**Sophos Intercept**

**Stopping Active Adversaries**

*An explanation of features in Intercept 2.0*

# Intercept 2.0 (Early Access Program June 2017)

With Intercept 2.0 the objective is to add machine learning to detect malicious executables and a number of mitigations for active adversary techniques. An active adversary is someone who has gained some access to device, a toe hold, and is looking to extend their compromise of the device to steal authentication credentials, elevate their privileges or establish persistence on the device. In many of these scenarios it is critical to understand that penetration has happened to some degree and the device is compromised. The objective is to detect when this has happened, terminate the process involved in the detected threat, and generate a root cause analysis report to facilitate additional administration actions.

The June Early Access Program(EAP) includes the active adversary mitigations and enablement of the techniques simply requires allowing the endpoint to participate in the EAP. Additional management controls for active adversary mitigations and full integration into incident report generation will not be included in the initial release of the EAP, these controls come in a future update. Later this year the EAP program will turn on the machine learning based malware detection.

During the EAP we encourage customers and partners to test the capabilities and to determine if they are detecting the attacks correctly without false positives. To assist with testing a test document is provided that explains how to trigger the mitigation and for many of the techniques a Sophos Test tool will be provided. Customers and partners familiar with penetration testing tools like Metasploit, Kali-Linux and other ‘hacking tools’ will be able to experiment and confirm the mitigations are effective against new and more sophisticated attack scenarios. The Early Access Program agent is designed to work on Windows Desktops, support for Windows servers will come in a future release.

## The June EAP for Intercept 2.0 includes the following new features

* **Credential theft protection** – Preventing theft of authentication passwords and hash information from memory, registry and off the hard disk.
* **New process protection techniques**
	+ **Code cave utilization** – Detects the presence of code deployed into another application, often used for persistence and AV avoidance.
	+ **Malicious process migration** – This detects a remote reflective dll injection used by adversaries to move laterally between processes running on the system.
	+ **Process privilege escalation**– This prevents a low privilege process from being escalated to a higher privilege, often used by an active adversary to gain system access rights.
	+ **APC protection (Atom bombing)** – This detects abuse of Application Procedure Calls often used as part of the new (2016) Atom Bombing exploit technique. Adversaries use this to get another process to execute their code.
* **New registry protections**
	+ **Sticky key protection** – Intercept 2.0 prevents replacement of the sticky key executable by an adversary often used for persistence.
	+ **Application verifier protection** – Intercept 2.0 prevents the replacement of application verifier dlls that would allow the adversary to circumvent AV and other normal process start up behaviour.
* **Improved process lockdown**
	+ **Browser behaviour lockdown** – Intercept 2.0 prevents the malicious use of PowerShell from browsers as a basic behaviour lockdown.
	+ **HTA application lockdown** – HTML applications loaded by the browser will have the lockdown mitigations applied as if they were a browser.

Level of notification, RCA, Alert/Event, heath state and if the offending process is killed or not

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Technique Detected**  | **User Notification** | **RCA** | **Admin action required** | **Security Health State** | **Process Terminated** |
| Credential theft | YES | YES | YES (ALERT) | RED | YES |
| Code cave | YES | YES | NO (Event) | GREEN | YES |
| Remote reflective DLL injection | YES | YES | YES (ALERT) | RED | YES |
| Privilege escalation | YES | YES | YES (ALERT) | RED | YES |
| APC protection | YES | YES | NO (Event) | GREEN | YES |
| Sticky keys | NO | NO | NO  | GREEN | N/A |
| Application verifier | NO | NO | NO  | GREEN | N/A |
| Lockdown (Browser powershell) | YES | YES | NO (Event) | GREEN | YES |
| Lockdown (ATA from browser) | YES | YES | NO (Event) | GREEN | YES |

### Integration with Synchronized Security

Some of the detection methods provide a strong indicator that an active adversary has penetrated the device and is still present. When this is determined the endpoint will be placed in a RED security health state.

