SOPHOS

Zoom sur les exploits :

Stratégie globale de prévention des exploits

Les exploits profitent des vulnérabilités de logiciels légitimes, tels que Adobe Flash ou Microsoft Office, pour infecter les systèmes. Ils sont habituellement utilisés par des cybercriminels dans le but de pénétrer les défenses des entreprises. Leurs objectifs sont divers : voler des données, les prendre en otage contre le paiement d'une rançon, réaliser une reconnaissance ou simplement déployer un malware plus traditionnel.

Des exploits sont couramment utilisés lors de cyber attaques : plus de 90% des violations de données déclarées montrent l'utilisation d'un exploit à un moment de l'attaque. C'est pourquoi intégrer la prévention des exploits dans sa stratégie globale de sécurité est extrêmement avantageux.

Les exploits existent depuis plus de 30 ans, il n'est donc pas étonnant que la grande majorité des éditeurs de sécurité revendiquent un certain niveau de prévention contre ces menaces. Cependant, l'étendue et le degré de protection varient de manière significative d'un éditeur à l'autre. Pour certains, il s'agit d'une case à cocher tandis que pour d'autres, il s'agit d'un élément majeur de leur stratégie. Consultez ce document pour en savoir plus sur les exploits et les différents niveaux de prévention proposés par les produits de sécurité phares du marché.

Sommaire

L'industrie des exploits : Crimeware-as-a-Service

Grâce aux kits d'exploits, les auteurs de malwares n'ont pas à se soucier de savoir comment trouver des bugs dans Java, Silverlight ou encore Flash, comment intégrer ces bugs dans des exploits opérationnels, comment trouver des serveurs web non sécurisés pour héberger les exploits, ni même comment attirer des victimes potentielles sur des pages web piégées.

De même, les auteurs des kits d'exploits n'ont pas à se soucier de l'écriture des malwares, de posséder des serveurs pour suivre l'évolution des systèmes infectés ou de collecter de l'argent auprès des différentes victimes, ni même à s'impliquer dans les processus d'extraction des données volées ou de leur revente.

Les cybercrimes représentant désormais une industrie de plusieurs milliards de dollars (estimée à près de 2 000 milliards de dollars d'ici 2019), chaque élément d'une attaque a été industrialisé.

Les criminels ont le luxe de pouvoir se spécialiser dans un ou plusieurs domaines des menaces, ce que l'on appelle ironiquement le CaaS pour « Crimeware-as-a-Service » (Logiciel criminel en tant que service).

Dans cette industrie désormais très lucrative, des courtiers en exploits sont apparus ; ils achètent les exploits aux personnes les découvrant et les revendent ensuite aux parties intéressées, qu'il s'agisse d'une instance gouvernementale ou d'un pirate crapuleux.

Les acheteurs gardent invariablement leurs objectifs secrets. Comme l'explique Kevin Mitnick, fondateur de Mitnick's Absolute Zero Day Exploit Exchange, à Wired : « Quand un de nos clients cherche une vulnérabilité Zero-Day, quelle qu'en soit la raison, nous ne posons pas de question, car de toute façon il ne nous répondrait pas. Les chercheurs trouvent les vulnérabilités, nous les vendent pour X, puis nous les revendons à nos clients pour Y en nous dégageant une marge. »

« Quand un de nos clients cherche une vulnérabilité Zero-Day, quelle qu'en soit la raison, nous ne posons pas de question, car de toute façon il ne nous répondrait pas. Les chercheurs trouvent les vulnérabilités, nous les vendent pour X, puis nous les revendons à nos clients pour Y en nous dégageant une marge. »

Kevin Mitnick

Techniques de prévention des exploits

Avec plus de 400 000 échantillons uniques de malware créés chaque jour et des milliers de nouvelles vulnérabilités découvertes chaque année, prévenir les attaques malveillantes est devenu un défi colossal. Cet essor incroyable du nombre de variantes de malwares requiert des approches innovantes si l'on veut pouvoir se défendre contre les cybercriminels.

Un examen attentif de l'industrie moderne du cybercrime montre l'opportunité offerte par des défenses asymétriques. Il s'avère que, malgré le défilé constant de nouvelles attaques, il n'existe qu'environ une vingtaine de techniques utilisées pour exploiter un logiciel. Une approche capable de contrer cette poignée de techniques d'exploits, au lieu de cibler individuellement chacun des exploits, serait extrêmement puissante.

De plus, selon la vulnérabilité, les pirates doivent souvent coupler plusieurs techniques d'exploits afin de réussir à distribuer leur malware. Ces techniques n'évoluent pas énormément d'une année sur l'autre, seuls un ou deux nouveaux stratagèmes sont ajoutés à la liste des techniques disponibles.

Il est parfois surprenant de ne pas trouver de techniques notables de prévention des exploits dans certains produits phares de sécurité. Et tandis que certains des nouveaux éditeurs offrent une technologie de nouvelle génération avec une prise en charge plus globale de la prévention des exploits, celle-ci reste parcellaire.

