Fichiers logs : preuves judiciaires et composant vital pour Forensics

BENSEFIA Hassina
Centre de Recherche sur l'Information Scientifique et Technique
Laboratoire des Logiciels de Base
hbensefia@mail.cerist.dz

1. Introduction

La sécurité des réseaux dépend de la technologie préventive. Elle consiste à mettre en place des systèmes de défense adéquats pour réduire le risque des attaques potentielles. Cependant, le spectre des menaces change. Malgré la présence des systèmes de défense, un réseau demeure vulnérable et sa sécurité est compromise par des menaces non identifiées. La technologie préventive s'avère insuffisante pour répondre aux besoins de la sécurité. Les attaques informatiques évoluent et envahissent Internet et le monde digital convergeant vers des crimes informatiques résultant en des pertes significatives de productivité. Alors d'autres éléments clés complémentaires à la technologie préventive sont introduits. La sécurité d'un système nécessite d'être surveillé de manière permanente à fin d'inspecter les activités suspectes. Aussi la mise en place d'un système judiciaire est indispensable pour identifier les actes illicites aux systèmes et leurs auteurs. D'où l'émergence de Computer Forensics ou l'informatique judiciaire. Apparue comme une réaction au crime informatique, Forensics utilise les techniques d'investigation et s'alimente des fichiers logs qui peuvent contenir le comportement de l'attaquant et les étapes de l'attaque. En effet ; ces fichiers constituent le moyen de surveillance de la sécurité et une source critique de preuve pour une investigation informatique.

2. la sécurité préventive

La sécurité des réseaux a pour objectif d'empêcher les éléments malicieux d'atteindre un réseau [3], elle est fondamentalement basée sur la prévention [2]. Les techniques préventives consistent à déterminer les menaces potentielles et monter des systèmes adéquats pour les empêcher. Les systèmes de sécurité préventifs tels que les firewalls agissent comme des obstacles électroniques dressant une barrière face aux attaques potentielles, ils prennent en compte un nombre limité de menaces [12]. Ils sont passifs devant les nouvelles menaces inconnues.

L'utilisation d'un système de sécurité n'est pas rassurante, même s'il représente la meilleure stratégie de sécurité. La propagation des virus et du code malicieux dans le e-mail et les pages web sont des exemples démonstrateurs combien il est possible de compromettre le meilleur système de sécurité créé par la technologie préventive [12].

La technologie préventive a contribué pour résoudre les problèmes de sécurité via une multitude de systèmes mais les menaces existent toujours et la majorité des réseaux demeurent vulnérables et sont sous attaques d'une façon constante. La prévention offre une sécurité statique, passive et fondamentalement fragile.

Pour atteindre une sécurité efficace, la sécurité préventive a besoin d'une couche supplémentaire de protection. Deux éléments qui étaient pratiquement absents dans le paradigme de la sécurité informatique s'imposent comme concepts clés complétant le rôle de la sécurité préventive. Ces deux éléments sont la surveillance de la sécurité et la mise en application de la loi.

3. La surveillance de la sécurité

Elle consiste à surveiller l'activité des systèmes et des réseaux. Elle se concrétise par la vérification en temps réel des traces du trafic entrant et sortant générés par un système tels que les firewalls, les routeurs, les serveurs, etc [2]. Ces traces contiennent l'information concernant les évènements qui arrivent dans le système. L'analyse de cette information permet de détecter les activités suspectes, les attaques en cours et de déterminer si la sécurité du système a été compromise ou en voie de compromission.

La surveillance de la sécurité d'un système de manière suffisante et en temps réel permet de détecter un attaquant indifféremment de quelle vulnérabilité il a exploité pour pénétrer dans le système. L'attaquant peut être repoussé au point d'entrée du réseau avant qu'il n'accomplisse son attaque.

4. La mise en application de la loi

Internet et l'emploi de l'ordinateur personnel représentent des outils permettant l'évolution de la société sur le plan technologique, économique, sociologique et culturel. Cependant ces mêmes outils sont devenus les armes des criminels durant la deuxième partie du vingtième siècle [4] [11]. Le comportement on-line étant anonyme, les humains peuvent voler, tricher, mentir et commettre des crimes sans être aperçus [1]. Un individu peut attaquer le réseau d'une organisation en s'installant derrière son ordinateur personnel et en se connectant à Internet. Il peut pirater un compte e-mail, envoyer un virus à un réseau ou compromettre un serveur vital.

Internet est encore sans loi [1] [2], plusieurs personnes et organisations se retrouvent victimes des criminels sur Internet sans aucun recours légal. La mise en application de la loi consiste en la poursuite judiciaire des attaquants et la condamnation des criminels. Elle envoie un message d'avertissement à tout le monde que l'utilisation illégale d'Internet sera sanctionnée. Elle permet de garder les criminels loin d'Internet pendant leur emprisonnement, de punir ceux qui ont endommagé les réseaux des organisations et de rendre Internet une société légale.
La surveillance de la sécurité et la mise en application de la loi complètent les insuffisances de la technologie préventive. La surveillance de la sécurité permet de détecter les nouvelles menaces inconnues. La mise en application permet d'agir aux attaques initiées.

