCP, permettant le vol de session par un pirate.

Une connexion est désynchronisée lorsque le numéro de séquence du prochain octet envoyé par système_a est différent du numéro de séquence du prochain octet à recevoir par système_b. Réciproquement, il y a désynchronisation lorsque le numéro de séquence du prochain octet envoyé par la machine système_b est différent du numéro de séquence du prochain octet à recevoir par système_a.

Concrètement, lorsqu'un pirate avec une machine système_c veut voler une session Telnet établie entre système_a et système_b, il procède de la façon suivante *(voir figure 1.22) :*

- Le pirate (système_c) sniffe le trafic Telnet (port TCP 23) entre système_a et système_b.
- Une fois qu'il estime que système_b s'est s'authentifié auprès du service Telnet de la machine système_b, il désynchronise la machine système_a par rapport à système_b en forgeant un paquet avec comme adresse IP source celle de système_a et comme numéro d'acquittement TCP celui attendu par système_b.
- système_b accepte ce paquet et permet au pirate de s'insérer dans la session préalablement établie par système_a.

Si système_a envoie un paquet à système_b, celui-ci n'est pas accepté du fait que le numéro de séquence n'est pas celui attendu par système_b. Cette attaque peut alors engendrer une série d'envois de paquets ACK entre système_a et système_b, qui les refusent tous deux du fait de la désynchronisation du numéro de séquence.

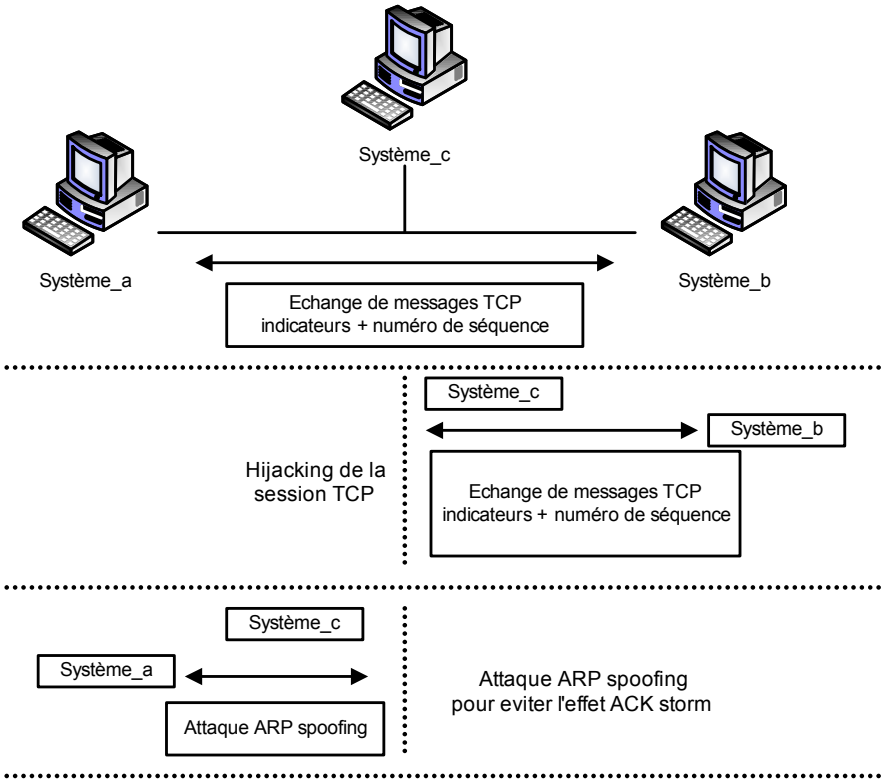


Figure 1.22

Hijacking d'une session TCP

Pour pallier ce problème dit du ACK storm, système_c peut utiliser l'attaque ARP spoofing vers système_a pour lui affirmer que l'adresse IP de système_b correspond à l'adresse MAC de système_c.

Attaque man-in-the-middle par modification du routage

Une des méthodes permettant à un pirate de se placer dans la configuration de l'homme au milieu repose sur la modification du routage.

Par diverses méthodes, selon le protocole de routage visé, le pirate peut influencer le comportement du réseau afin que les flux de celui-ci transitent par son ordinateur.

Modification par routage à la source

Le routage à la source vise à forcer un routage particulier pour un échange de données. Son principe de fonctionnement est simple : les paquets sont envoyés avec le chemin qu'ils doivent emprunter.

Prenons l'exemple illustré à la figure 1.23 :

- Une machine A échange des données avec une machine B. Normalement, cet échange de flux se fait par le biais des routeurs 1 et 2 (ligne en pointillés).
- L'agresseur envoie ses paquets vers la machine B en usurpant l'adresse IP source de la machine A et en utilisant un routage à la source.
- La machine B reçoit les données et renvoie les réponses *via* le chemin précisé dans le routage à la source.
- La machine de l'agresseur reçoit les données comme les aurait reçues la véritable machine A.

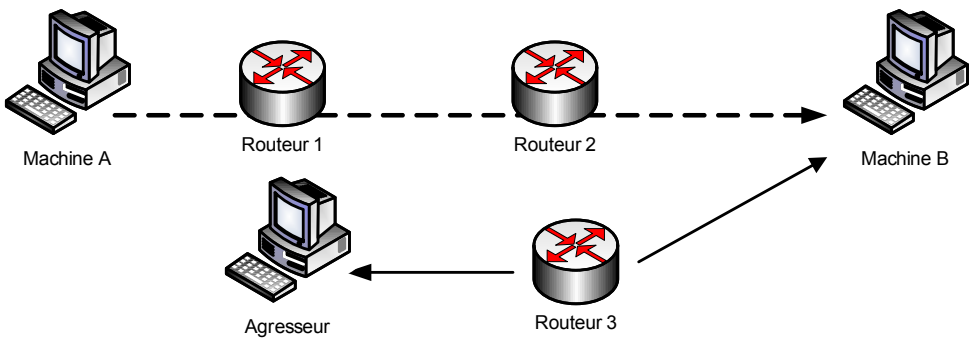


Figure 1.23

Attaque par routage à la source

Modification par ICMP redirect

Une variante de l'attaque précédente consiste à utiliser le type redirect du protocole ICMP.

Le principe de fonctionnement de cette attaque est le suivant *(voir figure 1.24) :*

- Une machine A échange des données avec une machine B.
- Normalement, cet échange de flux s'effectue par le biais des routeurs 1 et 2 (ligne en pointillés).
- L'agresseur s'installe entre les routeurs 3 et 4.
- Il convainc le routeur 1 que le meilleur chemin consiste à passer par le routeur 3 en lui envoyant des paquets ICMP redirect.
- Le routeur 1 envoie les paquets destinés à la machine B *via* le routeur 3.
- L'agresseur est placé en goulet d'étranglement entre les routeurs 3 et 4.