Vous trouverez ci-dessous une liste de techniques de prévention des exploits visant à éliminer l'ensemble des vulnérabilités et à bloquer les techniques d'exploit utilisées par les cybercriminels et les instances gouvernementales. Chaque technique de prévention varie selon les éditeurs. Il est important de noter que lorsqu'un éditeur prétend prévenir les exploits, cela veut souvent dire qu'il protège contre une infime partie des techniques d'exploits les plus couramment utilisées, qui ne s'appliquent souvent pas aux applications 64 bits. Seul Sophos offre réellement la prévention complète des exploits.

Application de la Prévention de l'exécution des données (DEP)

La prévention de l'exécution des données (DEP) est un ensemble de technologies matérielles et logicielles qui contrôlent l'utilisation de la mémoire pour empêcher les dépassements de la mémoire tampon. Sans la DEP, un pirate peut tenter d'exploiter la vulnérabilité d'un logiciel en allant sur le code malveillant (shellcode) à un emplacement de la mémoire, tel que le tas (heap) ou la pile (stack), où des données contrôlées par le pirate existent. Sans la DEP, ces zones sont normalement marquées comme exécutables, laissant le code malveillant s'exécuter.

La DEP est disponible dans Windows XP et les versions postérieures, comme option facultative qui doit être paramétrée par l'éditeur du logiciel au moment de la conception d'une application. De plus, les attaques sont capables de contourner la protection DEP intégrée et, par conséquent, il n'est pas recommandé de s'appuyer sur son implémentation native du système d'exploitation.

Application systématique de la technique ASLR (distribution aléatoire de l'espace d'adressage)

Certains exploits ciblent spécifiquement des emplacements de la mémoire connus pour être associés à des processus particuliers. Dans les versions anciennes de Windows (dont Windows XP), les processus clés avaient tendance à être chargés dans des emplacements prédictibles de la mémoire lors du démarrage du système. La distribution aléatoire de l'espace d'adressage (ASLR) renforce l'aléa des emplacements de la mémoire utilisée par les fichiers système et d'autres programmes, ce qui complique la tâche du pirate qui ne sait pas où chercher l'emplacement d'un programme donné, notamment la base de l'exécutable et la position de la pile, du tas et des bibliothèques.

La technique ASLR est uniquement disponible sur Windows Vista et versions ultérieures et, comme avec la DEP, elle doit être configurée par l'éditeur du logiciel au moment de la conception d'une application. De même que pour la DEP, les attaques sont capables de contourner la protection ASLR intégrée et, par conséquent, il n'est pas recommandé de s'appuyer sur son implémentation native du système d'exploitation.

Bottom-up ASLR

Si elle est activée, la technique dite du « Bottom-up ASLR » améliore l'entropie ou le caractère aléatoire de l'application systématique de l'ASLR.

L'avantage principal de l'application systématique de l'ASLR et du Bottom-up ASLR dans Sophos Intercept X est que les adresses de base des applications ne sont pas uniquement randomisées à chaque reboot, mais également à chaque démarrage d'une application protégée.

Null Page (déréférencement du pointeur Null)

À partir de Windows 8, Microsoft refuse aux programmes la capacité d'allouer ou de cartographier la « page NULL » (mémoire résidant à l'adresse virtuelle 0x00000000 dans l'espace d'adressage). En faisant cela, Microsoft réduit avec succès les risques d'une exploitation directe de toute une classe de vulnérabilités appelées « déréférencement du pointeur NULL ».

Sur les systèmes Windows XP, Vista et 7, l'exploitation d'une telle faille permettrait à un pirate d'exécuter du code dans le contexte du noyau (sous le niveau de privilège ring0 du CPU), aboutissant à l'élévation des privilèges vers l'un des niveaux les plus élevés.

De telles vulnérabilités offrent virtuellement aux pirates un accès à toutes les zones du système d'exploitation.

Heap Spray Pre-Allocation

Le 'heap spray' est une technique qui n'exploite pas en elle-même de vulnérabilité, mais qui est utilisée pour rendre une vulnérabilité plus facile à exploiter. À l'aide d'une technique appelée Heap Feng Shui[1], un pirate est capable de déterminer avec précision la position sur le tas de structures de données ou de shellcode visés, facilitant ainsi une exploitation fiable de la vulnérabilité d'un logiciel.

Un mécanisme typique de prévention de heap sprays consiste à réserver ou à préallouer les adresses mémoire couramment utilisées, pour qu'elles ne puissent pas être utilisées pour héberger des charges virales. Les attaquants plus créatifs connaissent généralement ces adresses, donc en réalité cette mesure d'atténuation aura peu d'effets. Également appelée 'Anti-HeapSpray Enforcement' ou 'Shellcode Preallocation', la pré-allocation de heap spray est généralement efficace contre les exploits utilisés par les organismes d'essais.

Dynamic Heap Spray

En comparaison avec le 'Heap Spray Pre-Allocation', le mécanisme de prévention du 'Dynamic Heap Spray' est généralement déclenché par une hausse soudaine de la consommation de mémoire.