Le crime informatique est un problème sérieux et épineux, plusieurs organisations ont perdu leur productivité et leur réputation à cause d'attaques variant entre attaques directes et indirectes dans diverses formes. Pour identifier les crimes informatiques et leur auteurs, les méthodes du Forensics sont nécessaires [3] [4]. La sécurité informatique a donc besoin d'un système judiciaire pouvant collecter et analyser des données relatives aux activités malicieuses et aux attaques pour identifier les attaquants et les poursuivre judiciairement. Ce besoin s'est traduit par l'émergence de Forensics dans le domaine de l'informatique pour

donner naissance à une nouvelle discipline dite « Computer Forensics » ou « l'informatique légale » ou encore l'informatique judiciaire.

Computer Forensics englobe aussi la surveillance de la sécurité à travers la collecte et l'analyse d'information, liées aux activités des systèmes et des réseaux, à la recherche de preuve d'attaque ou d'activité illicite qui est l'étape préliminaire dans le processus de Forensics.

4. Forensics

Le terme Forensics est très proche du terme forum. Un forum étant un endroit public pour la discussion et le débat, c'est aussi une cour. Ainsi la signification de forensics se rapporte aux cours et à la loi, forensics veut dire légal ou judiciaire.

Le terme Forensics désigne « Forensic science » ou « science légale », c'est l'application de la science à la loi [13]. Elle permet de résoudre les conflits légaux en utilisant des méthodes scientifiques. Elle consiste à rassembler des signes prouvant qu'un crime est survenu pour identifier le coupable. Un crime peut être un vol, une fraude, un meurtre ou un suicide. Les signes sont acquis en utilisant certains domaines scientifiques comme la biologie, la chimie, la science analytique, la psychologie et les mathématiques [13].

Forensics couvre un éventail de disciplines telles que la pathologie légale, la sérologie légale, la balistique légale, l'anthropologie légale, la psychologie légale, la géologie légale, …etc. qui ont toutes pour rôle la résolution d'un crime. Pour trouver des signes pouvant identifier un crime, l'expert en Forensics ou l'expert légiste examine les traces du matériel pouvant inclure ou exclure une association entre un suspect et une victime. Ces traces peuvent être le sang, la salive, les impressions de chaussures et de pneus, les poils, les armes à feu et les documents. L'expert légiste étudie aussi les causes des incendies, des explosions et des accidents de la route [13].

L'analyse faite par l'expert légiste est présentée sous forme d'un état écrit pouvant être lu dans un tribunal. Le résultat d'un cas civil ou criminel dépend de l'analyse faite par l'expert légiste et il est influencé par son interprétation. Pour cette raison l'expert légiste doit être doté d'une expérience professionnelle et d'une compétence ainsi que d'une intégrité personnelle. Il doit être impartial et ayant un intérêt neutre dans le résultat d'un cas.

5. Computer Forensics

Une discipline émergeante de la collection et de l'analyse de l'information concernant les ordinateurs et les réseaux d'ordinateurs pour détecter les activités suspectes ou malicieuses en vue de l'utiliser comme des preuves dans la cour de la loi [3]. *C'est l'art de découvrir et d'extraire l'information prouvant l'occurrence d'un crime informatique de telle façon à le rendre admissible dans le tribunal* [4].

Considérée comme l'autopsie du disque dur d'un ordinateur, Computer Forensics est une nouvelle science qui est apparue comme une réaction au crime informatique. Elle permet d'appliquer la loi à l'informatique, en employant les techniques d'investigation pour identifier la cause origine d'un crime informatique. Un crime informatique peut être une attaque, une intrusion ou toute activité malicieuse.

6. Network Forensics

Elle est au cœur de computer Forensics [12]. *C'est le pouvoir de collecter les données critiques nécessaires pour suivre et analyser l'utilisation illégitime des réseaux et des applications du réseau.*

« Network Forensics » étant un processus qui permet de collecter les données à partir des différents dispositifs du réseau et d'appliquer les techniques d'investigation afin de retracer les activités se passant dans le réseau et ce ci dans le but d'identifier une attaque et de découvrir l'identité de l'attaquant durant et après l'attaque [12]. Le but de Network forensics est de fournir la preuve suffisante pour permettre de poursuivre judiciairement avec succès l'auteur d'un crime informatique.

