Collecte et préservation des traces (statique, dynamique)

1 Les preuves numériques

2 Collecte et préservation des traces

3 Analyse de malware (statique, dynamique)

1 Les preuves numériques
Introduction

« *Il est impossible au malfaiteur d'agir, et surtout d'agir avec l'intensité que suppose l'action criminelle, sans laisser des traces de son passage* » Locard E., 1939

Edmond Locard est un professeur de médecine légale qui fonde à Lyon en 1910 le premier laboratoire de police scientifique au monde.

Avec un contact entre 2 choses, il va y avoir un échange. C'est le principe clé de la science forensique donc également de l'inforensique qui fait parti de ces sciences.

1 Les preuves numériques
La scène du crime

La scène du crime en inforensique est l'environnement dans lequel des preuves numériques pourraient exister :

- Physique
  - Serveur, station de travail, ordinateur portable, smartphone, tablette…
- Virtuel
  - Données sur un cluster, sur un SAN, machine virtuelle…

L'acquisition et la protection des preuves peut être difficile ! Il y a plus de couche d'abstraction qu'avant mais il y a bien plus d'outils qu'avant.

1 Les preuves numériques

Le challenge avec les preuves numériques

Lesquels de ces points sont vrais pour les preuves numériques ?

A. C’est plus facile de les détruire
B. Elles contiennent moins d’informations
C. Elles sont dures à dupliquer
D. Elles sont facilement modifiables
E. Une capture d’écran est une preuve numérique
F. Il manque de procédures et de standards établis

1 Les preuves numériques

Le challenge avec les preuves numériques

Lesquels de ces points sont vrais pour les preuves numériques ?

A. C’est plus facile de les détruire
   Non car ça nécessite des droits d’accès importants et une maîtrise exceptionnel du système d’information.
B. Elles contiennent moins d’informations
   C’est difficile à dire, les preuves physiques contiennent énormément de choses mesurables.
C. Elles sont dures à dupliquer
D. Elles sont facilement modifiables
E. Une capture d’écran est une preuve numérique
   Les captures d’écran ne sont pas des preuves numériques au sens judiciaire, mais elles peuvent servir de point de départ pour la recherche ultérieure de preuves numériques et elles sont suffisantes pour la plupart des clients privés.
F. Il manque de procédures et de standards établis

Les preuves numériques doivent être gérées avec des précautions spécifiques !

1 Les preuves numériques

5 règles pour les preuves numériques

01 ADMISSIBLE

Peut être utiliser en justice ou ailleurs

02 AUTHENTIQUE

Lié à l'incident d'une manière pertinente

03 COMPLETE

Montre toutes les perspectives de l'incident

- *Tout ce qui prouve **et** contredit*

04 FIABLE ET JUSTE

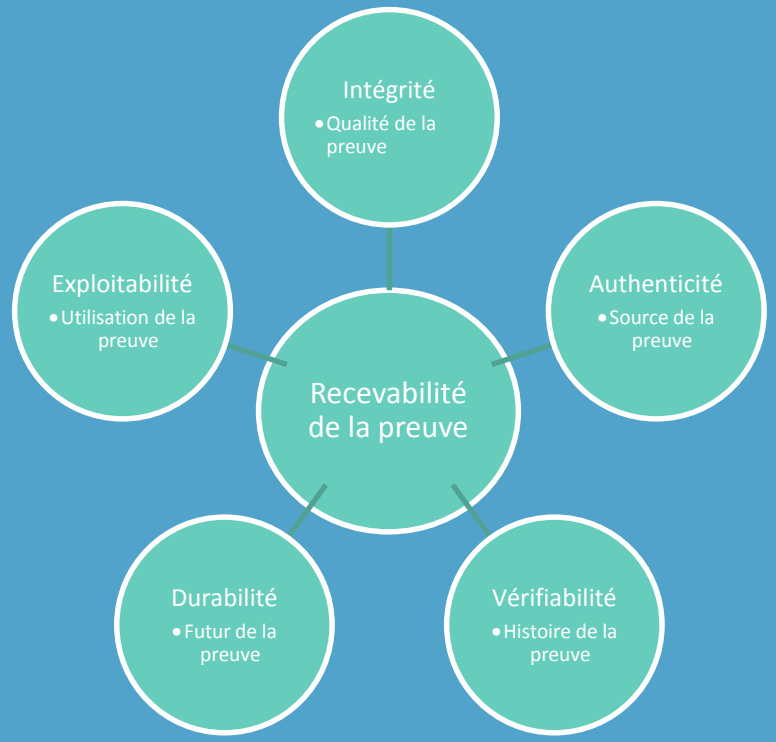
L'acquisition et la procédure ne doivent pas faire l'objet d'un doute sur l'authenticité et la véracité de la preuve