La prévention du Dynamic Heap Spray analyse le contenu des allocations de mémoire récentes dans le but de détecter des signatures indiquant la présence de heap sprays contenant des 'NOP sleds' (NOP : No Operation), des 'NOP sleds' polymorphes, des matrices JavaScript ou autres séquences suspectes, placés sur le tas de manière à faciliter une attaque d'exploit.

Stack Pivot (falsification de la pile)

La pile d'une application est une zone de mémoire contenant, entre autres choses, une liste des adresses mémoire (appelées adresses de retour). Ces emplacements contiennent le code dont le processeur a besoin par la suite pour s'exécuter.

La falsification de la pile est largement utilisée par les exploits pour contourner les protections telles que la DEP, en assemblant par exemple des manipulations particulières lors d'une attaque de type ROP (Return-Oriented Programming). Avec cette technique, les attaques peuvent pivoter depuis la vraie pile vers une fausse pile, qui peut être un tampon complètement contrôlé par le pirate, à partir duquel ce dernier peut contrôler le futur flux d'exécution d'un programme.

1 https://cansecwest.com/slides/2014/The%20Art%20of%20Leaks%20-%20read%20version%20-%20Yoyo.pdf

Stack Exec (MemProt)

Dans des circonstances normales, la pile contient des données et des adresses pointant vers du code pour que le processus puisse s'exécuter. En utilisant un dépassement de la mémoire tampon[2], les pirates peuvent écraser les données de la pile avec du code aléatoire. Pour que le code puisse s'exécuter sur le processeur, la zone de mémoire de la pile doit être rendue exécutable afin de contourner la DEP. Une fois que la mémoire de pile est exécutable, un pirate peut très facilement exécuter le code d'un programme.

Prévention Stack-Based ROP (Caller)

Pour contrer les technologies de sécurité telles que la DEP et l'ASLR, les attaquants piratent désormais le contrôle de flux des applications Internet vulnérables. Ces attaques en mémoire sont invisibles aux yeux des antivirus, de la plupart des produits Next-Gen et d'autres cyber défenses, car elles n'impliquent aucun fichier malveillant. Au contraire, ces attaques sont construites à l'exécution en combinant de petits morceaux de codes anodins faisant partie d'applications existantes, telles qu'Internet Explorer ou Adobe Flash Player, appelées 'attaques de réutilisation de code' ou 'de programmation orientée retour' (return-oriented programming ou ROP).

Au cours d'un contrôle de flux normal, les fonctions API sensibles (comme VirtualAlloc et CreateProcess) sont appelées par l'instruction 'CALL'. Au moment de l'appel d'une fonction API sensible, les défenses ROP stoppent l'exécution de code afin de déterminer l'adresse de l'API, en utilisant l'adresse de 'retour' localisée en haut de la pile. Si l'instruction de l'adresse de l'API appelée n'est pas un CALL, le processus est interrompu.

Le contenu de la pile étant accessible en écriture, un attaquant peut écrire des valeurs spécifiques sur la pile pour tromper l'analyse de la défense de ROP basée sur la pile. La défense contre le ROP basée sur la pile ne peut déterminer si le contenu de la pile est bénin ou manipulé par un attaquant.

Prévention des Branch-based ROP (Intégrité du contrôle de flux au niveau processeur)

Comme expliqué précédemment, les mécanismes de défense basés sur la pile contre les attaques de ROP sont peu précis et exposés aux manipulations. Pour améliorer cela, les données analysées à l'exécution doivent être plus affinées et résistantes aux manipulations.

Sophos Intercept X intègre le renforcement de l'intégrité du contrôle de flux au niveau processeur en exploitant une fonction matérielle inutilisée disponible dans les principaux processeurs Intel® (depuis 2008). Le processeur en lui-même offre des données en lecture seule pour augmenter la détection des attaques sophistiquées d'exploits au moment de l'exécution. Utiliser des traces suivies par matériel (branchement) a un avantage certain en matière de sécurité en comparaison de l'approche logicielle basée sur la pile. Les informations de branchement récupérées sur ces traces peuvent non seulement identifier la cible du branchement, mais également la source. Il montre donc l'origine du changement de contrôle de flux. Cette information spécifique ne peut pas être obtenue avec le même niveau de confiance en utilisant une solution basée sur la pile, comme Microsoft EMET ou Palo Alto Networks Traps.

2 https://en.wikipedia.org/wiki/Stack_buffer_overflow

Les informations de branchement dans les traces matérielles ne peuvent pas être manipulées. Elles ne peuvent donc pas être écrasées par des données contrôlées par un pirate. Les approches basées sur la pile s'appuient sur les données de la pile, qui sont - spécialement dans le cas d'une attaque de ROP - sous le contrôle du pirate, ce qui peut en retour induire le défenseur en erreur. En revanche, les données tracées par matériel et examinées par Sophos Intercept X sont plus fiables et inviolables.