Comme l'ouverture des réseaux et le degré de leur connectivité représentent le paramètre significatif dans l'évolution du crime informatique. Nous allons accentuer le point sur le Network Forensics

7. Les étapes principales de Network Forensics

7.1. La collecte de la preuve informatique

La preuve est l'information utilisée pour décider si une proposition ayant été sujet de débat est vraie [10]. La preuve informatique est l'ensemble des données collectées à partir des différents composants d'un réseau tels que les firewalls, les routeurs et les serveurs [4]. Les données en question sont les données concernant les paquets TCP/IP traversant le réseau. Chaque paquet contient une entête incluant le temps et la date de connexion, l'adresse IP source et destination, le type de session (FTP, Tel net, e-mail,…) et la durée de session. Ces données sont d'un intérêt majeur lors d'une attaque. L'adresse IP source peut révéler l'identité de l'attaquant, l'adresse IP destination peut révéler l'élément cible dans le réseau et la donnée dans le paquet identifie l'attaque survenue. Ces données permettent aussi de confirmer ou non l'occurrence d'une attaque et de documenter les vulnérabilités qui puissent exister dans le réseau [7]. La collecte de la preuve doit avoir lieu avant l'occurrence d'une attaque [6]. L'information collectée est considérée comme une preuve potentielle, après l'occurrence d'une attaque ou d'une activité malicieuse, elle devient une preuve réelle [4]. L'utilisation de la preuve informatique dans un tribunal est conditionnée par son intégrité et son authenticité [4]. Sa protection est primordiale dans Forensics.

7.2. L'investigation informatique

C'est une étape très importante dans Network Forensics. C'est une procédure qui permet la résolution d'une attaque après son occurrence. Elle consiste à vérifier et analyser la preuve informatique collectée afin d'aboutir à la preuve judiciaire affirmant ou réfutant l'occurrence d'un crime informatique [7]. Elle permet de déterminer si une attaque est survenue, la nature de l'attaque, l'auteur de l'attaque et les traces qu'il a laissées derrière lui.

L'expert en Network Forensics doit connaître l'emplacement de la preuve et doit avoir des compétences dans la collecte de la preuve [7]. Il doit mener la procédure d'investigation avec précaution et il doit être apte à suivre cette procédure jusqu'à la fin. Pour cela, il doit avoir les compétences suivantes [4] [11] :

- Il doit être expert en information et en administration des réseaux.
- Il doit être expert dans les techniques de piratage et familier avec les vulnérabilités des systèmes.

- Il doit avoir les compétences informatiques d'un attaquant.
- Il doit avoir les compétences investigatrices d'un détective.
- Il doit avoir les compétences judiciaires d'un juriste.

8. Emplacement de la preuve informatique

La preuve informatique dépend du type d'attaque [4], elle peut se trouver dans trois emplacements principaux [5]

1) Sur le composant victime du réseau.
2) Sur la machine de l'attaquant.
3) Sur les dispositifs du réseau situés entre le composant victime et la machine de l'attaquant.

Les données faisant objet de preuve informatique sont générée par le logging [6] [10]. Le mécanisme de logging ou l'enregistrement des activités est une fonctionnalité de la majorité des systèmes d'exploitation modernes. C'est la sauvegarde de la trace des événements qui arrivent pendant leur exécution. L'enregistrement peut avoir lieu dans un hôte ou dans un système fournissant un service réseau tel qu'un serveur de messagerie, un serveur Web, un serveur de nom de domaine (DNS) ou un firewall. L'enregistrement prend la forme d'un fichier dit « fichier log ». Les fichiers logs des différents composants d'un réseau représentent la source de preuve pour Network Forensics [4]. Ils contiennent des informations liées au comportement de chaque utilisateur spécifique dans le réseau ainsi que le temps et la durée de ses activités. Ils enregistrent toutes les activités et les événements se passant dans le réseau telles que les sessions Telnet et FTP, l'accès Web et messagerie. Le type d'information contenue dans les fichiers logs dépend des applications disponibles sur le réseau.

9. Les fichiers logs

L'expression « fichier log » signifie « le journal de bord des connexions », c'est l'historique des requêtes adressées à un système [9]. Un fichier log est un fichier créé par un logiciel spécifique installé sur un système. Il contient des informations concernant l'activité du système. Il a une structure ASCII (American Standard Code for Information and Interchange) qui est lisible par les humains. Le contenu d'un fichier log dépend du niveau d'enregistrement et du type d'activité du système.

Un fichier log contient des entrées. Chaque entrée est une ligne, elle représente une requête que le système a reçue, la réponse à cette requête et le temps de traitement de cette requête. Une entrée d'un fichier log est composée de champs élémentaires de données. Chaque champ désigne une information concernant la requête telle que le nom d'utilisateur, son adresse IP, la requête adressée par l'utilisateur au système, la réponse du système, la date et le temps de soumission de la requête, le protocole utilisé et d'autres informations spécifiques à la requête.
Un fichier log représente une base de données textuelle listée par le champ temps.

