Pivoter tel Bernard, ou comment monitorer des attaquants négligents

Daniel Lunghi
daniel_lunghi@trendmicro.com

Trend Micro

Résumé. Dans ce papier, nous allons présenter une investigation complète et sur le long terme d'un groupe d'attaquants ayant effectué des attaques ciblées.
L'idée est de montrer les différentes étapes qui ont amené à la publication d'un rapport complet sur une opération donnée.
L'investigation commence par l'analyse de souches malveillantes, ce qui permet d'en tirer des informations et des indicateurs qui pourront être exploités, via différentes méthodes qui sont expliquées dans le papier, pour obtenir des informations supplémentaires.
À l'arrivée, on obtient la liste des familles de malwares utilisées, une cartographie partielle de l'infrastructure de l'attaquant, une liste de victimes ainsi que de certains outils de post-exploitation.
Cette investigation s'appuie sur un cas réel, dont les résultats ont été publiés sur le blog de Trend Micro [7, 8].

1 DRBControl, un groupe qui aime les paris

Cette étude porte sur un groupe qui cible des entreprises de paris et de jeux en ligne en Asie du Sud-Est que nous avons analysé avec trois collègues de Trend Micro. Les premières attaques identifiées datent de juillet 2019, et les plus récentes ont été constatées en mars 2020. Pour information, un papier [8] analysant une des campagnes de ce groupe a été publiée sur le blog [7] de Trend Micro. Cependant, l'article en question ne s'attarde pas sur la méthodologie utilisée pour obtenir les données présentées, et en ce sens, il ne fait pas doublon avec cet article.

Le point de départ de cette investigation a été une proposition de collaboration de l'entreprise Talent-Jump Technologies [19] suite à une mission de réponse à incident qu'elle a effectuée au sein d'une entreprise philippine de pari en ligne.

2 Analyse des souches

L'investigation a commencé en juillet 2019. Nous avons reçu entre 15 et 20 souches sur une période d'un mois, au fur et à mesure des trouvailles

de l'équipe de réponse à incidents, puis encore deux souches début octobre 2019.

La première étape a été d'analyser les souches à l'aide de notre outil de rétro-ingénierie préféré. Les objectifs d'une telle analyse, au-delà de déterminer les fonctionnalités du malware, sont principalement de récupérer l'adresse du ou des serveurs de contrôle, puis de déterminer s'il s'agit d'une famille de malware connue. Dans le cas idéal, on tombe sur une famille de malware qui n'est utilisée que par un groupe d'attaquants, ce qui simplifie l'attribution, et permet d'emblée de s'appuyer sur les investigations précédentes. Dans notre cas, nous avons pu diviser nos souches en 4 familles :

— Type 1 : 11 souches
— Type 2 : 3 souches
— Type 3 : 5 souches
— HyperBro : 1 souche, arrivée en octobre

Les familles de type 1 à 3 nous étaient inconnues, et la seule famille que nous avons pu identifier, HyperBro, n'est arrivée que début octobre. Nous reviendrons dessus ultérieurement. Concernant le type 3, l'analyse montrera que sa seule fonction est de charger du code depuis Dropbox. Nous y reviendrons ultérieurement.

Nous allons nous attarder un peu sur le type 1, mais la démarche a été similaire pour le type 2.

Le malware utilise 3 fichiers pour son chargement :

— un premier exécutable légitime et signé par Microsoft, vulnérable à une attaque de DLL Side-Loading [2]
— une DLL malveillante, nommée mpsvc.dll, chargée par le binaire légitime ci-dessus
— un fichier binaire contenant le code final du malware, mais obfusqué et compressé

La DLL malveillante se charge de décompresser et décoder le fichier binaire puis de le charger en mémoire en lançant un binaire légitime en mode suspendu, puis en remplaçant en mémoire le code à exécuter par le code malveillant, avant de reprendre l'exécution du binaire (méthode appelée « process hollowing » [4]). Ces méthodes sont connues et ont pour but d'évader les solutions de sécurité. Pour obtenir une version « unpackée », il suffit de placer un point d'arrêt après que le malware ait été déchiffré et décompressé en mémoire.

Le malware est écrit en C++ et articulé de façon modulaire. On constate que les informations RTTI sont présentes, et on peut donc en extraire le nom des classes. Il existe une classe « CHPPlugin » virtuelle pour

les greffons ajoutant des fonctionnalités, et une classe virtuelle « CHPNet » implémentée par toutes les classes relatives aux communications réseaux (par exemple « CHPHttp », « CHPTcp » ou « CHPUdp »).

Les noms de classes des greffons sont assez explicites, et correspondent aux fonctionnalités d'un RAT classique, avec notamment toutes les capacités attendues d'une application d'espionnage : récupération des frappes clavier (« CHPKeyLog »), capture vidéo du contenu de l'écran (« CHPAvi »), capture d'écran (« CHPScreen »), exécution de commande distante (« CHPCmd ») ....

L'analyse montre également un numéro de version, ce qui permet de voir qu'une des souches a pour version « 1.0 », tandis que les autres sont de version « 8.0 ». Parmi les deux souches reçues en octobre, l'une d'entre elle est un malware de type 1 et a pour version 9.0. Si l'on en croit les dates de compilation, la version 1.0 a été compilée en mai 2019, la version 8.0 en juillet, et la version 9.0 en octobre, ce qui donne une idée de la vitesse de développement de l'attaquant. Les différences entre versions sortent du cadre de cet article, mais les fonctionnalités restent globalement les mêmes.

On note aussi un algorithme de substitution utilisant une table fixe de 256 octets pour obfusquer les données envoyées au C&C, que l'on retrouve sur l'image 1, et enfin une communication systématique vers le nom de domaine légitime « api.dropbox.com ».

A l'issue de l'analyse, on obtient le tableau 1

Famille de malware	**C&C**
Type 1	download.safedog.co safe.mircosofdevice.com office.support.googldevice.com 45.77.41.49 35.220.232.71 (souche reçue en octobre)
Type 2	update.mircosoftdefender.com store.microsoftbetastore.com
Type 3	api.dropbox.com
HyperBro	35.220.135.85 (souche reçue en octobre)

Tableau 1. Liste des différents C&C après analyse des souches initiales

On constate donc que plusieurs souches utilisent le même C&C, et que celui-ci est parfois un nom de domaine, parfois une IP.