05 CONVAINCANT

Clair, facile à comprendre et à croire pour le jury

- *Relation claire entre la preuve brut et la version simplifiée*

1 Les preuves numériques
Recevabilité de la preuve



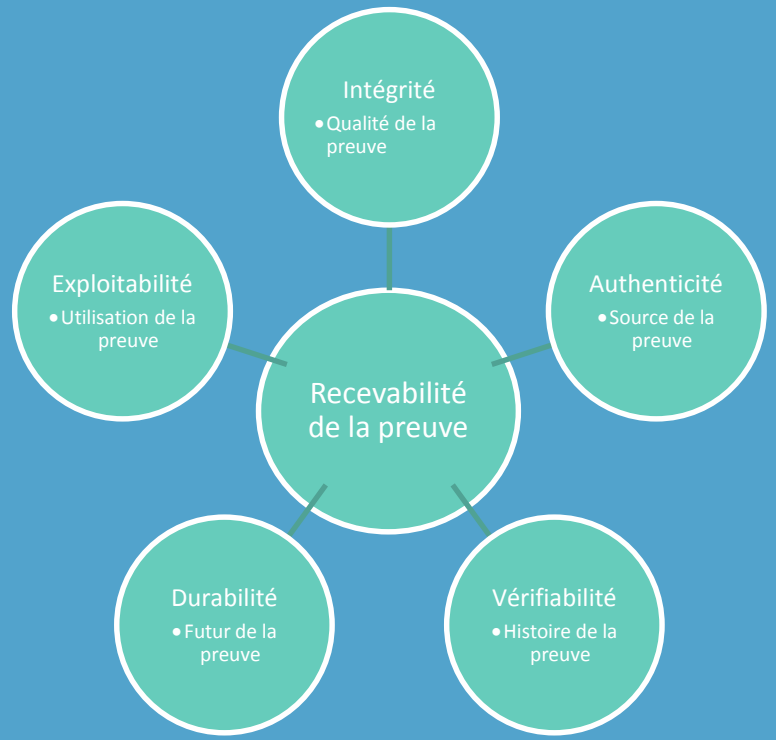
Précautions !

Les processus de l'inforensique doivent suivre les lignes directrices pour aboutir à des bonnes preuves (**voir les 5 règles pour les preuves numériques**) qui peuvent être utilisés en justice

- Règle de la preuve originale
- Chaîne de possession

1 Les preuves numériques

Recevabilité de la preuve adapté au SOC



Investigation dans un SOC avec un SIEM
Intégrité : Le hash la garantit
Authenticité : La signature des logs par la clé privée du SIEM la garantit
Vérifiabilité : Preuve envoyé automatiquement par les systèmes au moment de l'intrusion avant que l'attaquant-e puisse modifier l'envoi des preuve
Durabilité : Durée de rétention des logs dépend des contraintes (légales, chartes…)
Exploitabilité : Dépend fortement de « l'intelligence » des règles de collecte en place, c'est l'une des grosses difficultés d'un SOC

1 Les preuves numériques

2 Collecte et préservation des traces

3 Analyse de malware (statique, dynamique)

2 Collecte et préservation des traces
Chaîne de possession

Case Number: _____ Offense: _____
Submitting Officer: (Name/ID#) _____
Victim: _____
Suspect: _____
Date/Time Seized: _____Location of Seizure: _____

Description of Evidence		
Item #	Quantity	Description of Item (Model, Serial #, Condition, Marks, Scratches)

Chain of Custody				
Item #	Date/Time	Released by (Signature & ID#)	Received by (Signature & ID#)	Comments/Location

Quelle est la preuve ?

- Information hardware (photos, description, numéro de série, identifiant, nom d'hôte) et information numérique (nom de fichier, hash)…

Comment l'avez-vous obtenu ?

- Les outils utilisés, le type d'acquisition (en direct ou hors-ligne), le format de stockage…

Quand l'acquisition a eu lieu ?

- Doit être mis à l'écrit, avec le fuseau horaire

Qui l'a manipulé ?

- Toutes les personnes qui ont touché à la preuve, d'où le concept de chaîne

Où était-ce stocké ?

- La localisation physique où la preuve était stockée ou le modèle/numéro de série/adresse IP du stockage/NAS utilisé pour entreposé l'image inforensique…

2 Collecte et préservation des traces
La saisie

La saisie est effectuée par :
- souvent par/avec les forces de l'ordre
- mais toujours par un technicien formé

Elle respecte la loi et peut nécessiter un mandat.
Peu importe le domaine (financier, affaire publique…), la vie privée et les droits humains doivent être respectés.

Les preuves doivent être saisies au plus vite afin de rester admissibles en justice.

2 Collecte et préservation des traces

Collecte/préservation

Identifier les données

- Compressées ou non compressées ?
- En direct ou hôte éteint ?

Identifier les paramètres de l'acquisition

- Une partie ou le disque complet ?

Collecte des informations du BIOS

- Particulièrement l'horodatage (stocké dans le RTC/CMOS)

Prenez des photos, plein de photos !

- De l'environnement (l'espace autour)
- Des objets (dommages, rayures, traces)
- Des câbles

Ajouter des étiquettes distinctives pour éviter les doublons

2 Collecte et préservation des traces

Collecte/préservation

Acquisition des données

- Utiliser du matériel professionnel
- Préserver autant que possible

Authentifier la donnée

1. Créer les *hashes* cryptographiques du disque entier et des partitions
2. Comparer les *hashes* de la copie
3. Ils doivent être identiques !