When the endpoint is deployed with the Sophos XG Firewall and synchronized security is enabled, isolation from critical resources and the internet can be enforced by the Sophos XG Firewall policy.

When deployed with the Sophos SGN File encryption product and synchronized security is enabled a RED security health state will revoke encryption keys from the device.

NOTE: We are also working on a device isolation feature that will allow the endpoint to self-isolate itself from other devices and the internet when in a RED Security health state. Device self-isolation is currently targeted for late 2017 or early 2018.

### Competitive Anti-virus products

During development all of these attack techniques were tested with multiple traditional AV products and in almost every case the existing active adversary technique was un-detected by the AV product. As always Sophos Intercept 2.0 can run alongside competitive AV products or as a single integrated agent when deployed with Sophos Central Endpoint Advanced and managed from Sophos Central.

### Is your business a target

Suffice it to say everyone is vulnerable and given the robust infrastructure available to criminal syndicates, hacktivists and nation states, everyone is a target.

Given the proliferation of exploit kits and the hacking tools currently available we have reached a point where even un-sophisticated criminals can easily automate very advanced attacks. To complete an attack the adversary often has to string together a number of individual attacks successfully, such as initial exposure to the device (Getting a user to authorize something to run or browse to an infected website, insert a USB stick with malware etc). Once infected the software communicating back to the command and control server may not have sufficient privileges to complete the next stage of the attack, so additional techniques have to be used to elevate privileges, exploit other processes and establish a persistent and un-detected presence on the device. Once all of that is complete they can now deploy the malware that performs the ransom attack, steals the credentials, established the key logger etc. In advanced attacks the adversary may do all of this from a console, deploying each tactic and tool separately as the probe the defences and persist in their assault. When Intercept detects any of these tactics the malicious process is terminated and a root cause analysis report is generated, depending on the specific technique the machine may be placed in a RED security health state and isolated to prevent further damage.

For bulk attacks the whole effort can be scripted and fully automated, and this is where most small business fall victim. Automation is easy so with just a few clicks the criminals can target thousands of users and get a toe hold on hundreds of machines at multiple companies and establish persistence. Once the adversary has a good profile of the devices and companies compromised they then determine if they should proceed with a more sophisticated hacking effort, simply deploy a ransomware attack and wait for payment. , or sell the resource to another criminal.

The 2015 CSO article “Why criminals pick on small business” is truer today than it was then. Small businesses are easy targets, and cyber crime has been automated.

# *“Why criminals pick on small business*

## *Small and midsized businesses are now the preferred targets for cybercriminals – not because they are lucrative prizes individually but because automation makes it easy to attack them by the thousands, and far too many of them are easy targets.”*

## <http://www.csoonline.com/article/2866911/cyber-attacks-espionage/why-criminals-pick-on-small-business.html>

In Oct 2016 we see the same warning from a news article….

“*Small and mid-sized businesses are hit by 62 percent of all cyber-attacks, about 4,000 per day, according to IBM. Cybercriminals target small businesses because they are an easy, soft target to penetrate. They steal information to rob bank accounts via wire transfers; steal customers’ personal identity information; file for fraudulent tax refunds; and, commit health insurance or Medicare fraud.”*

<http://www.denverpost.com/2016/10/23/small-companies-cyber-attack-out-of-business/>

## Credential Theft Prevention

Intercept 2.0 detects when an adversary controlled process is attempting to extract user and administrator authentication credentials from the device. An adversary attempting credential theft can target multiple operating system components to steal the password or the hashed passwords of users and administrators for the device. Dozens of different tools are available to the adversary to achieve this but the most commonly used include mimikats – a credential extraction tool that targets LSASS (Local Security Authority Subsystem Service) memory and hashdump – a credential theft tool that extracts the hashed password from the SAM(Security Account Manager) database.