Dans des variantes de cette attaque, la machine de l'agresseur assure elle-même la fonction de routage à la place du routeur 3 ou fait croire à la machine B que sa machine est le meilleur chemin.

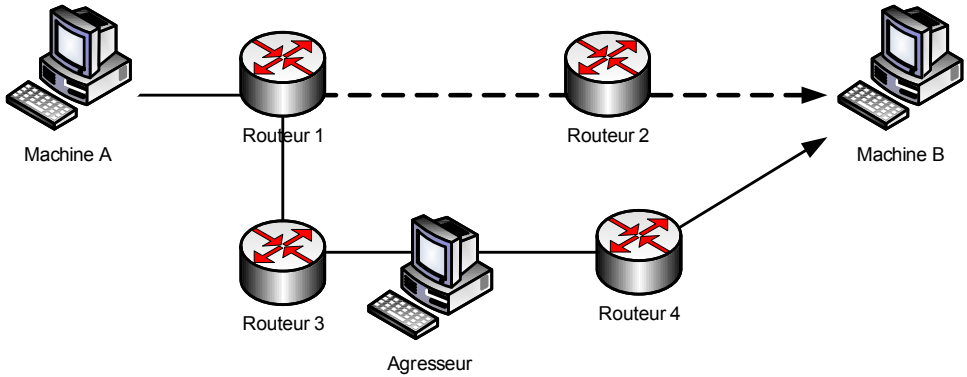


Figure 1.24

Attaque par ICMP redirect

Attaque man-in-the-middle sur le chiffrement SSL

Dans les attaques sur le chiffrement SSL, le trafic vers le port SSL (HTTPS sur le port 443 TCP) est dérouté de manière transparente vers la machine du pirate, cette dernière assurant les fonctions d'un serveur HTTPS (HyperText Transfer Protocol Secure sockets).

Le principe de fonctionnement de cette attaque est le suivant *(voir figure 1.25)* :

- La machine client A se connecte au serveur pirate.
- Ce dernier lui fournit un certificat apparemment digne de confiance (surtout lorsqu'il est accepté par un navigateur ou un utilisateur trop confiant).
- Le serveur HTTPS du pirate, qui assure toutes les fonctions d'un relais applicatif, reçoit les demandes de A de manière chiffrée.
- Le serveur du pirate les déchiffre et les réachemine vers le serveur B (la machine pirate s'est connectée au service HTTPS du serveur B et simule la connexion de la machine A).

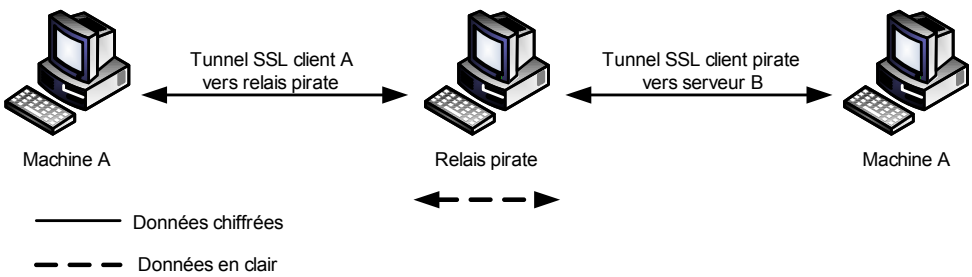


Figure 1.25

Relais pirate voyant les données SSL passer en clair

Le tunnel HTTPS semble techniquement irréprochable, sauf qu'il n'y a pas un tunnel entre A et B mais deux : un premier entre le client A et le serveur sur le relais pirate et un second entre le client relais pirate et le serveur B. Le relais pirate peut donc à loisir recopier ou modifier les données en transit malgré la sensation de chiffrement qu'a le client A.

Une telle attaque ne peut fonctionner s'il existe une complicité entre A et B, par le biais d'une clé privée, par exemple.

Attaques permettant de modifier le routage réseau

Tous les protocoles de routage ont pour objectif de maintenir les tables de routage du réseau dans un état intègre et cohérent. Pour y parvenir, les protocoles diffusent des informations de routage aux autres systèmes du réseau afin de transmettre les modifications des tables de routage. Ces protocoles réceptionnent en contrepartie les informations de routage d'autres systèmes du réseau afin de mettre à jour les tables de routage.

Dans les premiers grands réseaux, les tables de routage étaient statiques et donc maintenues à jour par des techniciens de bout en bout. De nos jours, les mises à jour des tables de routage et le calcul du meilleur chemin sont automatiquement propagés sur le réseau par les protocoles de routage.

IGP (Interior Gateway Protocol) et EGP (Exterior Gateway Protocol) sont les deux grandes familles de protocoles de routage dans les réseaux IP. Un réseau de routage est découpé généralement en systèmes autonomes, dits AS (Autonomous System). Dans un système autonome, le protocole de routage utilisé est de type IGP. Pour les échanges de routage entre systèmes autonomes différents, le protocole de routage utilisé est de type EGP.

On distingue aussi les protocoles de routage IGP et EGP en mode unicast *versus* les protocoles de routage IGP et EGP en mode multicast :

- Le mode unicast définit une connexion réseau point à point.
- Le mode multicast est utilisé pour désigner une méthode de diffusion de l'information d'un émetteur vers un groupe.

Sachant que toute attaque ou perturbation du routage peut impacter directement la disponibilité du réseau et de ses services, il est primordial de considérer les protocoles de routage comme des éléments-clés de la sécurité d'un réseau. D'autant qu'il est aussi possible de détourner du trafic par le routage à des fins de vol d'information.

Attaques sur le routage IGP

Le protocole OSPF (Open Shortest Path First) est un protocole de routage de type IGP permettant de gérer les routes internes à un AS.

Attaque du numéro de séquence maximal d'une annonce

Dans les spécifications d'OSPF v2, le champ réservé pour le numéro de séquence est un entier signé de 32 bits utilisé pour détecter les annonces vieilles ou dupliquées. Lorsqu'un

routeur reçoit un LSA (Link State Advertisement), la valeur du numéro de séquence est comparée à celle de l'annonce en cours afin de savoir lequel est le plus récent. Si les valeurs sont différentes, l'annonce avec le numéro de séquence le plus élevé est conservée.

Un routeur utilise 0x80000001 comme valeur de départ lorsqu'il envoie un LSA (la valeur 0x80000000 étant réservée), puis il incrémente le numéro de séquence d'une unité à chaque nouvelle annonce envoyée.