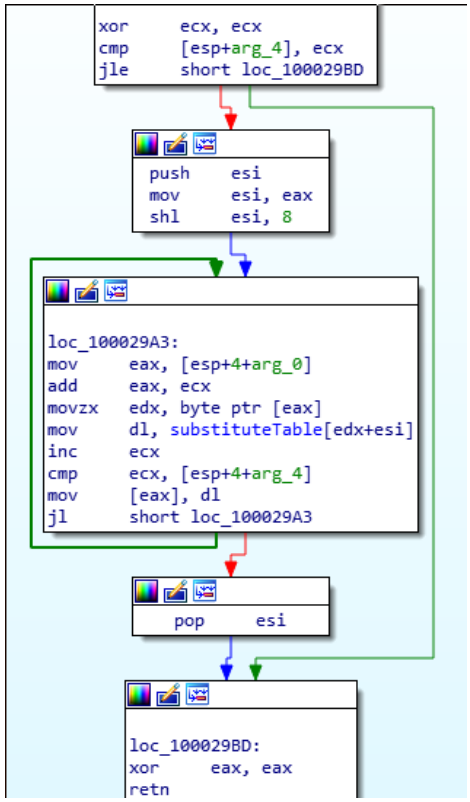

Fig. 1. Algorithme utilisé pour obfusquer les données envoyées au C&C

3 Pivots sur le code

Une fois cette première étape d'analyse terminée, nous allons nous intéresser à plusieurs méthodes qui nous ont permis ou non de trouver de nouvelles souches malveillantes liées à notre attaquant.

3.1 Recherche de chaînes de caractères ou octets peu courants

Généralement, la recherche la plus simple et rapide à effectuer consiste à identifier des chaînes de caractères peu courantes, puis à chercher leur présence dans une base de données de malware. Cette recherche peut commencer simplement par une requête sur un moteur de recherche, car de nombreuses sandbox en ligne stockent toutes les chaînes de caractères vues statiquement mais également dynamiquement lors de l'exécution d'un fichier soumis à leur service. Sinon, on peut utiliser le modificateur « content » [20] du service VirusTotal, qui va retourner tous les fichiers

correspondants au(x) motif(s) recherchés. L'approche est un peu similaire à « binacle » [10], présenté au SSTIC 2017. Pour des chaînes de caractères, cela s'avère largement suffisant, cependant on est limité dès lors que l'on souhaite exprimer des conditions par exemple.

Dans ce cas, on peut utiliser Yara,[1] qui dispose d'une grammaire beaucoup plus riche. Plusieurs services en ligne, ainsi que nos bases de données internes, sont compatibles avec cet outil. Cela veut dire qu'une fois que l'on a rédigé une règle, on peut l'ajouter à nos règles de « hunting ». Le principe de ces règles est qu'un email nous est envoyé dès qu'un fichier correspondant à la règle est envoyé sur une de nos sources de fichiers (base de données interne, VirusTotal ...).

Seulement, ces alertes ne portent que sur les fichiers qui seront analysés ou envoyés dans le futur. Si l'on souhaite chercher dans les fichiers passés, on peut alors utiliser les fonctionnalités de RetroHunt [21] de VirusTotal, qui vont permettre de chercher une règle Yara sur tous les fichiers envoyés ou analysés sur les trois derniers mois.

Dans le cas présent, les recherches sur des chaînes de caractères peu courantes, par exemple sur le nom de classe « CHPKeylog », n'ont rien donné. Ce n'est pas surprenant, puisque ces chaînes n'apparaissent qu'une fois le malware chargé en mémoire par les différents fichiers présentés en première partie.

3.2 Table fixe de substitution

Précédemment, nous avons identifié l'utilisation d'un algorithme de substitution de 256 octets pour obfusquer les données envoyées au C&C. Nous avons écrit une règle Yara contenant ces 256 octets, et nous l'avons ajouté à nos règles de « hunting », et nous avons également effectué un RetroHunt.

Après quelques heures, nous avons obtenu une liste de fichiers contenant la même table de substitution. Leur analyse a montré qu'il s'agissait bien de malwares de type 1, de version 4.0.

Nous avons pu ajouter deux C&C à notre liste :

— `test66.shopingchina.net`
— `update.google.com.updatesrvers.org`

Après quelques semaines, nous avons eu une alerte. En analysant le code, nous avons reconnu les mêmes classes C++ que celles vues dans notre malware initial, ce qui a confirmé le lien entre les deux malwares. Cependant, l'objectif de ce malware-ci est de faire de la redirection de

1. `https://virustotal.github.io/yara/`

trafic HTTP ou UDP d'un port en écoute vers une adresse IP et un port distant. On peut donc ajouter cet outil à l'arsenal de l'attaquant.

Un point important à noter : lors de l'analyse de cet algorithme, nous avons pensé que bien qu'il soit peu complexe (une simple substitution), il était très peu probable que l'on retrouve la même suite de 256 octets, dans le même ordre, par hasard. Cette impression s'est retrouvée confirmée par le faible nombre d'alertes (9 sur 6 mois) et la correspondance systématique après analyse des fichiers concernés avec notre investigation. Nous en avons déduit un marqueur fort de notre groupe d'attaquants.

Et pourtant... le 23 mars, bien après la soumission de cet article et la publication de notre blog, nous avons reçu une alerte qui s'est avérée être un faux-positif. Un fichier compilé le 29 mars 2015 et apparemment relatif à un serveur anti-triche contenait les fameux 256 octets, dans le bon ordre. Pourtant, ce fichier n'avait absolument rien à voir avec notre attaquant, en tout cas le code n'y ressemblait pas du tout. En poussant un peu les recherches, nous sommes tombés sur une question posée le 27 février 2015 sur CodeProject.com [1], dont on voit un extrait ci-dessous dans l'image 2.