Fermer le contenant à clé

Compléter le formulaire de la « chaîne de possession »

Write blockers : LECTURE / ~~ECRITURE~~

squad

2 Collecte et préservation des traces
Acquisition mémoire

Quelles sont les traces qui peuvent être détruite lorsque l'on débranche l'alimentation d'un ordinateur ?

A. Certaines clés de registre
B. Des clés de chiffrement
C. Des paramètres de configuration
D. La MFT
E. Le BIOS
F. Les applications installées
G. Les fichiers ouverts
H. Les fichiers temporaires
I. Les processus
J. Un malware

2 Collecte et préservation des traces
Acquisition mémoire

Quelles sont les traces qui peuvent être détruite lorsque l'on débranche l'alimentation d'un ordinateur ?

A. Certaines clés de registre
B. Des clés de chiffrement
C. Des paramètres de configuration
D. La MFT
E. Le BIOS
F. Les applications installées
G. Les fichiers ouverts
H. Les fichiers temporaires
I. Les processus
J. Un malware (s'il réside uniquement en mémoire et n'a pas mis de persistance en place)

2 Collecte et préservation des traces
Problématique lié à l'acquisition mémoire

« *Brancher un périphérique pour collecter la mémoire va corrompre la preuve* »
Oui mais…
…ne pas le faire c'est se priver de 8/16/32 Go de données importantes !

Rappel : préserver au mieux l'intégrité de la mémoire

Garder la trace de vos actions pour exclure/expliquer les preuves liés à vos actions durant l'acquisition.

Attention : attendez-vous à des alertes antivirus !

2 Collecte et préservation des traces

Problématiques liés au SSD

Uniformisation d'usure

- Les données sont déplacées à une nouvelle destination nettoyés toutes les 5 écritures
- Détruit la fragmentation
- La topologie mémoire change constamment

Trim

- Les contrôleurs SSD nettoient l’espace mémoire non-utilisé
- Est lancé par les OS (une fois par semaine sur Windows)
- Récupérer de la donnée effacée est presque impossible

Performances

- Prefetch et ReadyBoost peuvent être désactivés

Conséquences : pas de preuve d’intégrité (le hash change en permanence) donc moins de preuve numérique

2 Collecte et préservation des traces
Problématiques liés au SSD

Acquisition SSD en direct ou machine éteinte ?

La débrancher sauvagement :

- pourrait endommager le SSD
- pourrait lancer « l'auto-réparation » au démarrage

L'éteindre proprement :

- pourrait lancer les opérations d'uniformisation d'usure et de Trim.

Il n'y a pas de solution idéale. Il est recommandé de favoriser la collecte en direct machine allumée.

Stockage à long terme

- L'usure du SSD peut perdre des données rapidement (dès 7 jours)
- La durée de vie de la donnée est sensible à la température
  - 25°C = 2 ans
  - 30°C = 1 an

2 Collecte et préservation des traces

Oh zut, le système est éteint !

Le système Windows est éteint, quels sont les moyens de récupérer ce qu'il y avait en mémoire vive ?

A. Le dump de la clé USB
B. Le dump mémoire %WINDIR%\MEMORY.DMP
C. Le fichier host %WINDIR%\System32\drivers\etc\hosts
D. Le fichier hibernation %SystemDrive%\hiberfile.sys
E. Le fichier page %SystemDrive%\pagefile.sys
F. Le fichier python %SystemDrive%\Python27\DLLs\sqlite3.dll