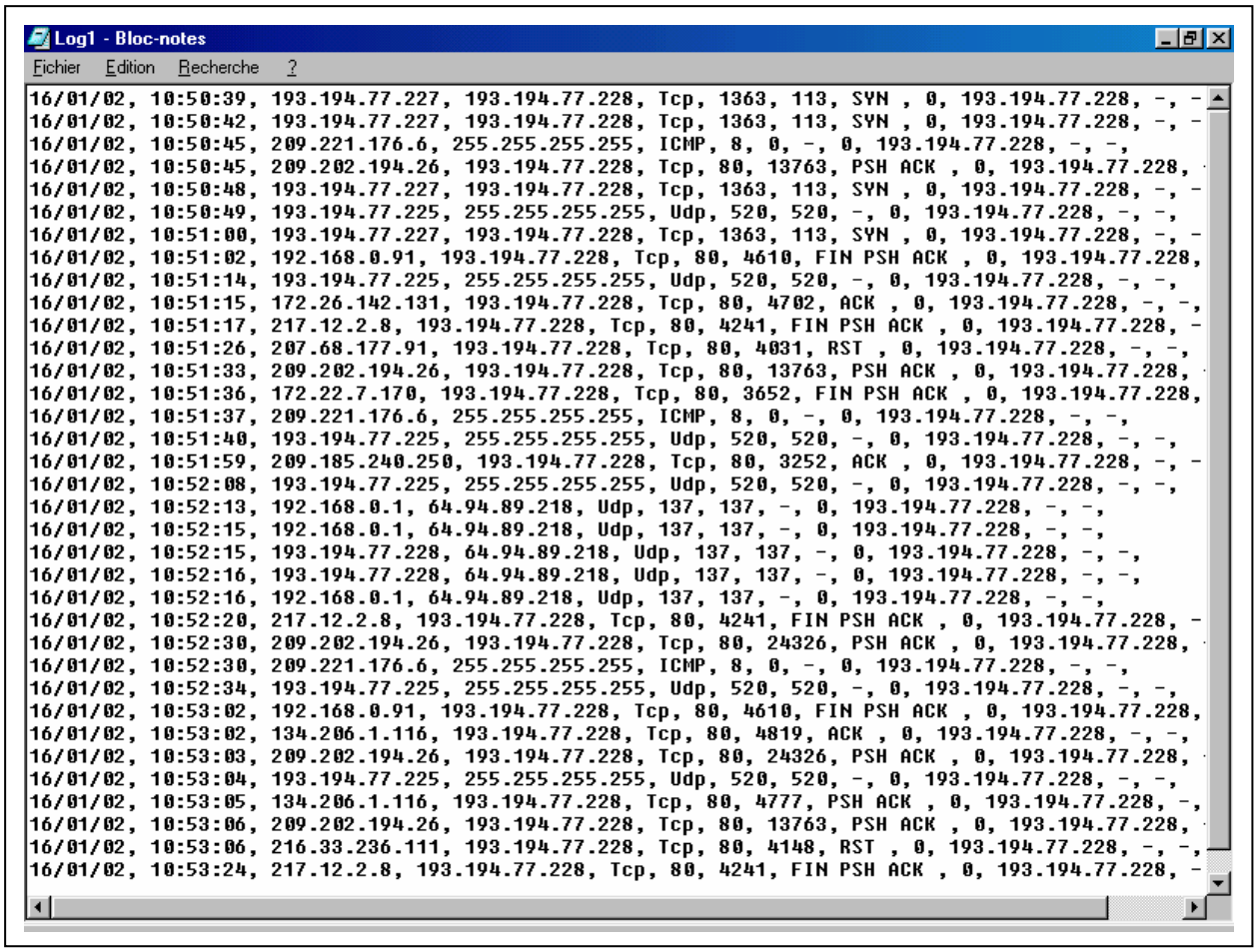

Figure 1 : Extrait d'un fichier log d'un serveur Proxy

On remarque que chaque entrée contient un ensemble de champs qui sont séparé par une virgule. Si on prend la première entrée de ce fichier :
16/01/02, 10:50:39, 193.194.77.227, 193.194.77.228, TCP, 1363, 113, SYN, 0, 193.194.77.228,

Voici la signification de chaque champ :

- 16/01/02 : est la date de réception du paquet.
- 10:50:39 : est l'heure de réception du paquet.
- 193.194.77.227 : est l'adresse source indiquant l'adresse IP de la machine émettrice du paquet.
- 193.194.77.228 : est l'adresse destination indiquant l'adresse IP de la machine réceptrice du paquet
- TCP : est le protocole utilisé.
- 1363 : est le port source indiquant le numéro de port relatif à l'application en cours sur la machine émettrice du paquet.
- 113 : est le port destination indiquant le numéro du service s'exécutant sur la machine réceptrice du paquet.
- SYN : est la valeur du drapeau TCP indiquant une requête d'établissement d'une connexion.
- 0 : ce champ indique la règle de filtrage et il prend pour valeur 0 ou 1. S'il est égal à 1, le paquet est accepté. S'il est à 0, le paquet est rejeté.
- 193.194.77.228 : est l'adresse IP de l'interface sur la quelle le paquet est reçu.

La majorité des systèmes génèrent des fichiers logs dans un format propriétaire qui est mystérieux et difficile à le déchiffrer. Il existe plusieurs formats de fichiers logs spécifiques

aux serveurs Web, serveurs de messageries et les firewalls. Seulement les développeurs peuvent comprendre leur contenu. L'analyse du contenu d'un fichier log exige une bonne compréhension du format et une structure appropriée facilitant son analyse. En effet ; des formats standards ont été développés. Parmi les quels, on peut citer : le format W3C Etendu et le format log commun.