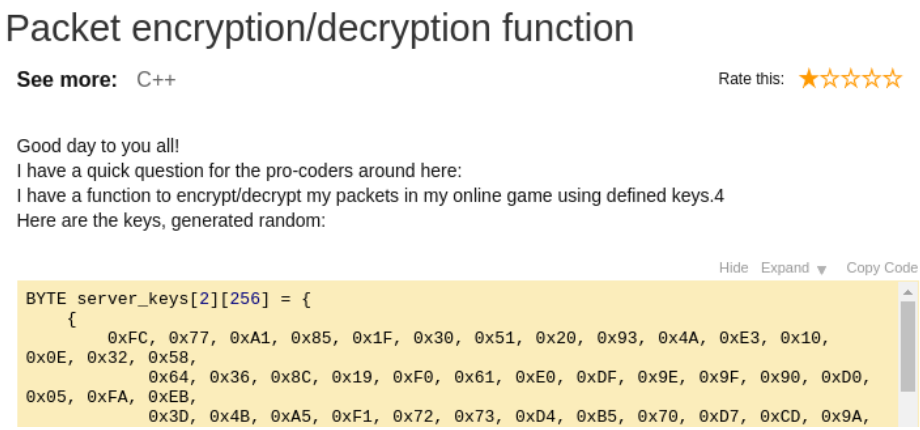

Fig. 2. Extrait de code trouvé sur Code Project

Nous en déduisons que notre attaquant a utilisé Google pour « coder » plus rapidement son algorithme de chiffrement réseau. Etant donné le faible nombre d'alertes que cette règle nous a remonté, et le nombre encore plus faible de faux-positifs (1 sur 9), le marqueur n'est pas mauvais pour autant. Mais c'est un bon rappel que les attaquants piquent souvent des

bouts de code à droite à gauche, et qu'il faut plus qu'une suite d'octets ou un algorithme pour une attribution correcte.

3.3 Utilisation des métadonnées

Une autre possibilité pour trouver du code relatif à un attaquant donné est d'utiliser les métadonnées des fichiers ou de documents. Nous allons en voir deux exemples ici.

Lorsqu'il s'agit d'un fichier exécutable au format PE, ces métadonnées peuvent être directement récupérable dans le fichier, comme les informations qui sont définies dans la ressource VERSIONINFO [16], telles que le nom du fichier, sa description, ou encore sa version. Parfois, ces métadonnées sont utilisées indirectement, comme le fameux « imphash » [14], qui calcule un condensat à partir de la table des imports d'un fichier exécutable.

Dans le cadre de cette investigation, nous nous sommes intéressés à la seule souche type 1 à notre disposition ayant pour version 1.0. Le champ « Internal Name » de cette souche est « HaoZipUpdate », qui s'avère être un gestionnaire de compression populaire en Asie. Une analyse du code a montré qu'il s'agissait d'un binaire légitime, dont quelques octets ont été patchés manuellement pour rediriger le flot d'exécution vers du code ajouté au milieu du fichier, comme on le voit dans les images 3 et 4.

Ce code se contente de résoudre quelques pointeurs de fonctions, puis il alloue une zone mémoire, et y recopie du code qui a été concaténé cette fois tout à la fin du fichier (ce qu'on appelle « l'overlay »), juste avant de rediriger le flot d'exécution vers cette zone mémoire. L'instruction utilisée pour cette dernière recopie est un ReadFile auquel on passe en argument l'offset à partir duquel on trouve l'overlay. Ceci se traduit par l'instruction « PUSH 15D08h », soit en assembleur x86 « 68 08 5D 01 00 ».

Revenons à nos métadonnées. Si l'on cherche [2] sur VirusTotal tous les fichiers non-signés, ayant pour nom « HaoZipUpdate », et ayant du code dans l'overlay, on obtient moins de 10 fichiers. On pourrait analyser manuellement ces 10 fichiers, mais une simple recherche binaire des octets « 68 08 5D 01 00 » dans ces fichiers nous renvoie un seul résultat.

Une analyse de ce fichier montre qu'il s'agit d'une version encore antérieure à la 1.0, avec des noms de classes similaire au type 1. Cette version n'a que très peu de greffons, et même pas de numéro de version. Elle date également de mai 2019.

2. name :"HaoZipUpdate" tag :"overlay" NOT tag :"signed"

```
                    loc_403AF7:                             ; CODE XREF: sub_403AB6+37↑j
56                                  push    esi
8B 35 20 D0 40 00                   mov     esi, ds:LoadLibraryA
57                                  push    edi
68 64 ED 40 00                      push    offset aComctl32Dll ; "COMCTL32.DLL"
C7 45 BC 08 00 00 00                mov     [ebp+7D0h+var_814], 8
C7 45 C0 FF 00 00 00                mov     [ebp+7D0h+var_810], 0FFh
FF D6                               call    esi ; LoadLibraryA
8B 3D 60 D0 40 00                   mov     edi, ds:GetProcAddress
68 74 ED 40 00                      push    offset aInitcommoncont ; "InitCommonControlsEx"
50                                  push    eax             ; hModule
89 45 C8                            mov     [ebp+7D0h+var_808], eax
FF D7                               call    edi ; GetProcAddress
3B C3                               cmp     eax, ebx
74 06                               jz      short loc_403B2F
8D 4D BC                            lea     ecx, [ebp+7D0h+var_814]
51                                  push    ecx
FF D0                               call    eax

                    loc_403B2F:                             ; CODE XREF: sub_403AB6+71↑j
68 8C ED 40 00                      push    offset aUser32Dll ; "User32.dll"
FF D6                               call    esi ; LoadLibraryA
68 98 ED 40 00                      push    offset aMessageboxw ; "MessageBoxW"
50                                  push    eax             ; hModule
89 45 C4                            mov     [ebp+7D0h+hLibModule], eax
FF D7                               call    edi ; GetProcAddress
```

Fig. 3. Version originale du fichier HaoZipUpdate

Une fois ce nouveau binaire trouvé, nous cherchons des bouts de code entiers de celui-ci via le modificateur « content » de VirusTotal, et finissons par trouver un document Word malveillant embarquant le fichier tel quel dans un objet OLE. Les métadonnées de ce document n'ont pas été enlevées, et le nom d'auteur est très spécifique : « Dell_20170514745 ».

En cherchant d'autres documents avec le même auteur, on en trouve 3 autres, également malveillants, qui nous serviront à trouver des cibles en les cherchant dans notre télémétrie.