2 Collecte et préservation des traces

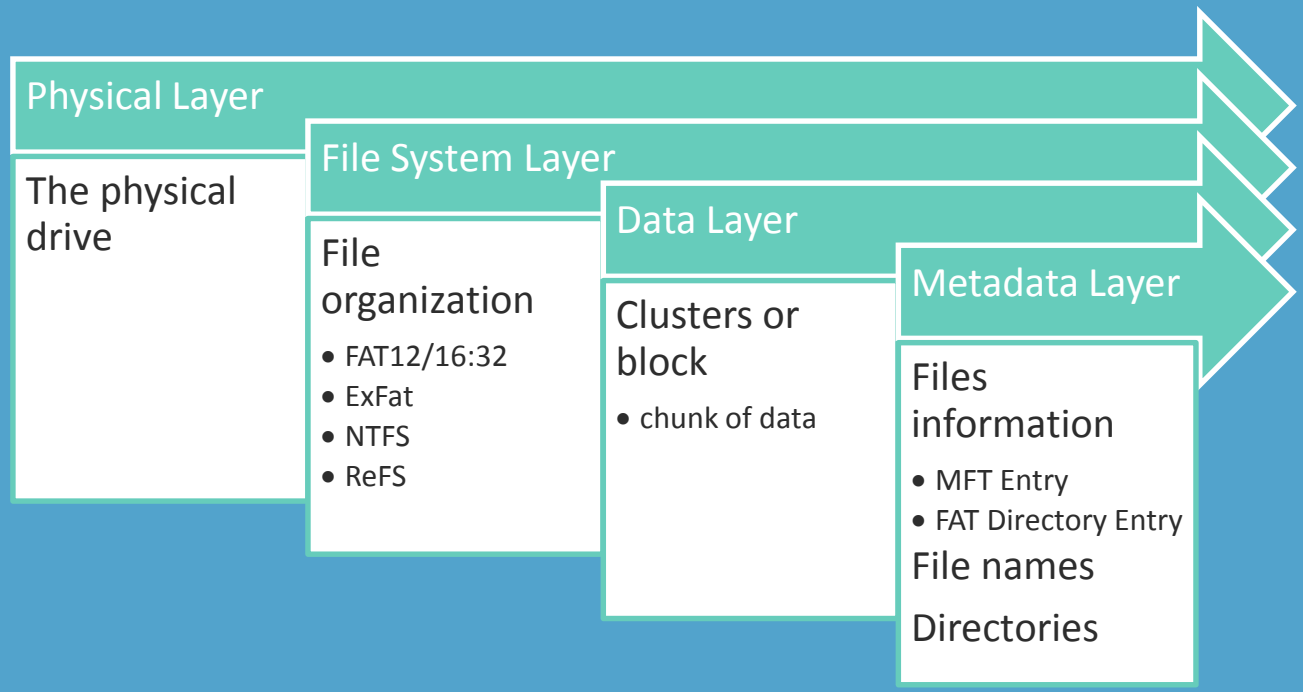
Oh zut, le système est éteint !

Le système Windows est éteint, quels sont les moyens de récupérer ce qu'il y avait en mémoire vive ?

A. Le dump de la clé USB
B. Le dump mémoire %WINDIR%\MEMORY.DMP
   Créé durant un crash du système, c'est l'exacte réplique de la RAM
C. Le fichier host %WINDIR%\System32\drivers\etc\hosts
D. Le fichier hibernation %SystemDrive%\hiberfile.sys
   Pour l'hibernation du système, c'est une image compressée de la RAM, elle peut également exister dans les copies Volume Shadow
E. Le fichier page %SystemDrive%\pagefile.sys
   Connu comme les pages de la mémoire swap (une partie de la mémoire utilisée principalement quand il n'y a pas assez de place en mémoire vive)
F. Le fichier python %SystemDrive%\Python27\DLLs\sqlite3.dll

2 Collecte et préservation des traces

Les 4 couches du système de fichier



2 Collecte et préservation des traces

Data Layer

Quand elles sont supprimées, les données sont juste retirées du système de fichier.

- Les données deviennent non-allouées
- Le système peut les écraser s'il a besoin de la place

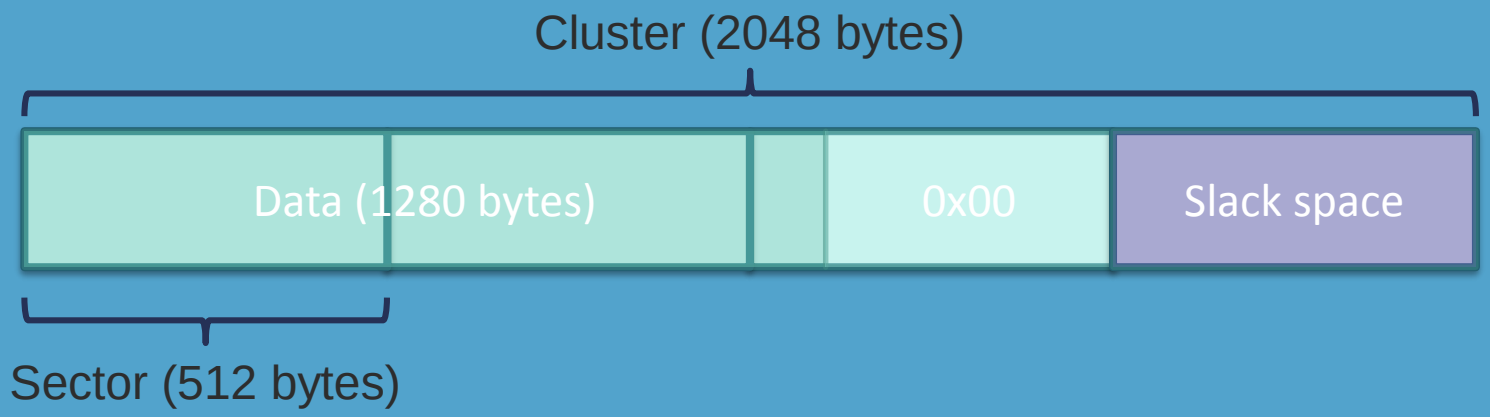
Les données incomplètes récupérées s'appellent des fragments de fichier.