9.1. Le format W3C Etendu (Word Wide Web Consortium) [8]

C'est un format ASCII adapté avec une variété de champs. L'utilisation de ce format permet d'inclure des champs importants comme elle peut omettre des champs non désirés. Les champs sont séparés par un espace, le temps est enregistré en GMT (Greenwich Mean Time). Ce format est disponible pour les serveurs Web et les serveurs FTP.

Exemple d'une entrée
172.16.25.10 02-05-1988 17:42:15 GET /default.html 200 HTTP/1.0

Cette entrée signifie que le 02 mai 1998 à 17h 42 mn 15s (GMT), un utilisateur ayant l'adresse IP 172.16.25.10 et en utilisant HTTP 1.0 a lancé une requête GET/default.html. Cette requête signifie le téléchargement de la page Web default.html. La requête est exécutée avec succès. La valeur 200 étant le code de l'état du service, cette valeur indique le succès de la requête.

9. 2. Le format log commun [8]

C'est un format ASCII fixé disponible uniquement pour les serveurs Web. Il a été développé par NCSA (National Center for Supercomputing Applications) à l'université d'Illinois à Urbana-Champaign. Chaque entrée contient les champs suivants :

- Le nom de l'hôte ou l'adresse IP de l'hôte
- Le nom de l'utilisateur
- La date de soumission de la requête
- Le temps de soumission de la requête
- Le contenu de la requête envoyée par le client
- Le code de l'état HTTP retourné à l'utilisateur : c'est le code de la réponse HTTP envoyé par le serveur au client.
- La taille en octets des informations envoyées par le serveur

Les champs sont séparés par un espace, le temps enregistré est le temps local.

Exemple d'une entrée
172.21.13.45 FRED 08-04-1998 17 :39 :10 GET/scripts/iisadmin/ism.dll 200 3401

Cette entrée indique qu'un utilisateur appelé FRED utilisant une adresse IP 172.21.13.45 a envoyé au serveur Web une requête en utilisant la commande HTTP qui est « GET » pour télécharger le fichier scripts/iisadmin/ismi.dll. Cette requête a été soumise au serveur le 08 Avril 1998 à 17h39mn 10 s temps local et elle a été retournée avec succès (code d'état HTTP= 200). Les données envoyées à l'utilisateur FRED ont une taille de 3401 octets.

Remarque

La réponse d'un serveur à un client peut être faite avec succès ou échec. Le code de l'état de la réponse peut renseigner sur ce résultat.

- Si le code d'état $\in$ {200,201,202,203,204,300,301,302,303,304, Alors le serveur a répondu au client avec succès.

- Si le code d'état ∈ {400, 401, 402, 403, 404, 500, 501, 502, 503}, Alors le serveur n'a pas répondu à la requête, il y a eu un échec et par conséquent la taille des données transférées au client est égale à 0.

10. L'importance des fichiers logs

Les fichiers logs tracent tous les événements qui arrivent pendant l'activité d'un système. Ils peuvent contenir la preuve en détail de toute activité exceptionnelle, suspecte ou non désirée. Les fichiers logs issus des différents composants d'un réseau peuvent indiquer si la sécurité du réseau est compromise ou en voie de compromission [6]. Ils sont la seule information que l'attaquant laisse derrière lui après son introduction dans un réseau, ils représentent l'empreinte de l'attaquant [12]. Lors d'une attaque, l'information contenue dans les fichiers logs peut être vérifiée pour définir les traces de l'attaque et aboutir à une preuve accusatrice.
Les fichiers logs sont un composant vital et critique pour Network Forensics [12]. Les informations pertinentes contenues dans les fichiers logs représentent la preuve qui est le besoin indispensable pour l'investigation [4] [6]. C'est le seul moyen pour identifier l'attaquant a fin de le poursuivre judiciairement.

11. Le traçage des attaque à travers l'analyse des fichiers logs

A travers les deux exemples suivants, nous allons exploiter des extraits de fichiers logs illustrant des activités malicieuses. Nous allons analyser et interpréter le contenu de ces logs pour déterminer les entrées qui représentent des activités malicieuses, identifier l'objectif de ces activités ainsi que leurs origines.

Exemple 1
Soit un extrait d'un fichier log « */etc/log/secure* » d'un serveur de log distant *syslog d'unix*. Ce serveur est destiné à recevoir et stocker tous les logs issus de tous les systèmes composant un réseau.

```
Apr 10 13:43:48 10.10.10.1 in.ftpd[6613]: connect from 192.168.1.225
Apr 10 13:43:51 10.10.10.2 in.ftpd[6613]: connect from 192.168.1.225
Apr 10 13:43:54 10.10.10.3 in.ftpd[6613]: connect from 192.168.1.225
Apr 10 13:43:57 10.10.10.4 in.ftpd[6613]: connect from 192.168.1.225
Apr 10 13:43:58 10.10.10.5 in.ftpd[6613]: connect from 192.168.1.225
Apr 10 13:43:59 10.10.10.6 in.ftpd[6613]: connect from 192.168.1.225
Apr 10 13:44:01 10.10.10.7 in.ftpd[6613]: connect from 192.168.1.225
Apr 10 13:44:04 10.10.10.8 in.ftpd[6613]: connect from 192.168.1.225
Apr 10 13:44:07 10.10.10.9 in.ftpd[6613]: connect from 192.168.1.225
Apr 10 13:44:09 10.10.10.10 in.ftpd[6613]: connect from 192.168.1.225
```

Toutes ces entrées montrent qu'un système identifié par l'adresse IP 192.168.1.225 scanne de manière séquentielle les machines du réseau de la classe IP 10.10.10.0. Il s'agit de connexions répétées au port 21 qui est réservé au service FTP(File Transfert Protocol). L'objectif de ce scan est de rechercher la présence de la vulnérabilité de wu-ftpd. Wu-ftpd étant le daemon de FTP pour les systèmes Unix.