4 Pivots sur l'infrastructure

Un autre moyen de trouver d'autres indicateurs et d'en savoir plus sur notre attaquant est de s'intéresser à son infrastructure. Plusieurs méthodes d'investigations existent, qui permettent généralement d'obtenir d'autres noms de domaines et adresses IP liées, qui eux-mêmes pourront amener à la découverte d'autres souches malveillantes, pour lesquelles on recommence les étapes précédentes. Cette partie se propose donc d'énumérer quelques méthodes que nous avons utilisées, avec succès ou non, pour récupérer un maximum de C&C liés à notre attaquant.

```
                    loc_403AF7:                             ; CODE XREF: sub_403AB6+37↑j
56                                  push    esi
8B 35 20 D0 40 00                   mov     esi, ds:LoadLibraryA
57                                  push    edi
68 64 ED 40 00                      push    offset aKernel32Dll_0 ; "kerneL32.DLL"
C7 45 BC 08 00 00 00                mov     [ebp+7D0h+var_814], 8
C7 45 C0 FF 00 00 00                mov     [ebp+7D0h+var_810], 0FFh
FF D6                               call    esi ; LoadLibraryA
8B 3D 60 D0 40 00                   mov     edi, ds:GetProcAddress
57                                  push    edi
90                                  nop
90                                  nop
90                                  nop
90                                  nop
50                                  push    eax
89 45 C8                            mov     [ebp+7D0h+var_808], eax
E8 18 8D 00 00                      call    resolvFunctions_LoadShellcode
90                                  nop
8D 4D BC                            lea     ecx, [ebp+7D0h+var_814]
51                                  push    ecx
FF D0                               call    eax
68 8C ED 40 00                      push    offset aUser32Dll ; "User32.dll"
FF D6                               call    esi ; LoadLibraryA
68 98 ED 40 00                      push    offset aMessageboxw ; "MessageBoxW"
50                                  push    eax             ; hModule
89 45 C4                            mov     [ebp+7D0h+hLibModule], eax
FF D7                               call    edi ; GetProcAddress
```

Fig. 4. Version modifiée du fichier HaoZipUpdate

4.1 Passive DNS

La première méthode courante qui est utilisée est le passive DNS. Certains services tels PassiveTotal[3] ou DNSDB[4] stockent toutes les associations « nom de domaine – adresse IP » qu'ils ont pu observer. Généralement, ces services disposent d'accords avec des serveurs DNS d'opérateurs, ou gèrent leurs propres infrastructures DNS.

La conséquence est qu'on obtient un historique des adresses IP associées à un nom de domaine (avec ou sans ses sous-domaines), et pour chaque adresse IP, on peut obtenir tous les noms de domaines que le service de passive DNS a vu pointer vers cette IP.

L'intérêt de telles bases de données est que certains attaquants réutilisent certains serveurs pour des campagnes d'attaques différentes. Parfois, on arrive donc à obtenir de nouveaux C&C que l'on pourra lier à notre attaquant, et qui pourront eux-mêmes être utilisés pour chercher d'autres souches malveillantes, ou d'autres machines compromises dans le cas d'une réponse à incident.

Il faut toutefois être vigilant à ce que ces autres noms de domaines ou IP soient utilisés dans le même intervalle de temps, car rien n'interdit à

3. `https://community.riskiq.com/`
4. `https://www.farsightsecurity.com/dnsdb-community-edition/`

deux groupes distincts d'utiliser un même serveur alors qu'ils n'ont aucun lien. Certains hébergeurs étant peu regardants sur l'utilisation qui est faite de leurs services, ils peuvent devenir un nid à APT, et on ne pourra alors effectuer aucune corrélation utile. De même, cette méthode a bien sûr ses limites dans le cas de serveurs partagés, puisque plusieurs domaines qui n'ont rien à voir vont pointer vers la même IP au même moment.

Prenons un exemple pour illustrer cette méthode. Dans l'étape précédente, nous avons identifié que le domaine update.mircosoftdefender.com est utilisé comme C&C par notre attaquant.

L'image 5 montre l'historique des adresses IP pointées par ce domaine dans PassiveTotal.

Resolve	Location	Network	ASN	First	Last
45.32.13.143	JP	45.32.8.0/21	20473	2020-03-31	2020-04-21
43.228.126.172	SG	43.228.126.0/24	133905	2019-07-19	2020-03-20

Fig. 5. Historique des adresses IP du domaine update.mircosoftdefender.com

On voit que l'IP 43.228.126.172 a été utilisée depuis juillet 2019, période à laquelle a été identifiée la première attaque.

L'image 6 montre tous les domaines qui ont été vus pointant vers cette IP.

Resolve	First	Last
update.microsoftdnsdown.com	2019-11-17	2020-03-31
support.microsoftdnsdown.com	2019-10-21	2020-03-31
update.mircosoftdefender.com	2019-07-19	2020-03-20
rollbackup.us	2018-05-22	2019-04-27
photon-sg-1.sakay.ph	2018-06-04	2018-10-06
owenysoo.cf	2018-04-17	2018-07-02
server.bego-meroty.ga	2018-06-01	2018-06-01
accept-idc638a898fdd25b31ae5d1d38e.us	2018-05-18	2018-05-23
client-idc638a898fdd25b31ae5d1d38e.us	2018-05-18	2018-05-23
customer-idc638a898fdd25b31ae5d1d38e.us	2018-05-18	2018-05-23
appleid.apple.com.accept-idc638a898fdd25b31ae5d1d38e.us	2018-05-19	2018-05-19
www.accept-idc638a898fdd25b31ae5d1d38e.us	2018-05-19	2018-05-19
support-midden-team.ga	2018-04-17	2018-05-04

1 - 13 of 13 ▾

Fig. 6. Historique des adresses IP du domaine update.mircosoftdefender.com

On s'intéresse aux noms de domaines utilisés en même temps que notre C&C, c'est-à-dire tous les domaines entre juillet 2019 et mars 2020.

On envisage les domaines update.microsoftdnsdown.com et support.microsoftdnsdown.com comme des candidats potentiels.

Plusieurs choses nous font pencher vers l'idée que ces domaines appartiennent à notre attaquant :

— Ils utilisent tous les deux la marque « Microsoft » (avec ou sans faute) dans le nom de domaine ;
— Ils utilisent tous deux « update » comme sous-domaine ;
— Une des souches malveillantes avait pour C&C le domaine « safe.mircosofdevice.com ». On retrouve ici le même sous-domaine « safe » ;
— Ils ont tous le même Registrar (GoDaddy) et NameServer (DomainControl.com). Cette combinaison est extrêmement courante, mais il est tout de même intéressant de le noter

Le plus pertinent aurait été de trouver un malware appartenant à une des familles analysées précédemment et ayant ces domaines comme C&C, mais cette recherche s'est malheureusement révélée infructueuse.