En théorie, le LSA avec la valeur maximale devrait être purgé du domaine de routage. Malheureusement, du fait de bogues d'implémentation, ce n'est pas le cas. Un pirate qui envoie un LSA avec un numéro de séquence maximal provoque l'ajustement du routage selon les informations fournies dans le LSA. De plus, toutes les mises à jour suivantes sont ignorées par les routeurs.

Attaque du numéro de séquence d'une annonce

Comme nous l'avons vu précédemment, selon la valeur du numéro de séquence d'une annonce, le routeur considère toujours l'annonce la plus récente.

Un pirate peut donc envoyer une annonce avec un numéro de séquence supérieur à celui de l'annonce en cours en écoutant le réseau. Le réseau ajuste alors son routage en fonction des informations fournies dans cette fausse annonce.

Attaques sur le routage EGP

BGP (Border Gateway Protocol) est un protocole de routage de type EGP permettant de gérer les routes externes d'un AS. Il s'agit du protocole utilisé sur Internet pour échanger les routes entre les différents AS constituant ce réseau.

Le protocole BGP n'a pas été conçu à l'origine de manière sécurisée. Il est donc possible, par le biais de diverses attaques, d'injecter, de modifier ou d'impacter d'une manière ou d'une autre un processus de routage.

Attaque man-in-the-middle

Entre deux AS différents, généralement plusieurs routeurs sont connectés afin d'échanger non seulement le trafic réel, mais aussi les annonces de routes. Ces routeurs constituent des cibles de choix pour des attaques visant à créer la situation de l'« homme au milieu ».

Attaque par attraction totale du routage

En cas de déni de service sur le véritable propriétaire des routes, tel qu'une entreprise, c'est l'intégralité des routes qui devient inaccessible depuis Internet, et non simplement quelques systèmes. Il s'agit alors d'une attaque de type black hole, ou trou noir.

Un des problèmes de sécurité engendrés par l'avènement de l'AS 7007 en 1997 a été qu'un routeur a publié des routes qui ne lui appartenaient pas et a fini par attirer vers lui tout le trafic Internet. La conséquence a été tout simplement un déni de service de l'ensemble du réseau Internet.

Attaque par attraction partielle du routage

Un des principes du routage réside dans la spécificité d'un préfixe annoncé. Par exemple, s'il existe dans une table de routage les deux entrées suivantes :

- 192.0.0.0/8
- 192.0.0.0/24

le trafic à destination du sous-réseau 192.0.0.0/24 est transporté par l'indication de routage relative à l'entrée de routage 192.0.0.0/24, et non par celle relative à 192.0.0.0/8.

Pour une annonce de route donnée, il est donc possible de faire converger une partie du trafic vers soi.

C'est ce qui est arrivé au Pakistan, en 2008, lorsque le gouvernement a souhaité interdire l'accès à YouTube et a redistribué sur Internet par mégarde une route plus spécifique à l'annonce de base de YouTube, attirant ainsi une grande partie du trafic mondial vers lui-même.

Attaque par attraction partielle et rebond du routage

Cette variante de l'attaque précédente permet de détourner un préfixe plus spécifique que celui attaqué, tout en redirigeant le trafic final vers la destination officielle. Elle a été présentée en 2008 à DEFCON par P. Kapela.

Basée sur l'annonce d'un préfixe plus spécifique modifiant volontairement l'AS-PATH (chemin des AS traversés) pour protéger le chemin de retour et ainsi éviter la détection de boucle par les AS intermédiaires (qui mettraient en échec l'attaque), cette attaque met en œuvre plusieurs techniques pour détourner du trafic.

Cette attaque couvre aussi la protection de l'attaquant en proposant de modifier le TTL (compteur décrémenté par chaque saut IP) des paquets transitant par ce nuage « pirate » afin de rendre invisible de l'extérieur le détournement de trafic par des outils traditionnels tel que « traceroute ».

Attaque par attraction temporaire du routage

Cette attaque consiste à annoncer des préfixes larges (de façon à avoir moins de chance de se faire filtrer par les routeurs) sur de courtes périodes, puis de monter, par exemple, des sessions smtp prenant une adresse source dans le préfixe annoncé, d'envoyer des messages de SPAM ou d'autres attaques et enfin de relâcher rapidement les préfixes annoncés.

Attaques sur le routage multicast

D'une manière générale, les attaques possibles sur un service de diffusion multicast peuvent être classifiées de la manière générique suivante :

- Selon le type. On distingue les attaques par déni de service, qui visent à engorger les capacités des réseaux ou à saturer les ressources des routeurs, et les attaques mettant à profit les vulnérabilités des protocoles par falsification de messages de signalisation multicast.

- Selon la cible. On peut distinguer les attaques par déni de service, dirigées contre le plan de transfert des routeurs, et celles dirigées contre le plan de contrôle des routeurs.
- Selon l'origine. Ces attaques peuvent provenir soit du cœur de réseau, soit de l'accès. Le cœur du réseau contient l'ensemble des routeurs multicast qui exécutent les protocoles de routage multicast. Le réseau d'accès est constitué des équipements qui exécutent les protocoles d'abonnement/désabonnement aux flux de diffusion multicast.

Voici deux types d'attaques possibles :

- Une source malveillante pourrait attaquer un arbre de distribution multicast existant en injectant un trafic parasite (des paquets quelconques dont l'adresse de destination est l'adresse multicast du groupe visé) sur le groupe de diffusion et violerait ainsi l'intégrité du flux de diffusion légitime en ajoutant ce trafic parasite à la communication de groupe existante. Ce trafic parasite, qui est reçu par tous les récepteurs du groupe, vise à perturber la diffusion existante du flux légitime.
- Un assaillant pourrait souscrire à des milliers d'adresses de groupes et à des milliers d'adresses sources. L'envoi de requêtes de souscription à un groupe par l'assaillant déclencherait de nombreux événements dans le protocole de routage multicast associé. L'énorme quantité d'entrées dans la table de routage multicast peut pénaliser les flux légaux. Cette attaque consomme aussi des ressources mémoire dans les équipements réseau gérant le trafic multicast afin de maintenir les états multicast créés et de traiter les messages de routage multicast. Elle est particulièrement dangereuse pour les équipements réseau situés aux racines (ou proches des racines) des arbres de distribution multicast puisque ce sont ceux qui ont à maintenir le plus d'états multicast.

Attaques permettant de mettre le réseau en déni de service

Le déni de service, ou DoS (Denial of Service), est une attaque qui vise à rendre indisponible un service, un système ou un réseau. Ces attaques s'appuient généralement sur une faiblesse d'implémentation, ou bogue, ou sur une faiblesse d'un protocole.