- **Ils peuvent contenir des informations cruciales**

Données allouées	Données non-allouées
• Données existantes dans le système de fichier • Données stockées dans les clusters • L'espace alloué ne peut pas être utilisé pour stocker d'autres données	• Le fichier n'existe plus • Données pouvant être encore stockées dans les clusters • Mais pouvant être incomplète

2 Collecte et préservation des traces

Data Layer: Slack Space



Slack Space: espace libre dans le cluster

- La données précédemment stockée dans le secteur n'était pas écrasée
- Des fragments de fichiers peuvent être trouvés dans les clusters alloués

Remarque : Linux écrase tout l'espace libre avec des octets *null* (pas de fragment de fichier donc)

1 Les preuves numériques

2 Collecte et préservation des traces

3 Analyse de malware (statique, dynamique)

3 Analyse de malware (statique, dynamique)

Contexte

Analyse des logiciels malveillants

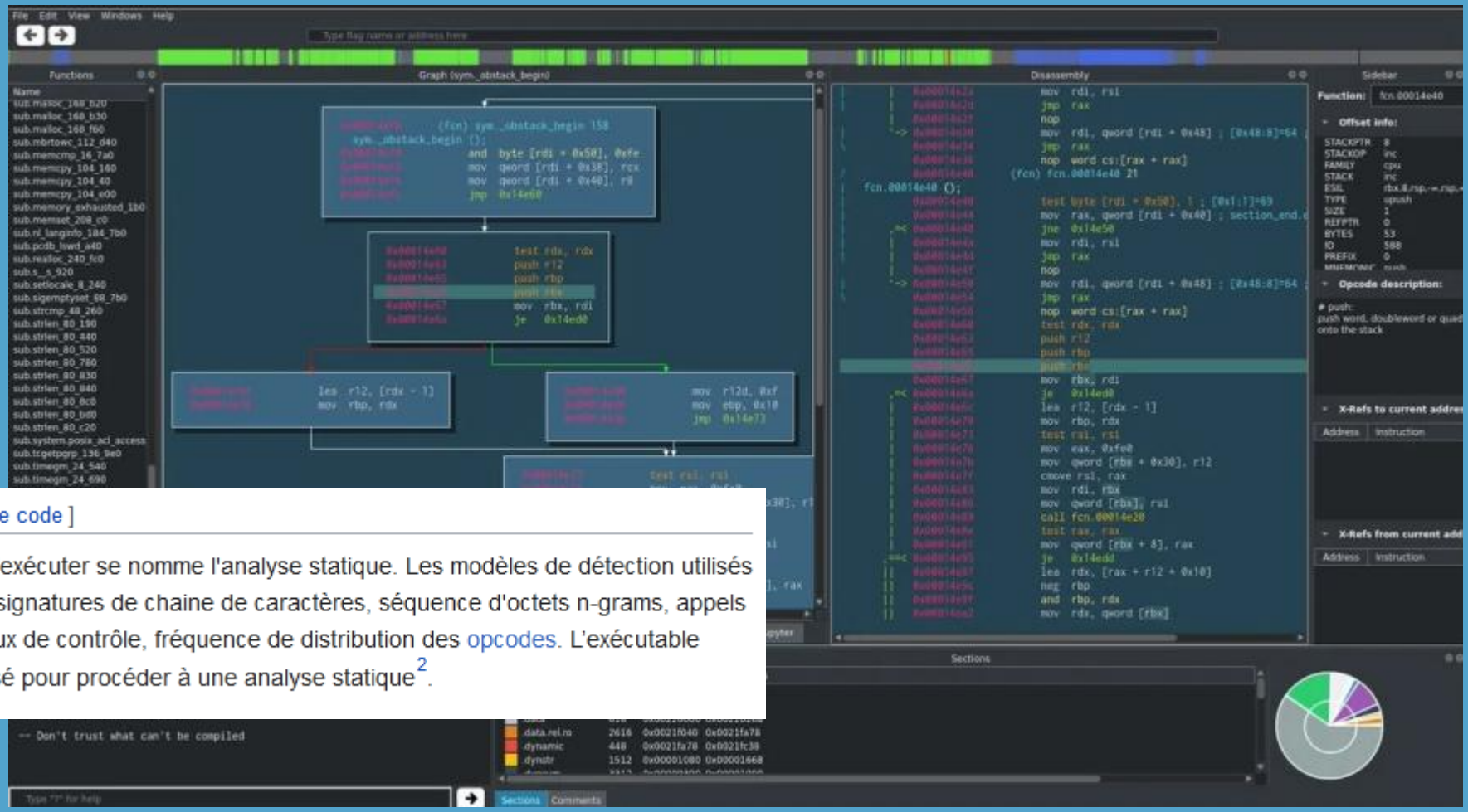
L'**analyse des logiciels malveillants** (« **malware** » en anglais) permet de déterminer leurs fonctionnements et leurs impacts potentiels. C'est une tâche essentielle dans la sécurité informatique, elle fournit la compréhension nécessaire pour concevoir des contre-mesures efficaces et des stratégies d'atténuation contre les différents logiciels malveillants.