Au passage, nous avons constaté que l'attaquant attribue une adresse IP à ses sous-domaines uniquement : update.mircosoftdefender.com a pointé vers une adresse IP pertinente, alors que mircosoftdefender.com a toujours pointé vers les IP par défaut de GoDaddy. Cela contourne les services de passive DNS qui se contentent de résoudre l'adresse IP des noms de domaines dont ils récupèrent les listes auprès d'organismes officiels tels l'AFNIC.

4.2 Corrélations d'infrastructure

Certains services tels Censys,[5] Shodan[6] ou Onyphe[7] (cocorico !) parcourent quotidiennement un grand nombre d'adresses IP en se connectant sur les ports « connus », et stockent dans leur base de données beaucoup de métadonnées, telles que les versions d'un serveur web, les certificats vus, les en-têtes HTTP retournées par un serveur web, etc. Il est parfois possible de se servir de ces informations pour effectuer des pivots et trouver de nouveaux C&C liés à un groupe d'attaquant.

Par exemple, si l'on identifie un certificat TLS lié à notre groupe d'attaquant, on peut requêter ces services pour obtenir une liste d'adresses

5. `https://censys.io/`
6. `https://www.shodan.io/`
7. `https://www.onyphe.io/`

IP sur lesquelles ce certificat a été vu, et ainsi enrichir notre liste d'indicateurs. Dans le cas de cette investigation, cette méthode n'a pas permis d'obtenir de nouvelles IP, mais il nous semble important de la citer car elle fonctionne dans de nombreux cas.

Nous avons constaté qu'en fin d'année 2019, l'attaquant a déplacé son infrastructure vers des IP hébergées chez Google Cloud. En filtrant sur les plages IP de cet hébergeur ainsi que sur certaines caractéristiques simples présentes dans les en-têtes HTTP, il a été possible de trouver de nouveaux C&C. Ces caractéristiques étaient par exemple le serveur web utilisé, sa version, ou encore la taille des données renvoyées. Nous avons ainsi pu ajouter 3 adresses IP de C&C qui présentaient des caractéristiques similaires.

De plus, cette étape a été cruciale car elle nous a permis de relier deux C&C de malwares de familles différentes. Pour rappel, en octobre 2019, nous avons reçu une nouvelle souche HyperBro, ainsi qu'une souche de type 1 de version 9.0. Ces deux malwares avaient été trouvés sur une même machine, cependant, l'analyse forensics n'a pas permis de trouver de lien direct entre ces deux malwares. Concrètement, nous n'étions pas sûr que les malwares étaient liés, étant donné qu'ils pouvaient très bien être le fruit de deux attaques de groupes différents, sans corrélation aucune. C'est donc l'analyse de l'infrastructure qui nous a permis de confirmer qu'un seul et même attaquant était derrière les deux infrastructures.

Au-delà de cette confirmation, cela nous a apporté un premier lien avec un groupe « connu ». En effet, à notre connaissance, le malware HyperBro est utilisé par un seul groupe, nommé Emissary Panda, Iron Tiger, LuckyMouse ou APT27 [3].

4.3 Enregistrement des noms de domaine

Certains services tels DomainTools [8] stockent l'historique de tous les enregistrements Whois vus sur la quasi-totalité des noms de domaines enregistrés, à part quelques TLD [9] peu connus ou réticents à donner ces infos. Ces registres contiennent des informations qui peuvent être utilisées pour effectuer des corrélations. Par exemple, si l'attaquant utilise la même adresse email pour enregistrer différents noms de domaine, il sera possible à l'aide de cette base de récupérer tous les noms de domaines enregistrés avec cette adresse. Cependant, cela est de moins en moins vrai, car au-delà des attaquants qui utilisent des services d'anonymisation

8. `http://whois.domaintools.com/`
9. `https://fr.wikipedia.org/wiki/Domaine_de_premier_niveau`

lors de l'enregistrement du nom de domaine, la loi GDPR entrée en vigueur en mai 2018 a entrainé le masquage de cette information dans les enregistrements Whois.

Pour autant, il reste des choses à faire. Par exemple, on peut toujours utiliser l'adresse email qui se trouve dans l'enregistrement DNS SOA [10] pour la corréler avec d'autres noms de domaine.

Parfois, certains attaquants enregistrent plusieurs noms de domaine en une fois, sur le même registrar, avec le même service d'anonymisation. La conséquence est que la date de création de tous ces domaines est très proche (quelques secondes d'écart). Une fois que l'on trouve un nom de domaine d'un tel attaquant, on peut alors récupérer la liste de tous les noms de domaines qui ont été enregistrés chez le même registrar à la même date, et après avoir filtré sur l'heure de création, on peut trouver des noms de domaine liés. Il faut toutefois faire attention car la probabilité d'un faux positif est élevée ! On peut alors se baser sur une nomenclature commune, un sujet d'intérêt pour les victimes et/ou l'attaquant, ou encore résoudre l'adresse IP liée à un domaine et chercher des corrélations dans l'infrastructure.

Cette méthode a fonctionné dans notre cas. Par exemple, dans l'étape d'analyse des malwares, nous avons identifié le domaine mircosoftdefender.com comme appartenant à notre attaquant.

Celui-ci a été enregistré auprès de GoDaddy le 2018-08-09 à 08 :40 :27. Le NameServer associé est Domaincontrol.com, et l'attaquant a utilisé le service d'anonymisation domainsbyproxy.com. Ce sont tous deux des services par défaut de GoDaddy. Etant donné la part de marché [9] importante de ce Registrar, on pourrait penser qu'il y aura trop de résultats. En effet, d'après DomainTools, 7661 domaines ont été enregistrés sur la seule journée du 2018-08-09.

Cependant, si on filtre les domaines ayant été créés entre 08h40 et 08h41, il n'en reste plus que 4, à savoir dinohonevice.com et luxespiremag.com créés à 08 :40 :10, mircosoftdefender.com à 08 :40 :27 et googleusermessage.com à 08 :40 :28.

Ce dernier nous semble un bon candidat, pour plusieurs raisons :

— Il a été créé une seconde après notre C&C ;
— Une marque d'entreprise informatique est présente dans le nom de domaine ;
— Seule une adresse IP par défaut de GoDaddy semble avoir été associée à ce domaine. Cela correspond aux cas vus précédemment, où les IP ne sont attribuées qu'aux sous-domaines.