Les premières attaques par déni de service sont apparues entre 1998 et l'an 2000. Elles visaient de grands sites Internet (Yahoo, eBay, eTrade, etc.). Le site Yahoo, a été attaqué en février 2000 et a été inondé (flood) sous 1 Go de données en quelques secondes, les données provenant d'au moins cinquante points réseau différents.

Attaque par inondation

L'inondation est la méthode la plus classique pour empêcher un réseau d'assurer sa mission.

Son principe de fonctionnement est le suivant :

- Une ou plusieurs machines inondent le réseau avec des paquets réseau afin de saturer la bande passante de celui-ci.
- Une fois que toute la bande est occupée, les autres machines ne peuvent plus travailler, ce qui génère une situation de déni de service.

L'inondation peut recourir à différentes méthodes. La plus classique est l'inondation ping (Ping flooding), une machine envoyant des paquets ping ICMP request et attendant en réponse un paquet ICMP reply. Sans mention d'un délai pour l'obtention de la réponse, la machine envoie ses paquets aussi vite qu'elle le peut, saturant ainsi le réseau.

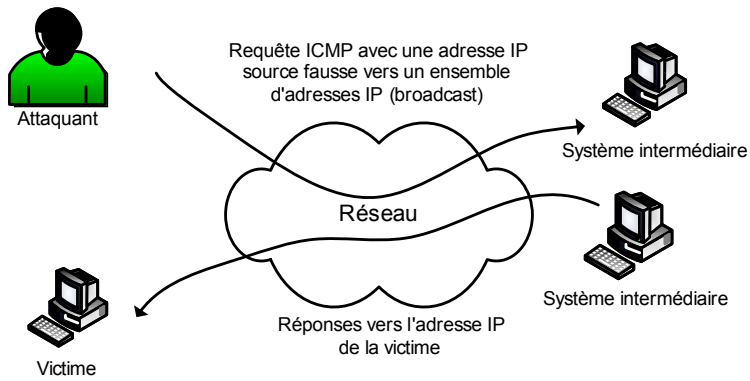
Attaques smurf et fraggle par amplification de l'inondation

Les attaques smurf et fraggle sont des variantes de la précédente qui s'appuient sur une faiblesse de configuration des routeurs.

Ces techniques consistent à inonder le réseau avec des ping qui n'utilisent que des adresses de broadcast. Pour un paquet envoyé, toutes les machines d'un réseau répondent, ce qui augmente la saturation du réseau, comme l'illustre la figure 1.26.

Figure 1.26

Attaques smurf et fraggle



Du fait de l'envoi des paquets ICMP avec une fausse adresse source vers une adresse de broadcast, chaque machine appartenant au réseau couvert par le broadcast répond aux systèmes victimes ou aux systèmes fictifs. Comme le pirate n'attend pas de trafic retour, il peut bombarder un ensemble d'adresses de broadcast et générer un trafic important par phénomène d'amplification.

La différence entre l'attaque smurf et l'attaque fraggle est que cette dernière utilise le protocole UDP.

Attaque par inondation TCP SYN

La technique d'inondation SYN est identique à celle du balayage SYN, à la différence près qu'elle est utilisée à des fins de déni de service.

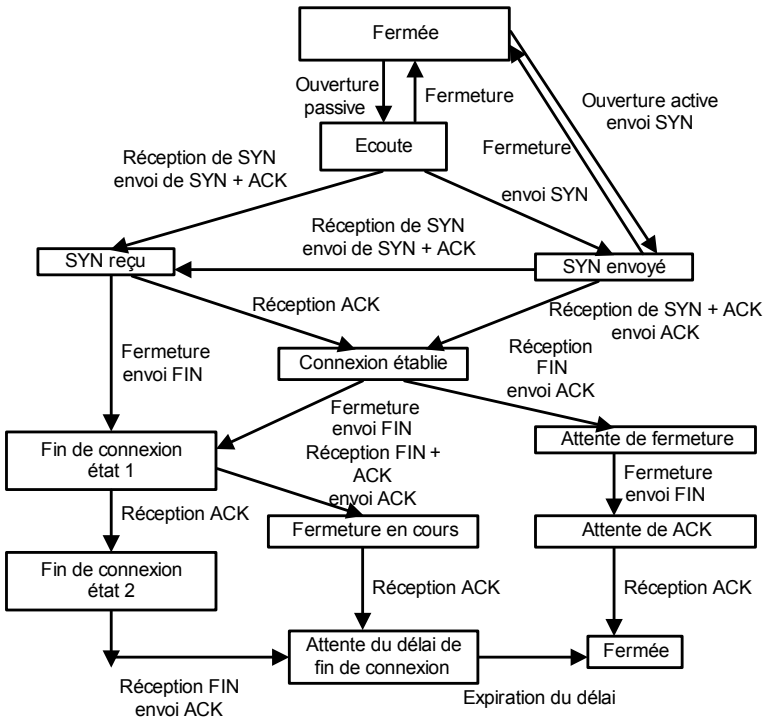
Nous avons vu que le principe du balayage semi-ouvert consistait à ce que le client ne termine pas la session TCP par l'envoi d'un paquet RST. Ainsi, le serveur reste dans un état intermédiaire, dans lequel la session n'existe pas réellement puisqu'elle est en cours d'établissement. Dans cet état, le serveur doit réserver des ressources (réseau, mémoire, CPU, etc.) pour le traitement de la session TCP et attendre la fin du handshake.

Tous les serveurs supportent un nombre maximal de sessions TCP en cours. Lorsqu'une session est terminée, les ressources associées à la session sont remises à disposition du système d'exploitation. Lorsque la session n'est pas encore établie, le système prend la peine de faire patienter les paquets manquants, estimant qu'ils sont simplement retardés par le réseau. Ce délai d'attente pour passer les différents états de libération de la session est paramétrable mais prend généralement une bonne minute.

Ces différentes étapes sont illustrées à la figure 1.27.

Figure 1.27

États d'une session TCP



L'envoi de paquets SYN par un pirate vers un serveur, une opération très rapide puisque le pirate n'attend pas de réponse de celui-ci, engendre une saturation des ressources réseau de la victime, laquelle ne peut plus dès lors assurer sa mission.

Attaque par épuisement de TCP

La technique par épuisement de TCP consiste à établir un nombre important de connexions (complètes) TCP entre deux acteurs, tout en créant un déséquilibre important d'allocation des ressources.