Exemple 2
Soit un extrait d'un fichier log d'un serveur Web IIS de Microsoft :

10.10.10.20 - - [14/Jan/2004:22:30:25 -0500] "GET / HTTP/1.1" 404 299
10.10.10.20 - - [14/Jan/2004:22:30:28 -0500] "GET / HTTP/1.1" 404 299
10.10.10.20 - - [14/Jan/2004:22:31:29 -0500] "GET / HTTP/1.1" 404 299
10.10.10.20 - - [14/Jan/2004:22:31:32 -0500] "GET / HTTP/1.1" 404 299
10.10.10.20 - - [14/Jan/2004:22:31:35 -0500] "GET / HTTP/1.1" 404 299
10.10.10.20 - - [14/Jan/2004:22:34:11 -0500] "GET / HTTP/1.1" 404 299
10.1.4.25 - - [18/Jan/2004:21:57:34 -0500] "GET
/default.ida?XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX%u9090%u6858%ucbd3%u780
1%u9090%u6858%ucbd3%u7801%u9090%u6858%ucbd3%u7801%u9090%u9090%u8190%u00c3%u0003%u
8b00%u531b%u53ff%u0078%u0000%u00=a HTTP/1.0" 404 285
10.118.25.74 - - [18/Jan/2004:22:02:17 -0500] "GET
/default.ida?XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX%u9090%u6858%ucbd3%u780
1%u9090%u6858%ucbd3%u7801%u9090%u6858%ucbd3%u7801%u9090%u9090%u8190%u00c3%u0003%u
8b00%u531b%u53ff%u0078%u0000%u00=a HTTP/1.0" 404 285
10.30.50.15 - - [18/Jan/2004:22:02:56 -0500] "GET /scripts/root.exe?/c+dir HTTP/1.0" 404 284
10.30.50.15 - - [18/Jan/2004:22:02:56 -0500] "GET /MSADC/root.exe?/c+dir HTTP/1.0" 404 282
10.30.50.15 - - [18/Jan/2004:22:02:57 -0500] "GET /c/winnt/system32/cmd.exe?/c+dir HTTP/1.0" 404 292
10.30.50.15 - - [18/Jan/2004:22:02:57 -0500] "GET /d/winnt/system32/cmd.exe?/c+dir HTTP/1.0" 404 292
10.30.50.15 - - [18/Jan/2004:22:02:57 -0500] "GET /scripts/..%255c../winnt/system32/cmd.exe?/c+dir
HTTP/1.0" 404 306
10.30.50.15 - - [18/Jan/2004:22:02:57 -0500] "GET
/_vti_bin/..%255c../..%255c../..%255c../winnt/system32/cmd.exe?/c+dir HTTP/1.0" 404 323
10.30.50.15 - - [18/Jan/2004:22:02:58 -0500] "GET
/_mem_bin/..%255c../..%255c../..%255c../winnt/system32/cmd.exe?/c+dir HTTP/1.0" 404 323
10.30.50.15 - - [18/Jan/2004:22:02:58 -0500] "GET
/msadc/..%255c../..%255c../..%255c/..%c1%1c../..%c1%1c../..%c1%1c../winnt/system32/cmd.exe?/c+dir
HTTP/1.0" 404 339
10.30.50.15 - - [18/Jan/2004:22:02:58 -0500] "GET /scripts/..%c1%1c../winnt/system32/cmd.exe?/c+dir
HTTP/1.0" 404 305
10.30.50.15 - - [18/Jan/2004:22:02:58 -0500] "GET /scripts/..%c0%2f../winnt/system32/cmd.exe?/c+dir
HTTP/1.0" 404 305
10.30.50.15 - - [18/Jan/2004:22:02:59 -0500] "GET /scripts/..%c0%af../winnt/system32/cmd.exe?/c+dir
HTTP/1.0" 404 305
10.30.50.15 - - [18/Jan/2004:22:02:59 -0500] "GET /scripts/..%c1%9c../winnt/system32/cmd.exe?/c+dir
HTTP/1.0" 404 305
10.30.50.15 - - [18/Jan/2004:22:02:59 -0500] "GET /scripts/..%%35%63../winnt/system32/cmd.exe?/c+dir
HTTP/1.0" 400 296
10.30.50.15 - - [18/Jan/2004:22:02:59 -0500] "GET /scripts/..%%35c../winnt/system32/cmd.exe?/c+dir
HTTP/1.0" 400 296
10.30.50.15 - - [18/Jan/2004:22:03:00 -0500] "GET /scripts/..%25%35%63../winnt/system32/cmd.exe?/c+dir
HTTP/1.0" 404 306
10.30.50.15 - - [18/Jan/2004:22:03:00 -0500] "GET /scripts/..%252f../winnt/system32/cmd.exe?/c+dir
HTTP/1.0" 404 306