Une implémentation alternative de l'intégrité du contrôle de flux au niveau processeur (HA-CFI) proposée par Endgame est basée sur l'apprentissage du contrôle de flux régulier pour détecter toute déviation du chemin de code prévu par le développeur. Il doit être formé en continu pour établir une liste blanche d'adresses valides du pointeur, reflétant l'ensemble des fonctions et des versions de l'application protégée. Sophos Intercept X ne requiert pas d'apprentissage préalable et fonctionne également correctement durant le changement de contexte du thread et la limitation du processeur.

Sophos Intercept X utilise automatiquement le contrôle de flux au niveau processeur contrôle suivies par matériel lorsqu'il détecte un processeur Intel® Core™ i3, i5 ou i7 (CPU). Si le processeur pris en charge n'est pas détecté, Sophos Intercept X reviendra automatiquement à des contrôles d'intégrité basés sur la pile, uniquement logiciels

Sophos Intercept X n'exploite pas seulement les traces matérielles pour renforcer la détection ROP. Il les utilise également pour le filtrage des tables d'import (IAF), afin de protéger l'accès aux tables d'import des applications protégées.

Remarque : Les correctifs des vulnérabilités de Spectre concernant la prédiction de branchement à l'intérieur du processeur Intel n'affectent pas le bon fonctionnement de Sophos Intercept X.

Structured Exception Handler Overwrite Protection (SEHOP)

Un pirate peut écraser, avec une valeur contrôlée, le gestionnaire d'exception sur la pile. Une fois qu'une exception se produit, le système d'exploitation remonte la chaîne de l'exception et appelle tous les gestionnaires sur chaque enregistrement d'exception.

Dès le moment où le pirate contrôle les enregistrements, le système d'exploitation sera à sa merci, lui donnant le contrôle du flux de l'exécution.

La protection SEHOP est une option facultative sur Windows Vista et versions ultérieures, et doit être paramétrée par l'éditeur du logiciel au moment de la conception de l'application. Les attaques contournent la protection SEHOP intégrée et, par conséquent, il n'est pas recommandé de dépendre de l'implémentation du système d'exploitation.

Filtrage des accès à la table d'import (IAF)

Un pirate a besoin à terme des adresses de fonctions système spécifiques (par ex. kernel32!VirtualProtect) pour être en mesure d'effectuer des attaques malveillantes.

Ces adresses peuvent être récupérées depuis différentes sources, l'une d'elle étant la table d'import (IAT) d'un module chargé. La table d'import sert de table de consultation lorsqu'une application appelle une fonction dans un module différent. Un programme compilé ne pouvant pas connaître l'emplacement de la mémoire des bibliothèques auquel il dépend, un saut indirect est requis à chaque fois qu'un appel d'API est émis. Comme le lien dynamique charge les modules et les réunit, il écrit les adresses réelles dans les emplacements de la table d'import, pour qu'elles pointent vers les emplacements de la mémoire des bibliothèques correspondantes.

Sophos Intercept X intègre le filtrage des accès à la table d'import assisté par matériel en exploitant des fonctions matérielles disponibles dans les principaux processeurs Intel® (depuis 2008). En plus des traces suivies par matériel (branchement) pour appliquer l'intégrité du contrôle de flux, Sophos Intercept X exploite également la prédiction de branchement matérielle pour améliorer la protection des tables d'import.

Remarque : Les correctifs des vulnérabilités de Spectre concernant la prédiction de branchement à l'intérieur du processeur Intel n'affectent pas le bon fonctionnement de Sophos Intercept X.

Load Library

Les pirates peuvent tenter de charger des bibliothèques malveillantes en les plaçant sur des chemins UNC. Contrôler tous les appels vers l'API LoadLibrary peut servir à prévenir ce type de chargement de bibliothèque.

Reflective DLL Injection

Normalement, lorsque vous chargez une DLL dans Windows, vous appelez l'API LoadLibrary. LoadLibrary prend comme entrée le chemin du fichier d'une DLL, puis se charge dans la mémoire.

Le chargement DLL réflectif fait référence au chargement d'une DLL depuis la mémoire plutôt que depuis le disque. Windows n'a pas de fonction LoadLibrary qui le prenne en charge, vous devrez donc écrire votre propre fonction pour en bénéficier. Un des avantages à écrire votre propre fonction est que vous pouvez omettre certains des éléments exécutés normalement par Windows, tels que l'enregistrement de la DLL comme module chargé dans le processus, ce qui a pour effet de rendre l'examen du chargement réflectif plus sournois. Meterpreter est un exemple d'outil qui utilise le chargement réflectif pour se cacher. La prévention est réalisée en contrôlant si une DLL est chargée de manière réflective dans la mémoire.