Ainsi, en exploitant au maximum les possibilités du protocole TCP de garder une connexion active *a minima* pour l'attaquant (utilisation du paramètre de débit permettant

de limiter les requêtes de maintien de connectivité en se plaçant dans un mode congestion), il est possible de créer un nombre important de connexions actives sur le système attaqué de telle manière que le système se sature lui-même.

La raison de la saturation générale est que la plupart des systèmes d'exploitation prennent en charge la gestion réseau et que, par effet de bord, l'attaque par épuisement de TCP impacte l'ensemble des applications jusqu'à obtenir un état de figement du système visé.

Attaques sur les bogues des piles IP/TCP

Les piles IP/TCP développées par différents constructeurs ou fournisseurs de services manifestent des différences de comportement malgré les définitions des RFC et contiennent de multiples faiblesses, qui peuvent être exploitées par des attaques bien ciblées.

Comme il est théoriquement impossible de vérifier l'absence de bogues dans un programme conçu avec les langages de programmation modernes, il existe une forte probabilité que des bogues permettent à des pirates de gagner des privilèges.

Les principales attaques qui s'appuient sur les erreurs de programmation associées aux piles TCP/IP sont le ping de la mort, le baiser de la mort, le win nuke, l'attaque land et l'attaque teardrop.

Le ping de la mort consiste à envoyer une suite de fragments d'une requête de type écho ICMP. Une fois à nouveau assemblés par la pile IP/TCP du système cible, ces fragments forment un paquet d'une taille supérieure à la taille maximale autorisée (65 507 octets) et peuvent faire déborder les variables internes, provoquant un comportement anormal du système *(voir figure 1.28).*

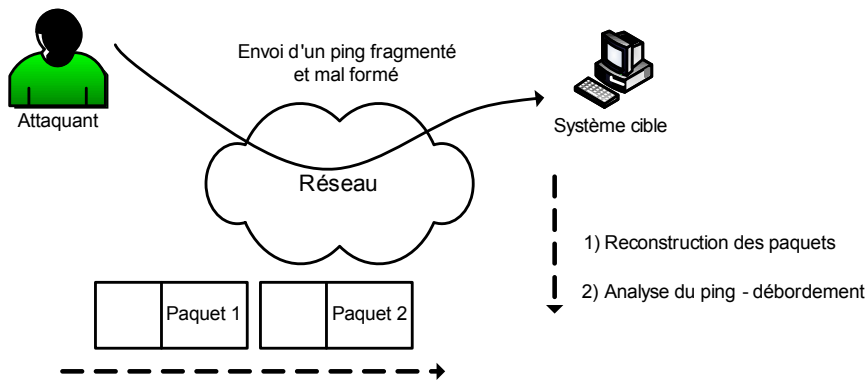


Figure 1.28

L'attaque ping de la mort

Le baiser de la mort consiste à envoyer un paquet IGMP (Internet Group Management Protocol) mal construit, mettant les machines Windows en refus de service.

Le win nuke envoie un paquet TCP mal construit avec des données OOB (Out Of Band), mettant les machines Windows en refus de service. L'impact de l'attaque peut provoquer des comportements indésirables du système cible.

L'attaque de type land demande une ouverture de session TCP avec l'adresse source du paquet égale à l'adresse destination et le port source égal au port destinataire. Cette attaque utilise principalement le port 139 TCP (NetBIOS Session) afin de viser le système d'exploitation Windows. L'impact de l'attaque peut provoquer des comportements indésirables du système cible.

L'attaque de type teardrop envoie un paquet fragmenté de telle façon que les en-têtes du second paquet écrasent ceux du premier, affolant la pile TCP/IP. Cette attaque a été conçue initialement pour les paquets ICMP fragmentés, mais de nombreuses variantes ont été développées depuis pour fonctionner avec n'importe quel type de protocole IP. L'impact de l'attaque peut provoquer des comportements indésirables du système cible.

Projection dans un monde IPv6

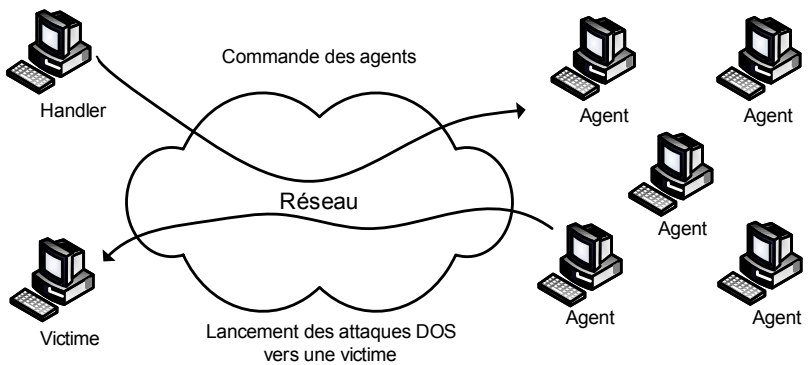
La programmation d'une pile réseau IPv6 n'est pas simple et doit suivre les recommandations internationales en termes de comportement. Il est donc à parier que l'ère d'IPv6 (encore trop immature) mettra en avant de nouvelles attaques exploitant des faiblesses de programmation et comportement de la pile.

Attaques par déni de service distribué (DDoS)

L'attaque DDoS (Distributed Denial of Service) est un dérivé de la précédente sous une forme distribuée, comme l'illustre la figure 1.29.

Figure 1.29

Attaque par déni de service distribué



La première étape consiste à pénétrer par diverses méthodes des systèmes dits handlers, ou maîtres (masters), et agents, ou esclaves (slaves). Le pirate contrôle ensuite directement un ensemble de systèmes handlers, qui contrôlent eux-mêmes un ensemble de systèmes agents. La dernière étape consiste pour le pirate à déclencher son attaque vers un ou

plusieurs systèmes cibles donnés. Cet ordre d'attaque aura été donné par les systèmes handlers, qui auront eux-mêmes reçu cet ordre du pirate.

Parmi les nombreuses attaques DDoS, citons TFN (Tribe Flood Network), historiquement la première, et Stacheldraht, qui chiffre les ordres de commandes échangés entre les handlers et les agents dans le champ données des paquets ICMP et que nous décrivons plus en détail ci-après.

Ces attaques ont fait des émules, et d'autres attaques sont apparues, telles que Trinoo, qui s'appuie sur UDP pour les communications des ordres entre handlers et agents, et TFN2K, une version entièrement revue de TFN, qui introduit des phénomènes de génération aléatoire des ports utilisés pour les communications des ordres entre les handlers et les agents, ainsi qu'un phénomène aléatoire dans le lancement des attaques vers les systèmes cibles.

L'attaque Stacheldraht : un cas d'école

Apparue en 1999, l'attaque Stacheldraht s'appuie sur le code source de TFN mais y apporte quelques variantes.