Si on analyse les deux entrées suivantes :

- 10.1.4.25 - - [18/Jan/2004:21:57:34 -0500] "GET
  /default.ida?XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
  XXX%u9090%u6858%ucbd3%u7801%u9090%u6858%ucbd3%u7801%u9090%u6858%ucbd3%u780

1%u9090%u9090%u8190%u00c3%u0003%u8b00%u531b%u53ff%u0078%u0000%u00=a HTTP/1.0" **404** 285

- 10.118.25.74 - - [18/Jan/2004:22:02:17 -0500] "GET /default.ida?XXX%u9090%u6858%ucbd3%u7801%u9090%u6858%ucbd3%u7801%u9090%u6858%ucbd3%u7801%u9090%u9090%u8190%u00c3%u0003%u8b00%u531b%u53ff%u0078%u0000%u00=a HTTP/1.0" **404** 285

Il s'agit d'une attaque de débordement de zone tampon (Buffer overflow). C'est une activité virale originaire des systèmes identifiés par les adresses IP 10.1.4.25 et 10.118.25.74. Le virus est un ver dit Code Red II exploitant une vulnérabilité ".ida" du service d'indexation de Microsoft Internet Information Server IIS[14]. La valeur 404 du code d'état indique que l'attaque a échoué.

Pour les entrées suivantes :

10.30.50.15 - - [18/Jan/2004:22:02:56 -0500] "GET /scripts/root.exe?/c+dir HTTP/1.0" **404** 284
10.30.50.15 - - [18/Jan/2004:22:02:56 -0500] "GET /MSADC/root.exe?/c+dir HTTP/1.0" **404** 282
10.30.50.15 - - [18/Jan/2004:22:02:57 -0500] "GET /c/winnt/system32/cmd.exe?/c+dir HTTP/1.0" **404** 292
10.30.50.15 - - [18/Jan/2004:22:02:57 -0500] "GET /d/winnt/system32/cmd.exe?/c+dir HTTP/1.0" **404** 292

10.30.50.15 - - [18/Jan/2004:22:02:57 -0500] "GET /scripts/..%255c../winnt/system32/cmd.exe?/c+dir HTTP/1.0" **404** 306
10.30.50.15 - - [18/Jan/2004:22:02:57 -0500] "GET /_vti_bin/..%255c../..%255c../..%255c../winnt/system32/cmd.exe?/c+dir HTTP/1.0" **404** 323
10.30.50.15 - - [18/Jan/2004:22:02:58 -0500] "GET /_mem_bin/..%255c../..%255c../..%255c../winnt/system32/cmd.exe?/c+dir HTTP/1.0" **404** 323
10.30.50.15 - - [18/Jan/2004:22:02:58 -0500] "GET /msadc/..%255c../..%255c../..%255c/..%c1%1c../..%c1%1c../..%c1%1c../winnt/system32/cmd.exe?/c+dir HTTP/1.0" **404** 339
10.30.50.15 - - [18/Jan/2004:22:02:58 -0500] "GET /scripts/..%c1%1c../winnt/system32/cmd.exe?/c+dir HTTP/1.0" **404** 305
10.30.50.15 - - [18/Jan/2004:22:02:58 -0500] "GET /scripts/..%c0%2f../winnt/system32/cmd.exe?/c+dir HTTP/1.0" **404** 305
10.30.50.15 - - [18/Jan/2004:22:02:59 -0500] "GET /scripts/..%c0%af../winnt/system32/cmd.exe?/c+dir HTTP/1.0" **404** 305
10.30.50.15 - - [18/Jan/2004:22:02:59 -0500] "GET /scripts/..%c1%9c../winnt/system32/cmd.exe?/c+dir HTTP/1.0" **404** 305
10.30.50.15 - - [18/Jan/2004:22:02:59 -0500] "GET /scripts/..%%35%63../winnt/system32/cmd.exe?/c+dir HTTP/1.0" **400** 296
10.30.50.15 - - [18/Jan/2004:22:02:59 -0500] "GET /scripts/..%%35c../winnt/system32/cmd.exe?/c+dir HTTP/1.0" **400** 296
10.30.50.15 - - [18/Jan/2004:22:03:00 -0500] "GET /scripts/..%25%35%63../winnt/system32/cmd.exe?/c+dir HTTP/1.0" **404** 306
10.30.50.15 - - [18/Jan/2004:22:03:00 -0500] "GET /scripts/..%252f../winnt/system32/cmd.exe?/c+dir HTTP/1.0" **404** 306

Nous constatons des tentatives d'accès aux fichiers exécutables de Windows. Ces tentatives prennent 05 secondes ce qui signifie que l'attaque est menée à l'aide d'un script ou d'un programme automatisé.