### How does Intercept 2.0 prevent

Instead of targeting the specific tools used by adversaries (there are lots of them) Intercept 2.0 instead looks for unauthorized interactions with the LSASS runtime memory, the SAM DB registry and direct extraction of credential data from the hard disk. As a prevention technique we have tested with a variety of malware, penetration, and hacking tools and found the mitigation to be extremely effective without generating false positive alerts for legitimate software that interacts with the LSASS and SAM DB.

### What happens when an attack is detected

When Intercept 2.0 detects an adversary attempted credential theft the process performing the attack will be terminated, and a notification to the end user will be presented.

This will also initiate a Root Cause Analysis, and will alert the administrator of the activity so it can be investigated. *(Not available during June EAP, will be available when shipped)*

The endpoint will be in a RED security health state until the administrator clears the alert notification after investigating. *(Not available during June EAP, will be available when shipped)*

### What should an admin do

The attack was detected at run time and though the attacking process was terminated the initial penetration technique could be repeated or the attacker may still have access to the device. Penetration of the device often involves tricking the end user to authorize the install of malicious software, or the enablement of macros or other actions to by the end-user, but in some instances the penetration involved no direct authorization by the end user.

The detection will generate an alert to inform the administrator that the credential theft attempt was detected and further examination of the incident is warranted. To aid in the investigation this detection will also request the generation of an incident report using Intercept 2.0’s root cause analysis capability. *(Not available during June EAP, will be available when shipped)*

## Process Protection (Code cave)

Code cave utilization is a technique used by adversaries where they modify what is likely legitimate software so that it contains an additional application. This additional application is inserted into what is called a code cave, a section of the target applications file that is un-used by the program. Code caves exist in most applications and adding code to these sections should not break the behaviour of primary application. Often the execution code inserted into a code cave is simply a remote shell launcher; these can be very small and simply grant the adversary access to the box where they can perform other actions. This type of attack requires the adversary to have established a presence on the device so they can deploy the software or to trick the user to download and install an application that has the code cave already exploited.

One of the primary reasons adversaries use code caves is to hide from detection by the general user and administrators, the expected application still works fine, and the inserted application is also running. If the application that has been modified is a legitimate business tool that the administrator expects to be on the device they are less likely to consider it malware if traditional AV detects a problem and administrators may simply add it to the exemption list, assuming the AV engine has generated a false positive. In this way the adversary establishes persistence on the endpoint and may have even tricked the admin to allow them to run.

### How does Intercept 2.0 prevent

A number of tools exist that can utilize the code cave technique to embed software into another application and most traditional anti-virus simply looks for the tell-tale indicators or signatures that these tools leave behind when they insert code into the code cave. For Intercept 2.0 we did not want to follow that approach and instead evaluate applications for any code cave utilization, this is done at initial execution of the software and when we detect the presence of the additional application residing in a code cave we terminate the application.

### What happens when an attack is detected

On detection of the use of a code cave the application will be terminated and the user notified.

This will also initiate a root cause analysis, and will alert the administrator of the activity so it can be investigated. *(Not available during June EAP, will be available when shipped)*

Sophos Clean will then remove the malware from the device. *(Not available during June EAP, will be available when shipped)*

### What should an admin do

On detection of a code cave utilization the administrator should check the RCA to determine how the infected application was deployed to the device, it may be that the adversary had already compromised the device by another means and was simply deploying the code cave to ensure persistence on the device. With this attack blocked the adversary is likely looking for other avenues of attack and persistence. If this was an end user that was tricked into downloading an application with a code cave it is likely the attack has been prevented, but understanding how they attempted to penetrate the device will help determine what training is required or if additional policy controls need to be put in place.

## Process Protection (Malicious migration – remote reflective DLL injection)

Process migration is a common technique performed by an adversary when they first establish presence on a device and want to move to another process to either escalate privileges or gain more enduring access. The adversary does not want to lose control when the end user simply closes their browser or terminates a process that has been compromised, so migrating to a system process is desired.