10. `https://fr.wikipedia.org/wiki/SOA_Resource_Record`

Nous ne pourrons malheureusement pas en savoir plus, car aucun sous-domaine n'a été trouvé. Mais nous notons tout de même le marqueur, car il pourrait servir si l'attaquant réutilise ce domaine dans le futur, ou si un chercheur tombe sur un malware ayant ce domaine comme C&C.

4.4 Sandbox publiques

Plusieurs solutions en ligne permettent d'analyser un exécutable ou document en sandbox, et fournissent en sortie une trace de l'exécution, avec notamment les connexions réseaux effectuées. On peut alors utiliser ces services pour retrouver des souches qui se seraient connectées à un C&C donné. Etant un éditeur antivirus, nous avons le même type de base de données en interne.

En requêtant ces bases de données avec tous les C&C récoltés lors des étapes précédentes, nous avons trouvé d'autres malwares liés à notre attaquant.

Par exemple, un malware de type 1 avait pour C&C test66.shopingchina.net. Si l'on cherche ce domaine dans VirusTotal,[11] on voit que deux autres sous-domaines sont connus de ce service. En allant sur chacun [12, 13] d'entre eux, on obtient une liste d'exécutables contactant ces sous-domaines. Nous n'avons pas trouvé d'utilisation légitime du domaine shopingchina.net, on peut donc raisonnablement penser qu'il n'est pas compromis, et que c'est donc bien l'attaquant qui le gère.

Par conséquent, on considère que ces autres malwares sont liés, et on les ajoute à l'arsenal de notre attaquant.

Par cette méthode, nous avons trouvé quatre familles de malwares supplémentaires :

- PlugX [15], un malware existant depuis au moins 2008 et utilisé dans de très nombreuses attaques ciblées ;
- Trochilus [12], un RAT public, dont des versions modifiées ont également été vues dans de précédentes attaques ciblées [6, 17] ;
- Un malware codé avec les classes MFC dont nous n'avons pas identifié la famille ;
- Cobalt Strike,[14] un framework offensif bien connu des équipes de Red Team.

11. `https://www.virustotal.com/gui/domain/shopingchina.net/relations`
12. `https://www.virustotal.com/gui/domain/jqb.shopingchina.net/relations`
13. `https://www.virustotal.com/gui/domain/fn.shopingchina.net/relations`
14. `https://www.cobaltstrike.com/features`

Contrairement à HyperBro, ces familles de malwares sont utilisées par plusieurs groupes différents, et ne permettent donc pas de rapprocher notre groupe d'attaquants d'un groupe connu. On peut tout au plus remarquer que plusieurs entreprises attribuent certaines de ces attaques à la Chine.

5 Utilisation de la télémétrie

Arrivés à cette étape, nous disposions des éléments suivants :
- Plusieurs versions de différentes familles de malware, connues et inconnues ;
- Des documents malveillants liés à un de ces malwares inconnus ;
- De nombreux noms de domaines et adresses IP liés à notre attaquant.

Il nous manque le vecteur d'accès, qui malheureusement n'a pas été identifié par l'entreprise effectuant la réponse à incident. Nous ne sommes pas sûrs non plus de la victimologie, étant donné que pour l'instant nous n'avons connaissance que de la compromission d'une entreprise de paris en ligne.

En cherchant les documents malveillants identifiés précédemment dans notre télémétrie, nous avons pu trouver des envois de mails vers une autre entreprise située en Asie du Sud-Est, appartenant elle aussi au domaine des paris en ligne.

Grâce à cette trouvaille, nous savons donc qu'un des vecteurs de compromission est le spear-phishing, et le secteur ciblé semble bien être celui des paris en ligne en Asie du Sud-Est.

En revanche, une recherche sur d'éventuelles détections des différents malwares identifiés a été effectuée et ne nous a pas permis d'identifier de nouvelles victimes. Cela montre le caractère particulièrement ciblé de cette attaque.

6 Cercle vertueux

Comme nous l'avons déjà indiqué, à chaque fois que l'on trouve un nouvel indicateur, il faut recommencer les étapes.
- Un nouveau nom de domaine ajoute potentiellement de nouvelles adresses IP et malwares liés ;
- Une nouvelle adresse IP ajoute potentiellement de nouveaux domaines et malwares liés ;
- Une nouvelle famille de malwares ajoute potentiellement de nouveaux C&C, donc des domaines et adresses IP.

Chacun de ces indicateurs permet potentiellement de détecter de nouvelles victimes via la télémétrie, et pourrait également permettre un rapprochement avec un groupe connu, et donc plus de contexte pour l'attaque.

Ci-dessous un exemple de corrélation effectuée une fois les étapes précédentes effectuées une première fois.

6.1 Mutex

Le code malveillant d'un des exécutables de la famille de malware « Trochilus », nommé « diskshawin.exe », [15] est chargé en mémoire par un exécutable [16] lancé lui-même par un fichier RAR auto-extractible. [17] Cet exécutable utilise plusieurs mutex aux noms apparemment aléatoires, « cc5d64b344700e403e2sse », « cc5d6b4700e403e2sse » et « cc5d6b4700032eSS ».

En cherchant ces mutex sur Google, nous avons trouvé un rapport de sandbox [18] correspondant à une souche appartenant à la famille BbsRat [11], nommée « diskwinshadow.exe » et contactant le nom de domaine « bot.googlerenewals.net ». En effectuant des recherches sur ce nom de domaine, on retrouve un rapport [18] de la société ClearSky publié en 2017 relatif au groupe Winnti [5].

Ce nom regroupe en fait plusieurs groupes en son sein, et plusieurs rapports publiés [13] sur ces groupes indiquent qu'il serait d'origine chinoise. La nomenclature similaire des noms de fichiers et des mutex, et la cohérence entre l'origine présumée du groupe et les victimes constatées de notre côté renforcent notre conviction que les deux groupes sont probablement liés.

7 Utilisation des fonctionnalités du malware

Nos connaissances sur ce groupe d'attaquants sont de plus en plus étendues, mais il est possible d'aller plus loin. Lors de l'analyse des malwares type 1, nous avons identifié une connexion au domaine api.dropbox.com.