Les quatre premières entrées indiquent des tentatives de connexion à la porte dérobée laissée par le ver Code Red II, la suite des entrées consiste en des tentatives d'exploitation de « Directory Traversal vulnerability ». Cette vulnérabilité permet à des attaquants visitant un

site Web IIS d'exécuter des commandes arbitraires avec le privilège « IUSR_nom machine » [14].

Il s'agit d'une activité virale du ver NIMDA exploitant une vulnérabilité dans Microsoft Internet Explorer initiée par le système identifié par l'adresse IP 10.30.50.15. Les valeurs 400 et 404 du code d'état indiquent l'échec de l'attaque.

12. Les problèmes liés aux fichiers logs

Malgré l'importance des fichiers logs, néamoins certains problèmes demeurent posés. Nous citons :

- Les fichiers logs consomment un espace disque très grand.
- Les fichiers logs contiennent beaucoup d'information, ils sont immenses et par conséquent l'analyse de leurs contenus devient une tâche très difficile.
- Les fichiers logs menacent la vie privée(privacy) de l'utilisateur. Un utilisateur refuse l'idée que toutes ses activités soient enregistrées.
- Les fichiers logs peuvent être menacés comme d'autres formes de données dans le réseau ou dans un système. Un attaquant qualifié pénétrant dans un système peut effacer les fichiers logs ou modifier leur contenu. Il peut même arrêter le mécanisme d'enregistrement.

13. Conclusions et recommandations

La sécurité informatique a besoin d'un système de surveillance et d'un système judiciaire. Les fichiers logs représentent le moyen de surveillance de la sécurité des systèmes et des réseaux. « Computer Network Forensics » étant le système judiciaire auquel la sécurité informatique a besoin.

Les fichiers logs constituent une source critique de preuve pour Forensics, ils peuvent contenir les empreintes des attaquants et indiquer les menaces et les attaques en cours. Ils représentent une forme de données qui peut-être effacée ou falsifiée par un attaquant afin d'effacer la trace de l'attaque ou modifier le cours de l'attaque. Si les fichiers logs sont effacés ou erronés, on perd toute preuve d'attaque et par conséquent le processus de Forenscis ne peut réussir. Des mesures de protection doivent être prises en compte vis à vis des fichiers logs. La politique de sécurité d'une organisation doit inclure les procédures de protection des fichiers logs des composants d'un réseau.

L'inspection et la révision des fichiers logs sont la meilleure façon pour maintenir la sécurité des systèmes et des réseaux. Un administrateur de sécurité est responsable de la protection du patrimoine informationnel. En l'occurrence d'une attaque, il doit faire une investigation en interprétant les données liées à l'attaque contenues dans les fichiers logs. Vu la nature ASCII du format des fichiers logs, il est très délicat de déchiffrer leur contenu et déterminer les étapes d'une attaque. L'information pertinente contenue dans les fichiers logs nécessite d'être

interprétée avec précaution pour accomplir une investigation et réussir le processus de Forensics.

Références

[1] The most dangerous place on earth in 2001 : A cyber ethnography
Terry M. Gudaitis
The Magazine on Information Impacts
May 2001

[2] FC : Bruce schneier on computer security :"Things are getting worse"
Bruce schneier
July 16, 2001
Testimony and Statement for the United States Senate

[3] Fingerprints, Broken Windows…Code everywhere
David Britt
July 2001
http://www.toplayer.com/content/cm/news66.jsp

[4] Policies to Enhance Computer and Network Forensics
Alec Yasinsac and Ynet Manzano
Workshop on information assurance and security
United states Military Academy
West point, NY, 5-6 June, 2001

[5] When security fails
By Paul Desmond
Network World
September 2000

[6] Network Forensics
Christopher Patrick Murray
University of Minnesota, Morris
http://web.archive.org/web/20040215100928/http://mrs.umn.edu/~lopezdr/seminar/fall2000/Murray.htm

[7] Techniques for catching the 'perp': the basics of computer forensics
By Don Walker
May, 2001
Unisys World

[8] About logging site activity
http://iishelp.web.cern.ch/IISHelp/iis/htm/core/iiabtlg.htm

[9] L'utilisation des "log files" en évaluation et en reconception : intérêts et limites
Christian Bastien.
2001
http://www.lergonome.org/pages/detail_articles.php?indice=9

[10] Downloads, Logs and Captures :
Evidence from cyberspace
Peter Sommer
Computer Security Research Center
London School of Economics & Political Science
http://www.giustizia.it/cassazione/convegni/dic2000/sommer_2.pdf

[11] Digital Evidence and Computer Crime:
Forensic Science, Computers, and the Internet
Eoghan Casey
Edition Academic Press, January
Janvier 2000

[12] La conception d'une base d'une base de connaissance pour l'aide à l'investigation dans Firewall Forensics
Bensefia Hassina
Mémoire de post-graduation spécialisée en sécurité Informatique
Mai 2002
Centre de Recheche sur l'Information Scientifique et technique
Alger, Algérie

[13] Forensics For Dummies
Douglas. P. Lyle
Edition Paperback
2004

[14] www.cert.org
CERT Coordination Center