Comme TFN, Stacheldraht est constituée de programmes maîtres, ou handlers, et esclaves, ou agents. Les attaques par déni de service sont lancées par le biais d'attaques ICMP flooding, SYN flood, UDP flood, smurf, etc.

L'une des faiblesses de TFN étant que les communications entre les programmes ne sont pas chiffrées, Stacheldraht chiffre la communication.

La phase initiale de l'attaque consiste à pénétrer un grand nombre de systèmes et à y propager les programmes handlers et agents par le biais d'une vulnérabilité quelconque. En 1999, par exemple, des vulnérabilités de type buffer overflow avaient été utilisées pour pénétrer des systèmes Solaris sur les services de type RPC tels que statd, cmsd et ttdbserverd.

Comme expliqué précédemment, l'objectif de ces attaques est de créer une hiérarchie de handlers et d'agents afin d'attaquer les systèmes cibles par déni de service.

La figure 1.30 illustre l'architecture de ces attaques par déni de service distribué.

Le pirate (client) contrôle un ou plusieurs handlers, et chaque handler contrôle plusieurs agents. Les attaques par déni de service sont coordonnées sur les plages d'adresses IP données par le handler responsable des agents.

Les communications entre le client et le handler s'effectuent par le biais de communications chiffrées utilisant un algorithme symétrique. Plus précisément, le client se connecte d'abord à un handler — une adresse IP suffit —, et un mot de passe est demandé au client. Le client entre le mot de passe par défaut, sicken, qui est crypté localement par le biais du programme Unix crypt. Le mot de passe est alors envoyé au handler. La communication entre le client et le handler est également chiffrée à l'aide de l'algorithme Blowfish avec la clé de chiffrement symétrique.

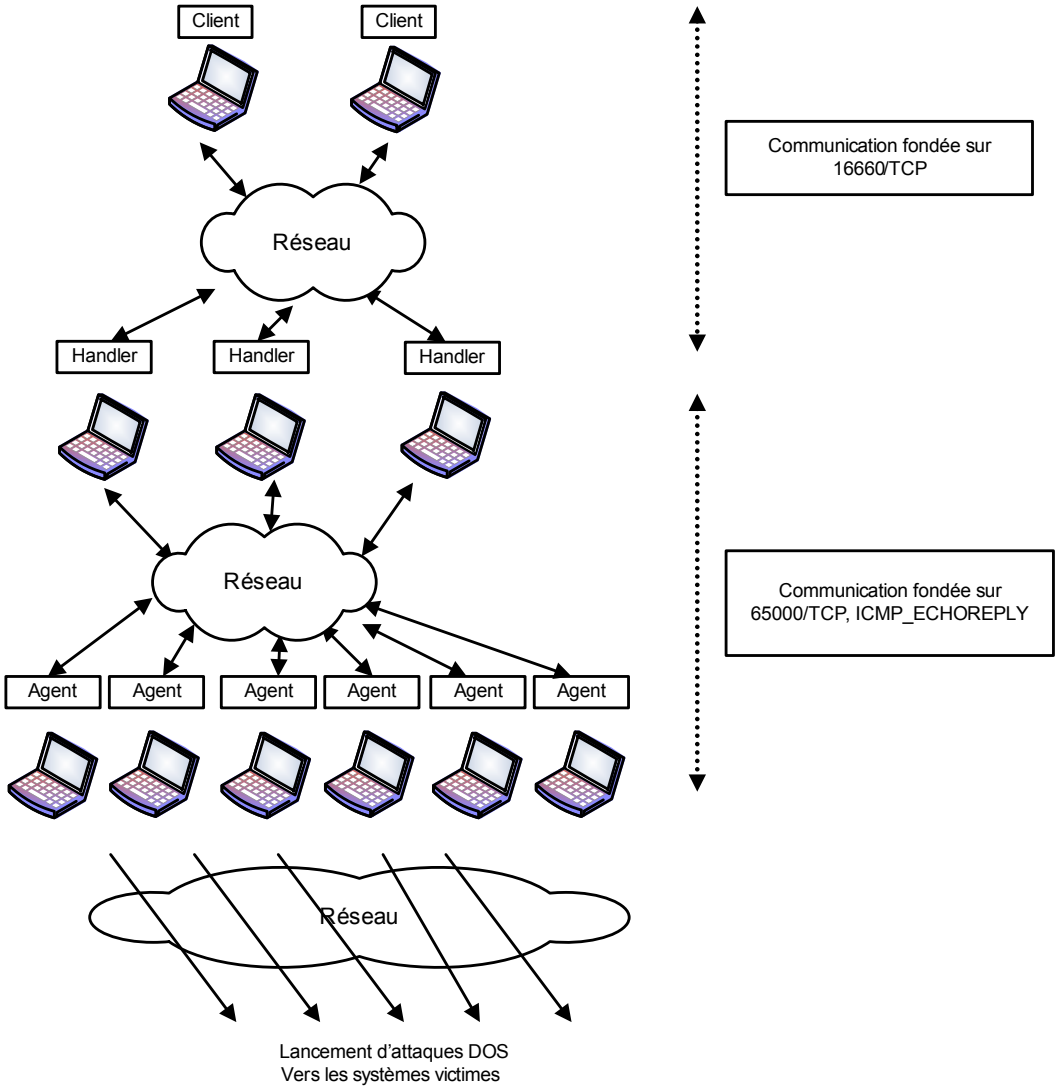


Figure 1.30

Architecture des attaques par déni de service distribué

Les commandes disponibles sur un handler sont les suivantes :

```
.help
Imprime l'aide associée aux commandes.

.killall
Tue tous les agents actifs.
```

```
.madd ip1[:ip2[:ipN]]
Ajoute les adresses IP suivantes dans la liste des systèmes cibles.

.mdie
Envoie une requête pour tuer tous les agents.

.mdos
Lance une attaque DOS.

.micmp ip1[:ip2[:ipN]]
Lance une attaque ICMP flooding sur la liste des systèmes donnés.

.mlist
Imprime la liste des systèmes cibles subissant une attaque DOS.

.mping
Ping les agents pour vérifier qu'ils sont vivants.

.mstop ip1[:ip2[:ipN]]
.mstop all
Arrête l'attaque sur les systèmes cibles.

.msyn ip1[:ip2[:ipN]]
Lance une attaque SYN flooding pour la liste des systèmes donnés.

.mtimer seconds
Fixe la durée de l'attaque.

.mudp ip1[:ip2[:ipN]]
Lance une attaque UDP flooding pour la liste des systèmes donnés.

.setisize
Définit la taille des paquets ICMP pour le flooding (par défaut 1024).

.setusize
Définit la taille des paquets UDP pour le flooding (par défaut 1024).

.showalive
Montre les agents vivants.

.showdead
Montre tous les agents morts.

.sprange lowport-highport
Permet de fixer le range des ports utilisés lors des attaques
➥ de SYN flooding (lowport 0, highport 140).
```

La communication entre le handler et un agent s'effectue à l'aide de ICMP echo-reply sur une connexion au port TCP 65000.