Shellcode

Un shellcode est un petit morceau de code utilisé comme charge utile lors de l'exploitation d'une vulnérabilité logicielle. Son nom provient du fait qu'à l'origine, il démarrait une commande shell à partir de laquelle l'attaquant pouvait contrôler la machine compromise, mais toute partie de code remplissant la même fonction peut être appelée 'shellcode'

Un exploit injecte habituellement un shellcode dans le processus ciblé avant ou au moment d'exploiter une vulnérabilité, dans le but de prendre le contrôle sur le pointeur d'instruction du processeur (EIP/RIP). Le pointeur d'instruction est reprogrammé pour pointer sur le shellcode, après quoi il est exécuté et accomplit sa tâche.

VBScript God Mode

Sur Windows, VBScript peut être utilisé dans les navigateurs ou le shell local. Lorsqu'il est utilisé dans le navigateur, les capacités de VBScript sont restreintes pour des raisons de sécurité. Cette restriction est contrôlée par l'option mode sans échec. Si cette option est modifiée, VBScript en HTML agit comme s'il était dans le shell local. Par conséquent, les pirates peuvent facilement écrire du code malveillant dans VBScript. La manipulation de l'option de mode sans échec sur VBScript dans le navigateur Web est appelée God Mode[3] (mode Dieu).

Par exemple, un pirate peut modifier la valeur de l'option mode sans échec en exploitant la vulnérabilité CVE-2014-6332[4], un bug causé par une mauvaise manipulation au moment du redimensionnement d'un objet Array dans le moteur VBScript d'Internet Explorer. En mode Dieu, le code aléatoire écrit sur VBScript peut casser la technologie de sandboxing du navigateur. Grâce au mode Dieu, la prévention de l'exécution des données (DEP), la distribution aléatoire de l'espace d'adressage (ASLR) et le graphe de contrôle de flux (CFG) n'entrent pas en jeu.

WoW64

Microsoft fournit une compatibilité rétroactive pour les logiciels 32 bits sur les éditions 64 bits de Windows, via la couche « Windows on Windows » (WoW). Certains aspects de l'implémentation du processus WoW ouvrent des pistes intéressantes pour les pirates, s'ils veulent compliquer l'analyse dynamique, dépaqueter des fichiers binaires ou contourner la prévention des exploits.

Le comportement d'une application 32 bits sous un environnement WoW64 est différent à de nombreux égards d'un vrai système 32 bits. La capacité de basculer entre différents modes d'exécution à l'exécution peut offrir aux pirates des méthodes pour l'exploitation, le camouflage et l'anti-émulation, notamment :

- Des techniques ROP supplémentaires absentes du code 32 bits.
- Des codeurs avec mode d'exécution mixte de la charge utile.
- Des fonctionnalités de l'environnement d'exécution qui peuvent réduire l'efficacité de la prévention.
- Des 'hooks' (code) de contournement insérés par les logiciels de sécurité, uniquement dans l'espace utilisateur 32 bits.

3 https://en.wikipedia.org/wiki/Glossary_of_video_game_terms#God_mode
4 https://www.rapid7.com/db/modules/exploit/windows/browser/ms14_064_ole_code_execution

La plupart des logiciels de protection Endpoint insèreront seulement un hook pour récupérer les fonctions API sensibles dans l'espace mémoire utilisateur si un processus est exécuté sous WOW64. Si un pirate arrive à passer au mode 64 bits, il peut alors accéder aux versions 64 bits des fonctions API sensibles non hookées, qui font autrement l'objet d'un hook dans le mode 32 bits.

Sur les éditions Windows 64 bits, Sophos Intercept X interdit au code de passer directement du mode 32 bits au mode 64 bits (par ex. en utilisant la technique ROP), tout en autorisant la couche WoW64 à réaliser cette transition.

Pour plus d'informations sur l'utilisation abusive de WoW64, consultez les recherches publiées par Duo Security : « WoW64 and So Can You »[5] et « Mitigating Wow64 Exploit Attacks »[6].

Syscall (Appel système)

Un appel système (syscall) est la manière dont un programme informatique fait appel à un service du noyau du système d'exploitation. Cela inclut les services liés au matériel, comme l'accès au disque local ou la création et l'exécution de nouveaux processus.

De manière générale, le système d'exploitation offre une interface de programmation (API) générique qui se situe entre les programmes normaux et le système d'exploitation. En situation normale, une application appellera toujours une API pour demander une tâche spécifique au noyau. Le logiciel de sécurité place des hooks (code) sur les fonctions API sensibles pour intercepter et réaliser des contrôles (par ex. une analyse antivirus), avant d'autoriser le noyau à répondre à la demande.

Un attaquant peut profiter du fait que :

- Toutes les fonctions API ne sont pas hookées par un logiciel de sécurité ; seules les fonctions sensibles le sont.
- Les fichiers de remplacement utilisés pour appeler les fonctions du noyau sont très similaires ; seul l'index des fonctions est unique.

En appelant une fonction de remplacement non-sensible et non contrôlée au niveau d'un offset (pour calculer de manière intentionnelle un service sensible du noyau), un attaquant peut réussir à contourner la plupart des logiciels de sécurité ou des analyses de sandboxing.