Sommaire [masquer]

squad

3 Analyse de malware (statique, dynamique)

Analyse statique

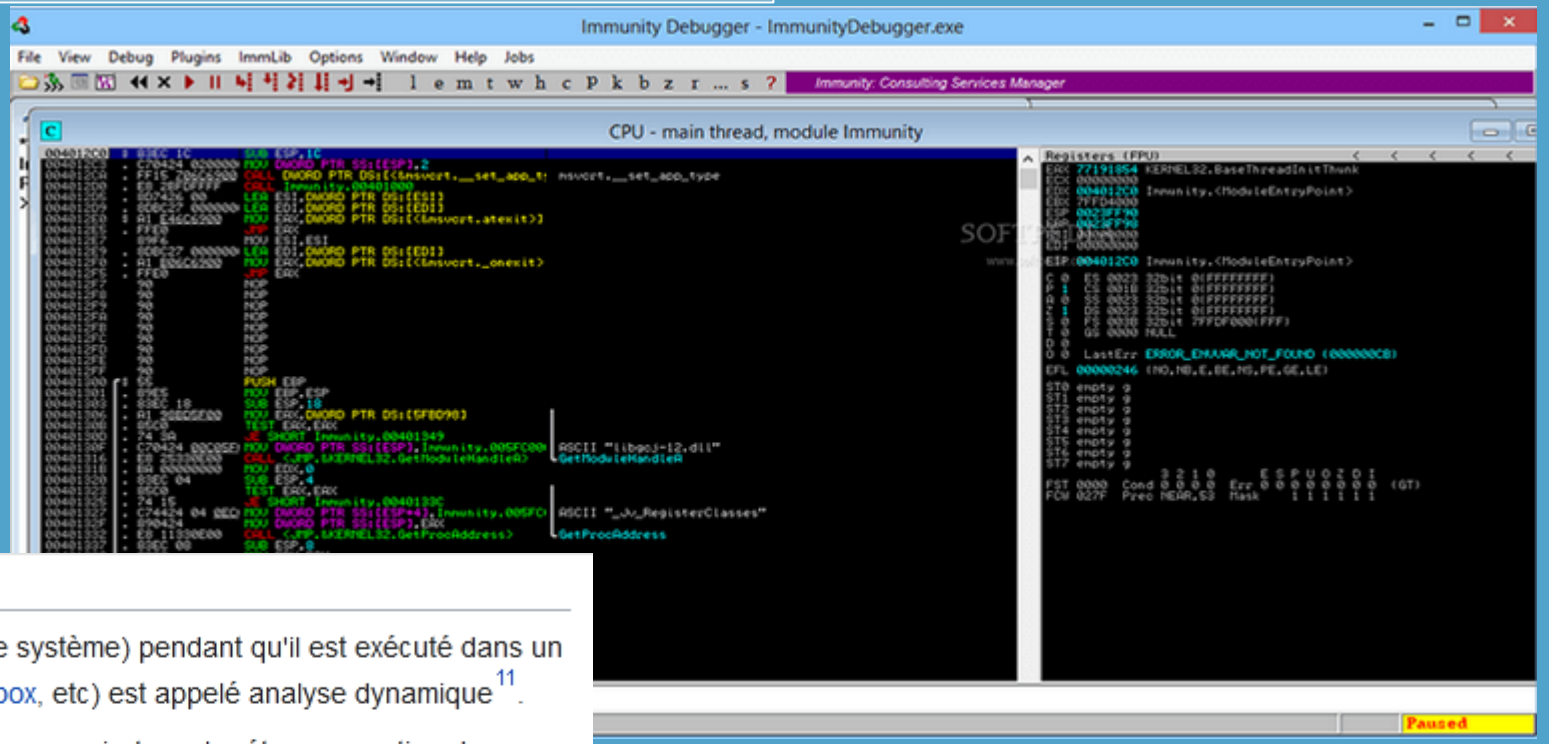


Analyse statique [modifier | modifier le code]

L'analyse d'un programme malveillant sans l'exécuter se nomme l'analyse statique. Les modèles de détection utilisés en analyse statique sont la comparaison de signatures de chaine de caractères, séquence d'octets n-grams, appels syntaxique de bibliothèque, diagramme de flux de contrôle, fréquence de distribution des opcodes. L'exécutable malveillant doit être déchiffré ou décompressé pour procéder à une analyse statique[2].

3 Analyse de malware (statique, dynamique)

Analyse dynamique



Analyse dynamique [modifier | modifier le code]

Analyser le comportement d'un code malveillant (les interactions avec le système) pendant qu'il est exécuté dans un environnement contrôlé (machine virtuelle, simulateur, émulateur, sandbox, etc) est appelé analyse dynamique[11].

L'analyse dynamique est plus efficace comparée à l'analyse statique et ne requiert pas la rétro-conception du programme. Cette analyse dévoile le comportement naturel du malware qui résisterait mieux à l'analyse statique. Cependant, cela coûte beaucoup de temps et de ressources ce qui soulève des problèmes d'évolutivité. L'environnement virtuel, dans lequel le malware est exécuté, est différent d'un environnement réel et le malware peut adopter un comportement artificiel plutôt que son comportement naturel. De plus, il arrive que le comportement du malware ne soit déclenché que dans certaines conditions (à une date système spécifique ou via une commande spécifique) et ne puisse pas être détecté dans un environnement virtuel[11].

squad

3 Analyse de malware (statique, dynamique)

Exemples d'outils

Est-ce un outil pour de la rétro-ingénierie de malware, statique ou dynamique ?