La première étape, lorsqu'un agent démarre, consiste à lire un fichier contenant l'adresse IP du ou des handlers dont il dépend. Sinon, il essaye les adresses IP 1.1.1.1 et 127.0.0.1. Une fois la liste déterminée, il envoie un paquet ICMP echo-reply contenant dans le champ données un ID=666 avec la phrase skillz. Si un handler reçoit ce paquet, il répond avec un paquet ICMP echo-reply contenant dans le champ données un ID=667 avec la phrase ficken.

Dans un second temps, l'agent envoie un paquet ICMP echo-reply avec une adresse source fausse (3.3.3.3) contenant dans le champ données un ID=666 et l'adresse IP du système sur lequel il se trouve. Ce test permet de vérifier que des attaques par déni de service peuvent être lancées.

Une fois que le handler reçoit le message, il répond à l'agent avec l'adresse contenue dans le champ données du paquet ICMP echo-reply et envoie alors un paquet ICMP echo-reply avec un ID=1000 dans le champ données.

Une fois la communication établie, de nombreux autres types de paquets peuvent être échangés afin de transmettre les commandes entre le handler et l'agent, que nous ne détaillerons pas davantage.

Les faiblesses des protocoles sont donc des vecteurs d'attaque qui peuvent être exploités de diverses manières. De plus, les protocoles réseau n'ayant prévu aucun mécanisme d'authentification véritable, ils subissent des attaques qui s'appuient sur ces faiblesses d'authentification.

L'évolution des attaques par déni de service

La maîtrise des logiciels distribués a rendu plus facile la création de botnets regroupant un ensemble de machines zombies exploitées à des fins malveillantes. La puissance de feu d'attaques par déni de service est donc considérable et difficilement maîtrisable par les opérateurs de télécommunications.

Basées généralement sur des attaques classiques (par inondation SYN), les attaques par déni de service peuvent mettre en péril une infrastructure, un réseau, un serveur ou une application donnée (service voix/ip). À titre d'exemple, une attaque de grande envergure avait visé les serveurs DNS racine en 2007. Déjà mis en péril par de telles attaques, les serveurs DNS racine ont adopté une stratégie d'invisibilité par un adressage de type anycast. La technologie anycast, permettant la présence de plusieurs machines ayant une même adresse IP en plusieurs points du réseau Internet, a masqué aux attaquants le nombre de systèmes rendant le service ainsi que leur répartition sur les différents lieux géographiques.

Il faut cependant noter que cet exemple reste à part et illustre surtout les surfaces d'attaques possibles sur les systèmes/services mis en œuvre sur le réseau. Dans ce contexte, l'opérateur de télécommunications joue un rôle crucial, car la suppression de telles attaques à sa périphérie reste une des meilleures contre-mesures.

Attaques spécifiques à IPv6

Bien que la plupart des attaques IPv4 restent valables en IPv6, ce dernier apporte son lot spécifique d'attaques liées à cette évolution protocolaire.

Attaque par manipulation des en-têtes

Sachant que l'en-tête d'un paquet IPv4 n'est pas *a priori* de taille fixe, IPv6 y remède en imposant un en-tête initial de taille fixe et introduit la notion d'extensions remplaçant la notion d'options en IPv4.

Comme l'illustre la figure 1.31, en IPv6, un champ next header indique soit un protocole de niveau supérieur, tel que TCP, soit une extension.

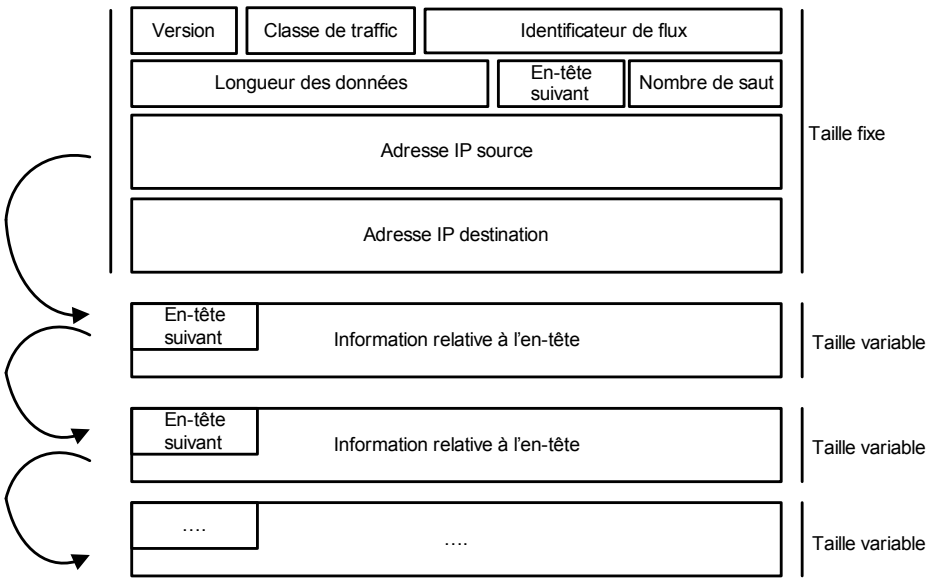


Figure 1.31

Architecture des attaques par déni de service distribué

Dans le cas d'une extension, cette dernière a un champ next header autorisant le chaînage. Dans le cas ou plusieurs extensions sont chaînées, il est recommandé qu'elles apparaissent dans l'ordre suivant :

- Valeur du next header = 0 : il s'agit de l'extension hop-by-hop.
- Valeur du next header = 43 : il s'agit de l'extension routing.
- Valeur du next header = 44 : il s'agit de l'extension fragment.
- Valeur du next header = 51 : il s'agit de l'extension authentication.

- Valeur du next header = 50 : il s'agit de l'extension encapsulating security Payload.
- Valeur du next header = 60 : il s'agit de l'extension destination.

On peut imaginer toutes sortes d'attaques utilisant un chaînage ne respectant pas l'ordre recommandé, un chaînage ne respectant pas les normes, un chaînage le plus long possible, etc., autant de nouvelles menaces à la fois pour le système destinataire et les systèmes en charge de la sécurité, tels que pare-feu ou systèmes de détection d'intrusion.