Sophos Intercept X offre une nouvelle approche pour empêcher les attaquants de toucher aux fonctions sensibles du noyau par le biais de fonctions API non protégées.

Pour plus d'informations sur l'utilisation abusive des appels système, consultez l'article de blog de BreakDev.org intitulé « Antivirus Real-time Protection From The Inside »[7].

Process Hollowing (processus creux)

La technique des processus creux consiste à charger un processus fiable (par ex. explorer.exe ou svchost.exe) sur un système, dont le rôle sera uniquement de servir de conteneur de code malveillant. Un processus creux est typiquement créé en état suspendu, puis le mapping de sa mémoire est supprimé et remplacé par du code malveillant. De manière similaire à l'injection de code, l'exécution de code malveillant est masquée sous un processus légitime et peut ainsi contourner les défenses et les analyses de détection.

5 https://duo.com/blog/wow64-and-so-can-you
6 https://hitmanpro.wordpress.com/2015/11/10/mitigating-wow64-exploit-attacks
7 https://breakdev.org/defeating-antivirus-real-time-protection-from-the-inside

Process Doppelgänging

La plupart des ordinateurs Windows utilisent le système de fichiers NTFS. En 2007, Microsoft a introduit une nouvelle fonction appelée Transactional NTFS (TxF). Celle-ci permet de traiter de nombreuses opérations de fichiers comme un tout : elles peuvent soit réussir comme un tout et être menées à bien, soit échouer comme un tout et être annulées. De cette manière, une application peut apporter de nombreux changements à plusieurs fichiers sur le disque et restaurer tous les fichiers vers leur état d'origine si une erreur est détectée. La fonction TxF est couramment utilisée lors de l'installation des mises à jour de Windows.

Le 'Process Doppelgänging' exploite le mécanisme TxF pour cacher un malware. Il choisit un fichier inoffensif, l'écrase puis exécute le malware via une API de bas niveau pour, par exemple, se faire passer pour un fichier de confiance (de manière similaire au processus creux). Juste avant d'exécuter le malware, il annule et restaure tous les changements pour empêcher le logiciel antivirus d'analyser le contenu du fichier réellement exécuté. S'il est ouvert, le fichier présent sur le disque ne contiendra donc aucun contenu douteux. De plus, ce fichier peut tout à fait être une application populaire signée numériquement.

Détournement de DLL

En raison d'une vulnérabilité communément appelée « détournement de DLL » (DLL hijacking, DLL spoofing, DLL preloading, binary planting), de nombreux programmes chargent et exécutent un fichier DLL malveillant contenu dans le même dossier que les fichiers légitimes qu'ils sont amenés à ouvrir.

Échange dynamique de données (DDE)

L'Échange dynamique de données (DDE) de Windows est un protocole client-serveur pour la communication inter-processus (IPC) entre les applications. Les attaquants peuvent utiliser le DDE pour exécuter des commandes arbitraires. Par exemple, les documents Microsoft Office peuvent être infectés avec des commandes DDEAUTO et être utilisés pour exécuter des commandes PowerShell via des campagnes de spear phishing ou du contenu Web hébergé, évitant ainsi l'utilisation de macros Visual Basic pour Applications (VBA). Il est également possible d'intégrer des commandes DDEAUTO dans le corps des emails ou des demandes de réunion, qui seront alors exécutées au moment de l'envoi d'une réponse ou de l'acceptation dans Microsoft Outlook.

Grâce à la technologie de verrouillage des applications, Sophos Intercept X empêche naturellement l'exécution de code malveillant par le biais de l'Échange dynamique de données.

Application Lockdown

Dans l'éventualité où un pirate réussirait à exploiter et à contourner toutes les techniques de prévention, Sophos Intercept X limite ses capacités. Cette fonctionnalité, appelée « Application Lockdown » (verrouillage des applications), a pour but d'empêcher les pirates d'introduire du code indésirable.

Le verrouillage des applications stoppe les attaques dont le mode opératoire ne repose généralement pas sur les failles des applications. Une telle attaque pourrait être, par exemple, l'utilisation d'une macro artisanale (malveillante) dans un document Microsoft Office jointe à un email de phishing. Les macros sont potentiellement dangereuses, car elles sont créées dans le langage de programmation du Visual Basic pour Applications (VBA), qui inclut la capacité de télécharger et d'exécuter des fichiers binaires à partir du Web et qui permet également l'utilisation du PowerShell et d'autres applications de confiance.

Cette fonctionnalité imprévue (ou exploit logique) offre aux pirates un avantage certain, car ils n'ont pas besoin d'exploiter une faille d'un logiciel ou de trouver le moyen de contourner les défenses de la mémoire et du code pour parvenir à infecter des ordinateurs. Ils exploitent simplement une fonction standard offerte par une application de confiance largement utilisée, et avec un peu d'ingénierie sociale ils n'ont besoin que de persuader la victime d'ouvrir le document malveillant.