Migration techniques can leverage a remote reflective DLL injection. For more information on DLL injections in general MITRE provides a great resource. <https://attack.mitre.org/wiki/Technique/T1055> A remote reflective DLL attack is similar, but harder. Suffice it to say the adversary has already compromised one process and from there they are manipulating another process to load dlls, and run arbitrary code.

### How does Intercept 2.0 prevent

Intercept 2.0 is monitoring process activity for the behaviour of allocating memory in a remote process and the injection of dlls into that process. This behaviour is not something that should be happening and when we detect the behaviour we have high confidence it is malicious and indicates an active adversary or malware script running on the compromised system.

### What happens when an attack is detected

When Intercept 2.0 detects an adversary attempting to migrate to another process in this way the attacking process will be terminated, and a notification to the end user will be presented.

This will also initiate a root cause analysis, and will alert the administrator of the activity so it can be investigated. *(Not available during June EAP, will be available when shipped)*

The endpoint will be in a RED security health state until the administrator clears the alert notification after investigating. *(Not available during June EAP, will be available when shipped)*

### What should an admin do

Because the attack was detected at run time it is possible that an adversary is still active on the device, and though the attacking process was terminated the initial penetration technique could be repeated or the attacker may still have access from another process

The detection will also generate an alert to inform the administrator that process migration with remote reflective DLL injection was detected and further examination of the device is warranted. To aid in the investigation this event will also request the generation of an incident report using Intercept 2.0’s root cause analysis capability. *(Not available during EAP, will be available when shipped)*

## Process Protection (Privilege escalation)

When an adversary has gained access to a system often they are not running at the privilege level they want or need to complete the rest of their attack. A number of methods exist to elevate privileges for the adversary from credential theft to process migration but with these doors now locked by Intercept 2.0 the adversary has to use other techniques. One that comes to mind is stealing the authentication token of a privilege process and inserting it into another process to elevate privileges.

All processes running on the device have an authentication token that the operating system uses to determine the privileges of the process. With this technique the adversary is likely looking to steal the authentication token of a SYSTEM process. If an adversary can steal the authentication token of a process with system privilege and use it they have what they want and did not need to crack the admin user password or perform a process migration to get it. By taking advantage of known system kernel vulnerabilities in unpatched windows devices the adversary has a number of well document techniques to capture a privileged token from a process and use it for their own purpose. Given the number of methods available for privileged token theft it is likely more as yet unknown vulnerabilities in the operating system and kernel remain.

### How does Intercept 2.0 prevent

Instead of trying to protect from the numerous known vulnerabilities that allow privileged token theft Intercept 2.0 is instead looking for when a process has a privileged authentication token inserted into it to elevate privileges. This behaviour is simply not used by legitimate software and when spotted we can be fairly sure it is an active adversary attack. By detecting this escalation of privilege Intercept 2.0 is able to protect against this technique regardless of what vulnerability, known or unknown, was used to steal the authentication token in the first place.

### What happens when an attack is detected

We will terminate the process and notify the end user.

This will also initiate a sophos clean to remove the malware.

On detection an RCA will be generated to determine how the attacking process started and what else may have been happening on the device that is related to the root cause or detected escalation. *(Not available during June EAP, will be available when shipped)*

The endpoint will be put into a RED Security health state as this attack indicates an adversary has likely penetrated the device and more investigation is recommended. *(Not available during June EAP, will be available when shipped)*

### What should an admin do

Like similar exploit prevention detections, administrators should review the RCA report to determine how the attack unfolded and where it came from.

Once the investigation is complete the administrator can clear the alert to allow normal operation of the device.

## Process Protection (malicious APC use – Atom Bombing)

Atom Bombing is a technique used by adversaries to trick another application into running malware or other code. The technique is fairly complex and new and involves abuse of the operating systems ATOM tables and asynchronous procedure calls. See <https://breakingmalware.com/injection-techniques/atombombing-brand-new-code-injection-for-windows/> for a better description than I can write.