De manière plus spécifique et à titre d'exemple :

- L'extension routing indique un mécanisme permettant de mener des attaques en imposant un routage par la source du message, ce qui peut être extrêmement dangereux. Cependant, cette extension étant aussi utilisée pour offrir la mobilité IPv6, il importe de « trier le bon grain de l'ivraie ».
- En dehors des autres extensions, seule l'extension hop-by-hop est lue par tous les sauts IP par lesquels transite le paquet IPv6. Cette extension peut donc être utilisée à des fins néfastes et peut pénaliser les ressources mémoire et processeur des équipements réseau en transit.

Attaque par les dual stack

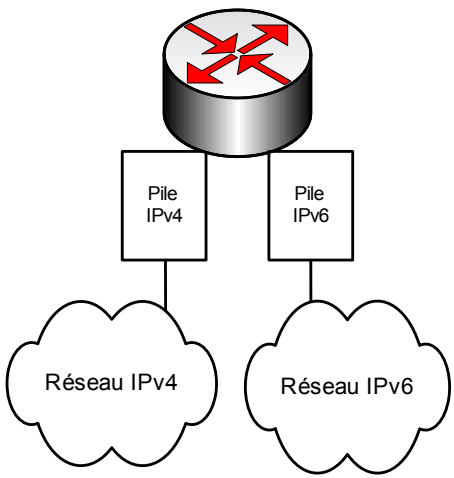
IPv4 et IPv6 devront cohabiter pendant un certain temps, si bien que l'un pourra mettre en péril l'autre, compte tenu du principe du maillon le plus faible.

Deux techniques sont proposées pour cette migration (IPv4 vers IPv6) :

- dual stack *(voir figure 1.32)* : il s'agit d'un mécanisme offrant un support complet des deux protocoles aux systèmes et équipements réseau traversés. La sécurité repose sur l'étanchéité du code de la pile/stack réseau, mais aussi sur la configuration globale des systèmes.

Figure 1.32

Architecture des attaques par déni de service distribué



- 6to4 *(voir figure 1.33)* : il s'agit d'un service de tunnel automatique reliant des nœuds IPv6 par l'intermédiaire d'un réseau IPv4. Cette technique doit gérer proprement l'encapsulation IPv6 relative au trafic et au routage. De plus, la configuration de l'encapsulation nécessite la plus grande attention et peut être sujette à de nombreuses erreurs.

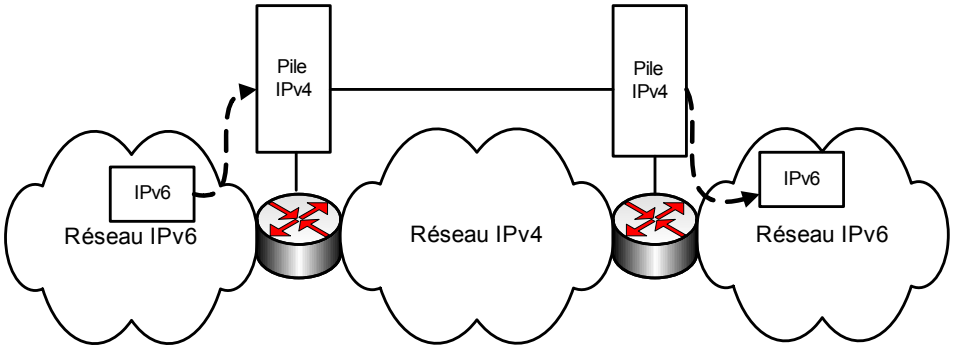


Figure 1.33

Architecture des attaques par déni de service distribué

Les débuts de la cohabitation seront probablement houleux et sujets à de nombreuses attaques :

- Comment s'assurer de l'étanchéité d'IPv6 *via* la pile IPv4 ?
- Comment s'assurer de l'étanchéité d'IPv4 *via* la pile IPv6 ?

Ces deux questions sont délicates et dépendent des implémentations sur les systèmes d'exploitation considérés. Que ce soit les systèmes d'exploitation Windows, Linux, FreeBSD, IOS, etc., tous devront assurer une isolation et une étanchéité parfaite entre les mondes IPv4 et IPv6.

Autres formes d'attaques

L'accès physique aux équipements réseau permet de prendre la main en tant qu'administrateur sur pratiquement tous les systèmes actuels. Cela peut impacter fortement le réseau si un pirate en profite pour modifier directement les tables de routage internes.

La copie des configurations des équipements réseau est une attaque redoutable, qui permet au pirate de reconstituer tout le réseau logique ainsi que les protections mises en place. La configuration des équipements réseau est, par nature, une information confidentielle du réseau.

L'écoute électronique pour récolte d'information peut permettre de mener des attaques ciblées. Les diverses techniques d'écoute disponibles actuellement permettent d'écouter n'importe quel type de média.

Le vol de secret se rencontre plus fréquemment dans l'ingénierie sociale. Par exemple, l'agresseur entre en contact avec la personne qu'il veut usurper en se faisant passer pour un technicien en intervention bloqué dans son travail par une demande d'authentification ou une permission trop forte. Pour peu qu'il soit convaincant, l'agresseur peut obtenir les couples compte/mot de passe ou permissions qu'il désire, voire directement ceux de l'administrateur système.

Une variante de cette attaque consiste à obtenir un compte privilégié créé directement par un administrateur trouvant cette procédure plus « sécurisée »...

En résumé

Les attaques réseau reposent sur un ensemble de faiblesses de sécurité touchant différents domaines, tels que les protocoles réseau, les implémentations des piles réseau et les systèmes d'exploitation des systèmes réseau.

Les attaques réseau touchent beaucoup d'autres protocoles, que nous n'avons pas décrits dans ce chapitre, tels que les protocoles VoIP (voix sur IP), qui n'implémentent pas non plus de couche de sécurité et qui s'exposent en premier lieu aux attaques par usurpation d'identité. Citons également le protocole DNS, qui subit lui aussi des attaques répétées pouvant mettre hors de fonctionnement les services d'un réseau.

La mise en place de couches de sécurité telles que IPsec ou SSH pour créer des tunnels chiffrés et authentifiés ne met pas à l'abri d'attaques. Ces dernières visent généralement les faiblesses d'implémentation des piles de sécurité. D'autres attaques, profitant des faiblesses des protocoles de sécurité, ont permis de faire évoluer ces derniers.

Le chapitre 2 détaille les méthodes et techniques d'intrusion permettant de prendre le contrôle d'un système réseau. Ces attaques reposent sur les faiblesses de sécurité des systèmes d'exploitation liés aux équipements réseau.