Sans avoir besoin de maintenir une liste noire de dossiers, Sophos Intercept X va bloquer automatiquement une application protégée en fonction de son comportement. Par exemple, lorsqu'une application de bureautique est utilisée pour lancer PowerShell, accéder à l'Interface de gestion Windows (WMI), exécuter une macro pour installer un code arbitraire ou manipuler une zone critique du système, Sophos Intercept X va bloquer l'action malveillante - même si l'attaque n'engendre pas de processus enfant.

Java Lockdown

Auparavant, les kits d'exploits étaient un élément clé du téléchargement passif de malwares. Ils exploitaient des vulnérabilités dans l'environnement d'exécution Java (JRE) pour injecter une charge virale Windows PE. Le JRE est chargé comme plug-in ou module d'extension dans les principaux navigateurs.

Sophos Intercept X empêche le JRE d'exécuter des applications qui ne sont pas en Java. Par exemple, Sophos Intercept X arrêtera une application Java si elle tente de s'introduire et de s'exécuter dans un Windows PE binaire. En plus de cela, les attaquants ne peuvent pas exploiter Java pour manipuler les registres de démarrage, dont le dossier Démarrage, Run, RunOnce et toute autre clé de registre.

Remarque : Avec l'introduction de Java 8 Update 20 en 2014, le niveau de sécurité des applications Java est par défaut paramétré comme Élevé. Il est alors beaucoup plus difficile pour les attaquants d'exécuter des exploits Java avec suffisamment de permissions pour infecter un système d'extrémité. En conséquence, les exploits Java sont tombés en disgrâce, ce qui rend le verrouillage de Java quelque peu obsolète.

Code Cave

Le Code Cave est une technique utilisée par les attaquants pour modifier une application légitime en y ajoutant une application supplémentaire. Cette dernière est insérée dans ce que l'on appelle le Code Cave, une section du fichier de l'application ciblée qui est inutilisé par le programme. Les Codes Cave existent dans la plupart des applications, et l'ajout de code dans ces sections ne modifie généralement pas le comportement de l'application initiale.

Le code exécutable inséré dans le Code Cave est souvent un simple shell launcher distant ou une porte dérobée. Ce dernier peut être très petit et permettre un accès direct à l'attaquant au système d'extrémité, d'où il pourra alors réaliser d'autres actions malveillantes. Pour ce type d'attaque, l'attaquant doit au préalable établir une présence sur le système d'extrémité afin de déployer une application dotée d'une porte dérobée ou tromper l'utilisateur pour qu'il télécharge et installe une application dont le Code Cave est déjà exploité.

Les attaquants utilisent principalement le Code Cave pour se cacher aux yeux des utilisateurs ou des administrateurs. L'application originale téléchargée fonctionne comme attendue, mais l'application insérée s'exécute également.

Si l'application modifiée est à l'origine un outil professionnel légitime et qu'un antivirus traditionnel détecte un problème, l'administrateur ne remettra pas immédiatement en question cette application. Il se peut que les administrateurs ajoutent l'application à la liste d'exemption, présumant que l'antivirus a généré un faux positif. Dans ce cas, l'attaquant a réussi à se maintenir durablement sur le système d'extrémité et il a même trompé les administrateurs qui ont facilité l'exécution de l'application intégrée.

Dans le cas d'une attaque de chaîne d'approvisionnement, un cybercriminel peut également attaquer les serveurs de mise à jour du logiciel en liant du code malveillant à une mise à jour, pour infecter de manière silencieuse les clients à l'aide, par exemple, d'un ransomware ou d'un wiper.

Sophos Intercept X bloque automatiquement l'exécution des applications qui sont liées à une porte dérobée. Il détecte même le shellcode ajouté lorsque l'exécution du code ne va pas vers un Code Cave ou une section ajoutée dans le fichier PE infecté. Il offre une protection élargie contre les outils d'injection de shellcode, tels que Shellter et Backdoor Factory.

Migration des processus – remote reflective DLL injection

La migration des processus est une technique courante réalisée par un attaquant lorsqu'il établit pour la première fois une présence sur un appareil, et qu'il veut migrer vers un autre processus pour élever ses privilèges ou pour obtenir un accès durable. L'attaquant ne veut pas perdre le contrôle lorsque l'utilisateur ferme son navigateur ou quitte un processus compromis, c'est pourquoi il va chercher à migrer vers un processus système.

Une attaque d'injection Reflective DLL à distance est similaire à une migration des processus. L'attaquant a déjà compromis un processus et, à partir de celui-ci, il va manipuler un autre processus pour charger des DLL et exécuter un code arbitraire.

Local Privilege Escalation (LPE)

Sophos Intercept X empêche l'élévation de privilèges d'un processus dont les droits d'accès sont faibles par le biais d'un jeton volé à processus doté de privilèges plus élevés. Cette technique est souvent utilisée en tandem avec une autre vulnérabilité pour installer et exécuter le code malveillant d'un attaquant disposant des autorisations du système.