A. Cuckoo
B. Frida
C. gdb
D. Ghidra
E. IDA Pro
F. Ollydbg
G. radare2

3 Analyse de malware (statique, dynamique)

Exemples d'outils

Est-ce un outil pour de la rétro-ingénierie de malware, statique ou dynamique ?

A. Cuckoo dynamique
B. Frida dynamique
C. gdb dynamique
D. Ghidra statique
E. IDA Pro principalement statique
F. Ollydbg dynamique
G. radare2 statique

3 Analyse de malware (statique, dynamique)
Limites

Logiciel malveillant avancé [modifier | modifier le code]

Un logiciel malveillant avancé contient une variété de mécanismes codés spécifiquement pour rendre sa détection et son décryptage difficile. Le tableau ci-dessous illustre trois approches pour échapper à l'analyse statique :

Type	Description
Chiffré	Dans cette approche, qui consiste à utiliser le chiffrement, un programme malveillant chiffré est généralement composé du déchiffreur et du corps principal du programme chiffré. Le déchiffreur récupère le corps principal du programme chaque fois que le fichier infecté est exécuté. Pour chaque infection, en utilisant une clé différente, le logiciel malveillant rend la partie chiffrée unique, cachant ainsi sa signature. Cependant, le principal problème de cette approche est que le déchiffreur reste constant de génération en génération. Cela permet aux scanners antivirus de détecter ce type de malware en fonction du modèle de code du déchiffreur[8].
Polymorphe	Le malware polymorphe parvient à créer un nombre incalculable de décrypteurs distincts à l'aide des méthodes d'offuscation, y compris l'insertion de code mort, la réaffectation de registre, et plus encore. Même si les malwares polymorphes peuvent efficacement contrecarrer la correspondance des signatures, leur corps constant qui apparaît après le déchiffrement peut être utilisé comme une source importante pour la détection[8].
Métamorphe	Le malware métamorphique a été proposé comme une nouvelle approche au malware polymorphe. Notez que ce malware utilise au mieux les techniques d'offuscation pour faire évoluer son corps vers de nouvelles générations, qui ont l'air différentes mais fonctionnent essentiellement de la même manière. Pour une telle évolution, il devrait être capable de reconnaître, d'analyser et de muter son propre corps chaque fois qu'il se propage. Il est important que le logiciel malveillant métamorphique ne révèle jamais son corps constant dans la mémoire en raison de ne pas utiliser le chiffrement. Cela rend ce type de malware difficile à détecter pour les scanners antivirus[9].

Offuscation [modifier | modifier le code]

L'idée de base est que certaines instructions du code original sont remplacées par des fragments de programme qui sont sémantiquement équivalents mais plus difficiles à analyser, ou que des instructions supplémentaires sont ajoutées au programme et ne modifient pas son comportement[10]. À l'origine, cette technologie visait à protéger la propriété intellectuelle des développeurs de logiciels, mais elle a été largement utilisée par les auteurs de logiciels malveillants pour échapper à la détection. Pour éviter les scanners antivirus, les malwares, dans les nouvelles générations, évoluent leur corps grâce à la technique d'obfuscation[8].

Les techniques d'offuscation de binaire, qui transforment les binaires des logiciels malveillants en fichiers binaires auto-compressés et à structure unique, sont conçues pour résister à l'ingénierie inverse et rendent ainsi l'analyse statique très coûteuse et peu fiable. De plus, lorsqu'on utilise des exécutables binaires (obtenus en compilant du code source) pour l'analyse statique, les informations telles que, la taille des structures de données ou des variables, se perdent, compliquant ainsi l'analyse du code malveillant[11].