15. SHA256 : 4e3e9e4613d414ba671fd35d7d70d0c3093cd322f5f297281a502420741c03c8
16. SHA256 : 02d6ae0039abf9b042c60c6b0eb84f6af1283a25932c1d69a9646bc7dea34984
17. SHA256 : 60a7b03a7776a26aea2d0d64246ea24d204dd9db815b47e1c32171783a50b27b
18. `https://www.hybrid-analysis.com/sample/f5dc823aa3c51d96ac632e311fa198a6a7bbc37bf771b2c0dad7d7532f8f9d7f?environmentId=120`

7.1 Nouvelle charge utile

Après une analyse plus poussée, on constate que Dropbox est utilisé comme un canal de contrôle supplémentaire. En effet, lors de sa première infection, la machine génère un ID unique, et s'en sert pour créer un répertoire sur un dépôt Dropbox. Ensuite, le malware va régulièrement vérifier si un fichier d'extension « asc » est présent dans ce répertoire, et si c'est le cas, il le déchiffre, passe quelques pointeurs de fonction sur la pile, puis redirige le flux d'exécution vers le code déchiffré. Il nous semble donc important de récupérer ce fichier, pour pouvoir le détecter et le bloquer.

Pour pouvoir accéder à ce dépôt Dropbox sans interaction utilisateur, le malware embarque une clé d'API de Dropbox. Nous avons alors extrait cette clé, et via l'implémentation Python officielle de l'API Dropbox,[19] nous avons pu accéder au dépôt de l'attaquant afin d'en observer le contenu.

Nous avons fait les constats suivants :

— 142 répertoires différents existent à la racine ;
— 129 de ces derniers contiennent des fichiers d'extension « asc » ;
— Il y a 26 versions de fichiers « asc » différents, visant les architectures x86 et x64.

Comme nous le soupçonnions, ces fichiers d'extensions « asc » contiennent du code malveillant. Il s'agit en fait d'une nouvelle famille de malware, qui mérite donc une nouvelle analyse. Cette analyse sort du cadre de cet article, mais un résumé des fonctionnalités se trouve dans l'article publié sur notre blog [8].

On retient surtout que cette porte dérobée utilise Dropbox comme C&C. Ainsi, si un attaquant veut exécuter une commande chez une victime, il doit écrire cette commande dans un fichier nommé « yasHPHFJ ». De façon régulière, le malware regarde si un fichier à ce nom existe sur le dépôt, et si c'est le cas, il exécute la commande, et écrit le résultat, également sur le dépôt, dans un fichier nommé « Csaujdnc ». Il est donc possible, en ayant accès au dépôt, de voir les différentes commandes exécutées.

7.2 Outils de post-exploitation

De même, on trouve une cinquantaine d'exécutables supplémentaires présents sur le dépôt. Leur analyse montre qu'il s'agit principalement d'exécutables malveillants qui correspondent à des outils utilisés généralement lors de la phase de post-exploitation. Ces outils incluent des logiciels

19. `https://github.com/dropbox/dropbox-sdk-python`

de récupération de mots de passe tels Mimikatz ou QuarksPwdDump, des outils pour contourner l'UAC, des outils d'énumération de partages réseaux tels nbtscan, ou encore des outils d'élévation de privilèges via l'exploitation de vulnérabilités et des outils de brute-force.

7.3 Analyse des commandes

L'analyse des commandes passées par l'attaquant sur 67 ID différents a montré que la plupart d'entre elles consistaient à lister le contenu de répertoires, puis à lancer par différents moyens les outils mentionnés ci-dessus. Les résultats de cette analyse de commandes ont également été publiés sur notre blog.

Certaines de ces commandes sont particulièrement intéressantes : sur une machine, l'attaquant a téléchargé une charge malveillante qui est directement stockée sur une adresse IP distante. En regardant le passive DNS de cette adresse IP, nous avons retrouvé un nom de domaine qui est lié à Winnti. Cela a donc renforcé notre conviction que notre attaquant avait des liens avec ce groupe.

De même, les commandes avec lesquelles l'attaquant envoie des fichiers vers la victime contiennent le chemin absolu des fichiers. Nous avons donc pu voir un nom de répertoire « DRBControl » dans l'arborescence de l'attaquant, ce qui nous a inspiré le nom de cette investigation.

Nous avons également confirmé la victimologie, car nous avons retrouvé les entreprises de certaines des victimes infectées par ce malware utilisant Dropbox, et elles sont également liées aux paris en ligne.

Pour finir, nous avons fait le constat que dans les fichiers téléchargés par l'attaquant, il y avait de très nombreuses bases de données, des fichiers contenant les frappes claviers, des documents PDF et Office, des cookies, et même une base de données Keypass.

8 Chronologie

Avant de conclure, faisons une petite chronologie des événements liés à cette attaque.

— 10 juillet 2019 : début de l'investigation
— Fin juillet 2019 : finalisation d'une première passe d'analyse des différentes familles de malwares et de l'infrastructure de l'attaquant
— Courant août 2019 : analyse plus poussée des chargeurs de code, découverte des documents malveillants et de spear-phishing envoyés en mai 2019

— Fin août 2019 : début de l'analyse des commandes de l'attaquant et identification de nouvelles cibles
— Courant Septembre 2019 : première corrélation avec Winnti trouvée. Toutes les clés d'API Dropbox sont invalidées
— 2 Octobre 2019 : réception de deux nouvelles souches, dont une nouvelle version d'HyperBro
— Courant octobre 2019 : pause complète de l'investigation
— Courant décembre 2019 : reprise, investigation plus poussée de l'infrastructure de l'attaquant, lien avec EmissaryPanda confirmé
— Fin décembre 2019 : première version du rapport
— Courant janvier 2020 : légères modifications du rapport suite aux remarques du marketing technique
— 2 février 2020, 23h44 : soumission au SSTIC
— 18 février 2020 : publication du post blog [7] et du papier d'investigation [8] sur le blog de Trend Micro
— 10 mars 2020 : notification de l'acceptation du SSTIC
— 25 avril 2020 : envoi d'un draft du papier SSTIC

A noter que cette investigation a été effectuée à 99% par trois personnes ayant des compétences différentes, sur des fuseaux horaires différents, et avec des niveaux de disponibilité différents, c'est-à-dire qu'un chercheur n'est pas forcément dédié à une seule investigation.