### How does Intercept 2.0 prevent

Intercept 2.0 is looking for abuse of APC calls. Like many of the exploit protection methods already available in Intercept X, the product is able to monitor process activity at the kernel level and as far as we can see this type of behaviour is never good.

### What happens when an attack is detected

We will terminate the exploiting application and notify the end user.

This will also initiate a sophos clean to remove the malware and trigger a root cause analysis evaluation to determine how the attacking process started and what else may have been happening.

### What should an admin do

Like similar exploit prevention detections, administrators should review the RCA report to determine how the attack initiated and if other actions are required.

## Registry Protection (Sticky Key persistence)

[Sticky keys](https://en.wikipedia.org/wiki/Sticky_keys) is a feature of windows operating systems. This feature will launch an application when the user presses shift 5 consecutive times. The application to be launched is identified in the registry. Most organization will simply disable this windows feature, but when it has not been disabled the adversary can use this registry trick to launch whatever they want. To make matters worse, sticky-keys works from the login page and the application will launch with system privileges. To make this registry change the adversary needs to have access to the device or have an application they want run executed by the end-user.

### How does Intercept 2.0 prevent

Using Sticky keys and a number of other registry modifications to establish persistence on a device is a old-school technique used by hackers for years. Intercept 2.0 simply disables the ability to change the windows sticky-key executable. This prevents the adversary from using the feature to start malware or to launch an otherwise legitimate remote shell connection.

### What happens when an attack is detected

We do not notify the end-user or generate an alert when the registry has been modified to launch another application on sticky key activation; we simply ensure the authorized MS windows utility is launched.

## Registry Protection (Application verifier mitigation - Double Agent)

This is another registry trick that adversaries have available in their tool kit. The attack involves modification of the registry to identify software that should run whenever an application is started. The feature from Microsoft is intended to enable developers to monitor and diagnose application activity, but when used by an adversary it is often to ensure that they have access to the box and can circumvent the protection capabilities of the application being run. This attack made the [news in 2017](https://cybellum.com/doubleagent-taking-full-control-antivirus/) when it was noted that many anti-virus products were susceptible to having the registry for the anti-virus software is modified to run an adversaries application as well. In reality the attack is much broader than just targeting AV products; an application verifier registry change can be used for any application on the operating system. See a sophos naked security article for more information. <https://nakedsecurity.sophos.com/2017/03/27/doubleagent-vulnerability-just-how-bad-is-it/>

### How does Intercept 2.0 prevent

Intercept 2.0 will enforce the authorized windows dlls when application verification is used. In this way even if the adversary managed to tamper the registry and set it to launch their attack the application will instead ignore these illegitimate registry changes.

It is important to note that when Intercept 2.0 is deployed alongside a competitors AV we will protect that AV product from attacks that use the Double Agent (Application Verifier) technique.

### What happens when an attack is detected

We do not notify the end-user or generate an alert when the registry has been modified to launch another application on sticky key activation; we simply ensure the authorized MS windows utility is launched.

## Improved Process Lockdown (browsers and HTML applications)

Intercept 2.0 already includes Process Lockdown where we prevent various malicious behaviours of identified process types. With Intercept 2.0 we are extending the lock down capability to prevent web browsers from launching power shell and extending the browser lockdown capability to HTML applications that are run by the browser ([HTA applications](https://en.wikipedia.org/wiki/HTML_Application)).

### How does Intercept 2.0 prevent

Intercept 2.0 will auto classify an application as a browser and when so classified application lockdown is enforce to prevent malicious behaviours, like powershell . Lockdown leverages the fact that Intercept 2.0 is able to monitor the applications activity at the kernel and is always running with the application.

### What happens when an attack is detected

When Intercept 2.0 detects an application misbehaving in this way it is prevented from completing the activity and the user is notified.

An event is generated for the administrator to review

Detection of this type of malicious behaviour will also request the generation of an RCA for review by the admin. *(Not available during June EAP, will be available when shipped)*