Injection de code DoublePulsar

DoublePulsar était à l'origine une porte dérobée développée par l'équipe Equation Group de la NSA (National Security Agency) aux États-Unis, qui a été divulguée par The Shadow Brokers en début d'année 2017. L'implant contient une technique d'injection innovante retrouvée dans plusieurs exploits de la NSA, notamment EternalBlue et EternalRomance. Ces exploits ont également été utilisés dans le composant du ver à auto-propagation des épidémies de WannaCry et NotPetya.

La technique d'injection de code de DoublePulsar utilise un appel APC (Asynchronous Procedure Call) pour exécuter le code arbitraire (shellcode) à l'intérieur d'un processus fiable courant. Sophos Intercept X arrête naturellement la méthode fondamentale utilisée par DoublePulsar et stoppe ainsi les attaques qui font appel à la même technique pour injecter du code.

Injection de code AtomBombing

Les injections APC (Asynchronous Procedure Call) impliquent le fait d'attacher du code malveillant à la file d'attente des appels APC du thread d'un processus. Les fonctions APC mises en attente sont exécutées lorsque le thread entre en état modifiable. AtomBombing est une variante qui utilise les appels APC pour invoquer du code malveillant précédemment écrit sur la table Atom globale.

Injection de code DoubleAgent

DoubleAgent exploite un outil légitime de Windows appelé Microsoft Application Verifier. Cet outil est inclus dans toutes les versions de Microsoft Windows et est utilisé comme outil de vérification à l'exécution afin de découvrir et de réparer tous les bugs des applications. Application Verifier peut être configuré pour charger n'importe quelle bibliothèque à partir du disque, ouvrant ainsi la possibilité de charger une bibliothèque malveillante dotée des autorisations du processus victime.

DoubleAgent est qualifié de vulnérabilité et d'attaque Zero-Day par les antivirus, mais en réalité, l'objectif d'Application Verifier est de charger du code arbitraire dans n'importe quelle application choisie, y compris les processus de productivité et Windows de confiance.

Sophos Intercept X empêche l'injection de code via Application Verifier.

Fonctionnalités d'Intercept X

Fonctions	
PREVENTION DES EXPLOITS	
Application de la Prévention de l'exécution des données	✓
Distribution aléatoire de l'espace d'adressage (ASLR)	✓
Bottom-up ASLR	✓
Null Page (déréférencement du pointeur Null)	✓
Allocation de heap spray	✓
Dynamic Heap Spray	✓
Stack Pivot	✓
Stack Exec (MemProt)	✓
Prévention Stack-Based ROP (Caller)	✓
Prévention des Branch-based ROP (assisté par matériel)	✓
Structured Exception Handler Overwrite Protection (SEHOP)	✓
Filtrage des accès à la table d'import (IAF)	✓
Load Library	✓
Reflective DLL Injection	✓
Shellcode	✓
VBScript God Mode	✓
Wow64	✓
Syscall (Appel système)	✓
Hollow Process	✓
DLL Hijacking	✓
Squiblydoo AppLocker Bypass	✓
Protection APC (Double Pulsar / AtomBombing)	✓
Processus d'élévation de privilèges	✓
PRÉVENTION ACTIVE ADVERSARY	
Protection contre le vol des codes d'accès	✓
Prévention du Code Cave	✓
Détection MITB (Safe Browsing)	✓
Détection du trafic malveillant	✓
Détection des Meterpreter Shell	✓

Fonctions	
PREVENTION ANTI-RANSOMWARE	
Protection des fichiers contre les ransomwares (CryptoGuard)	✓
Restauration automatique des fichiers (CryptoGuard)	✓
Protection de l'enregistrement de démarrage et contre la réinitialisation du disque (WipeGuard)	✓
VERROUILLAGE DES APPLICATIONS	
Navigateurs Web (y compris HTA)	✓
Plugins navigateur Web	✓
Java	✓
Applications Média	✓
Applications Office	✓
DEEP LEARNING	
Détection des malwares par Deep Learning	✓
Blocage des applications potentiellement indésirables (PUA) par Deep Learning	✓
Suppression des faux positifs	✓
Live Protection	✓
RÉPONSE INVESTIGATION SUPPRESSION	
Analyse de l'origine de l'attaque (RCA)	✓
Sophos Clean	✓
Synchronized Security Heartbeat	✓
DÉPLOIEMENT	
Peut fonctionner comme un agent autonome	✓
Peut fonctionner en plus de l'antivirus existant	✓
Peut fonctionner comme composant de l'agent Sophos Endpoint actuel	✓
Windows 7	✓
Windows 8	✓
Windows 8.1	✓
Windows 10	✓
macOS*	✓

* Fonctions prises en charge : CryptoGuard, détection du trafic malveillant (MTD), Synchronized Security Heartbeat, analyse détaillée des attaques (RCA).

Essayez Sophos Intercept X gratuitement

Sur la page sophos.fr/intercept-x

Équipe commerciale France :
Tél. : 01 34 34 80 00
Email: info@sophos.fr

06/03/18 WP-FR (DD)

SOPHOS