9 Conclusion

Nous sommes partis d'une quinzaine de souches appartenant à 4 familles de malwares distinctes, avec 5 noms de domaines et 3 adresses IP de C&C différents.

A l'arrivée, nous avons :

— 4 familles de malwares additionnelles
— 14 noms de domaines supplémentaires
— 6 adresses IP supplémentaires
— Des dizaines de souches différentes (liste complète des condensats disponible dans notre papier [8])
— Un vecteur de compromission (4 documents malveillants utilisés pour du spear-phishing)
— Une liste de nombreux outils de post-exploitation utilisés par l'attaquant
— Une victimologie précise
— Des liens avec deux groupes d'attaquants connus et documentés

Pour une victime, l'intérêt de telles méthodes pour enrichir sa liste d'indicateurs et obtenir plus de contexte est indéniable. Cependant, il est certain qu'une telle investigation prend du temps et nécessite des ressources importantes. De même, il faut parfois attendre plusieurs semaines avant d'avoir plus de contexte et/ou d'informations, suite par exemple à la découverte d'une nouvelle famille de malware.

Concernant plus particulièrement cette présentation, il faut préciser que la plupart des concepts évoqués ici sont appliqués par de nombreuses entreprises de « threat intelligence ». Cependant, si dans certains cas, ces concepts sont applicables par tous, dans d'autres, ils requièrent un accès à des plateformes payantes. Cela est particulièrement vrai en ce qui concerne VirusTotal, alors que d'autres plateformes comme Passive Total proposent un accès gratuit mais avec certaines limitations, notamment sur le nombre de requêtes quotidiennes. D'autres concepts eux, requièrent de la télémétrie. Sachant que tous les acteurs de la sécurité informatique ont une vision différente, du fait d'un placement, de clients ou d'usages différents, il faut accepter que la vue est forcément partielle. C'est également l'intérêt d'établir des liens de confiance entre chercheurs d'entreprises différentes, afin de pouvoir compléter sa vision d'une attaque.

Terminons sur un point non négligeable : même si nous avons montré quelques méthodes permettant de suivre un attaquant, que nous avons qualifié dans le titre de négligent, ce dernier est loin d'être idiot. Il ne se privera pas d'utiliser l'information disponible en source ouverte, que ce soit les publications d'éditeurs de sécurité le concernant, mais également les mentions de chercheurs sur Twitter par exemple, pour s'améliorer.

Cette investigation n'a pas failli à cette règle. Après avoir publié notre investigation en février 2020, nous avons identifiées de nouvelles attaques liées à ce groupe en mars 2020. Le même malware de type 1 a été identifié, pointant vers une infrastructure totalement différente, mais utilisant toujours Dropbox comme C&C. Après extraction de la clé d'API, nous avons constaté que ses permissions ont été ajustées, et il n'est désormais plus possible de lister le contenu du dépôt. Il faut donc faire preuve de créativité, mais également de discernement lorsqu'on rédige un papier d'investigation !

Références

1. Packet encryption/decryption function. `https://www.codeproject.com/Questions/881460/Packet-encryption-decryption-function`.
2. Mitre ATT&CK. DLL Side-Loading. `https://attack.mitre.org/techniques/T1073/`.

3. Mitre ATT&CK. HyperBro. https://attack.mitre.org/software/S0398/.
4. Mitre ATT&CK. Process Hollowing. https://attack.mitre.org/techniques/T1093/.
5. Mitre ATT&CK. Winnti Group. https://attack.mitre.org/groups/G0044/.
6. K. Kohei B. Sy, CH Lei. ChessMaster Makes its Move : A Look into the Campaign's Cyberespionage Arsenal. https://blog.trendmicro.com/trendlabs-security-intelligence/chessmaster-cyber-espionage-campaign/, 2017.
7. K. Lu J. Yaneza D. Lunghi, C. Pernet. Uncovering a Cyberespionage Campaign Targeting Gambling Companies in Southeast Asia. https://www.trendmicro.com/vinfo/us/security/news/cyber-attacks/operation-drbcontrol-uncovering-a-cyberespionage-campaign-targeting-gambling-companies-in-southeast-asia, 2020.
8. K. Lu J. Yaneza D. Lunghi, C. Pernet. Uncovering DRBControl - Inside the Cyberespionage Campaign Targeting Gambling Operations. https://documents.trendmicro.com/assets/white_papers/wp-uncovering-DRBcontrol.pdf, 2020.
9. ICANN. ICANN Contractual Compliance Performance Report. https://features.icann.org/compliance/registrars-list.
10. Guillaume Jeanne. Binacle : indexation "full-bin" de fichiers binaires. https://www.sstic.org/2017/presentation/binacle_indexation_full-bin_de_fichiers_binaires/, 2017.
11. Malpedia. BBSRAT. https://malpedia.caad.fkie.fraunhofer.de/details/win.bbsrat.
12. Malpedia. Trochilus RAT. https://malpedia.caad.fkie.fraunhofer.de/details/win.trochilus_rat.
13. Malpedia. Winnti. https://malpedia.caad.fkie.fraunhofer.de/details/win.winnti.
14. Mandiant. Tracking Malware with Import Hashing. https://www.fireeye.com/blog/threat-research/2014/01/tracking-malware-import-hashing.html, 2014.
15. Trend Micro. PlugX : New Tool For a Not So New Campaign. https://blog.trendmicro.com/trendlabs-security-intelligence/plugx-new-tool-for-a-not-so-new-campaign/, 2012.
16. Microsoft. VERSIONINFO resource. https://docs.microsoft.com/en-us/windows/win32/menurc/versioninfo-resource.
17. PWC. Operation Cloud Hopper - Technical Annex. https://www.pwc.co.uk/cyber-security/pdf/cloud-hopper-annex-b-final.pdf, 2017.
18. ClearSky Research Team. Recent Winnti Infrastructure and Samples. https://www.clearskysec.com/winnti/, 2017.
19. Zero Chen Theo Chen. CLAMBLING - A New Backdoor Base On Dropbox. http://www.talent-jump.com/article/2020/02/17/CLAMBLING-A-New-Backdoor-Base-On-Dropbox-en/, 2020.
20. VirusTotal. Content search (VTGrep). https://support.virustotal.com/hc/en-us/articles/360001386897-Content-search-VTGrep-.
21. VirusTotal. Retrohunt. https://support.virustotal.com/hc/en-us/articles/360001293377-VT-Retrohunt.