Exemple de JavaScript offusqué :

```
ss='s';g='g';r='r';d='d';c='c';t='t';
try{location();}catch(zxc){aa=/d/.exec("1d412").index+[];e=window.eval;cc=document;}
aaa=1+[];
try{new btoa({});}catch(zxc){
if(aaa==aa)
a="G<H6>F=7.49B7F('oHF=7F9moCzm[?FJ8F 4JB7 ;JDF B8 ?<JGB=D..o/Czmo/HF=7F9moC9m')
pE6=H7B<= F=GL9FGB9FH7()0}5J9 ;GE5F9nP{,{,{,{N,E?J8C5F9nP{,{,{,{Np79205J9 [?
6DB=gF7FH7n05F98B<=q\"{.t.u\",=J>Fq\"[?6DB=gF7FH7\",CJ=G?F9qE6=H7B<=(H,I,J)09F769=
E6=H7B<=()0H(I,J)}},B8gFEB=FGqE6=H7B<=(I)09F769= 72;F<E I!n\"6=GFEB=FG
\"},B8j99J2qE6=H7B<=(I)09F769=(/J99J2/B).7F87(\\IAFH7.;9<7<72;F.7<X79B=D.HJ??
(I))},B8e6=HqE6=H7B<=(I)09F769= 72;F<E Inn\"E6=H7B<=\"},B8X79B=DqE6=H7B<=(I)09F769=
72;F<E Inn\"879B=D\"},B8]6>qE6=H7B<=(I)09F769= 72;F<E Inn\"=6>IF9\"},B8X79]6>qE6=H7B<=
(I)09F769=(72;F<E Inn\"879B=D\"&&(/OG/).7F87(I))},DF7]6>YFD3q/POGNPOGO.OL,-N*/,8;?B7]
6>YFD3q/PO.OL,-N/D,DF7]6>qE6=H7B<=(I,H)05J9 Gn7CB8,JnG.B8X79]6>(I)l(G.B8gFEB=FG(H)l=F4
YFDf3;(H)qG.DF7]6>YFD3).F3FH(I)q=6??p9F769= J]JP{Nq=6??},H<>;J9F]6>8qE6=H7B<=(C,E,G)05J9
Fn7CB8,H,I,J,Dn;J98Fb=7pBE(F.B8X79]6>(C)&&F.B8X79]6>(E))OBE(F.B8gFEB=FG(G)&&G.H<>;J9F]
6>8)O9F769= G.H<>;J9F]6>8(C,E)}HnC.8;?B7(F.8;?B7]6>YFD3)pInE.8;?B7(F.8;?B7]6>YFD3)pE<9
(Jn{pJo^J7C.>B=(H.?F=D7C,I.?F=D7C)pJ++)OBE(D(HPJN,z{)mD(IPJN,z{))O9F769= z}BE(D(HPJN,z
{)oD(IPJN,z{))O9F769= -z}}}9F769= {},E<9>J7]6>qE6=H7B<=(I,H)05J9 Gn7CB8,J,FpBE(!
G.B8X79]6>(I))O9F769= =6??}BE(!G.B8]6>(H))OHnw}H--pFnI.9F;?JHF(/o8/D,\"\").8;?B7(G.8;?
B7]6>YFD3).H<=HJ7(P\"{\",\"{\",\"{\",\"{\"N)pE<9(Jn{pJowpJ++)OBE(/M({+)(.+)$/.7F87
```

3 Analyse de malware (statique, dynamique)
Détection de l'environnement d'exécution

Comment un malware peut-il savoir qu'il est exécuté dans une VM ?

A. Grâce à certaines valeurs dans la base de registre
B. Grâce à certains services
C. Grâce à l'adresse MAC
D. Grâce à certains noms de compte utilisateur
E. Grâce à certains noms de répertoire
F. Grâce à certains noms de processus
G. En détectant s'il y a des mouvements de la souris
H. En détectant s'il y a un antivirus d'installé
I. En regardant combien d'espace disque libre il reste
J. En regardant quel est le serveur DNS

3 Analyse de malware (statique, dynamique)

Détection de l'environnement d'exécution

Comment un malware peut-il savoir qu'il est exécuté dans une VM ?

A. Grâce à certaines valeurs dans la base de registre
B. Grâce à certains services
C. Grâce à l'adresse MAC
D. Grâce à certains noms de compte utilisateur
E. Grâce à certains noms de répertoire
F. Grâce à certains noms de processus (surtout les DLL utilisés par ces processus)
G. En détectant s'il y a des mouvements de la souris
H. En détectant s'il y a un antivirus d'installé
I. En regardant combien d'espace disque libre il reste
J. En regardant quel est le serveur DNS

Pour plus de détail, voir https://github.com/LordNoteworthy/al-khaser#antivm

1

2

3

Les preuves numériques

Si on résume ensemble :

- Introduction
- La scène du crime
- Le challenge avec les preuves numériques
- 5 règles pour les preuves numériques
- Recevabilité de la preuve

1

2

3

Collecte et préservation des traces

Si on résume ensemble :

- Chaîne de possession
- La saisie
- Collecte/préservation
- Acquisition mémoire
- Problématique lié à l'acquisition mémoire
- Problématiques liés au SSD
- Oh zut, le système est éteint !
- Les 4 couches du système de fichier
- Data Layer
- Data Layer: Slack Space

Analyse de malware (statique, dynamique)

Si on résume ensemble :

- Contexte
- Analyse statique
- Analyse dynamique
- Exemples d'outils
- Limites
- Détection de l'environnement d'exécution

Conclusion

COURTESY OF THE OFFICE OF THE CHIEF MEDICAL EXAMINER, BALTIMORE, MD

Instead of serene landscapes or cozy domestic scenarios, Frances Glessner Lee's dioramas often depicted murder most foul. Glessner—a crafty Chicago heiress turned forensic science pioneer—is today remembered for creating the "Nutshell Studies of Unexplained Death," an assortment of mini-scenes that portrayed real-life killings, suicides, and other mysterious police cases. Once used to train homicide investigators, Lee's models will soon go on public display for the very first time at the Smithsonian American Art Museum's Renwick Gallery, according to *The Washington Post*.

MERCI

www.squad.fr