How to write malware and learn how to fight it!

Antonio ‘s4tan’ Parata

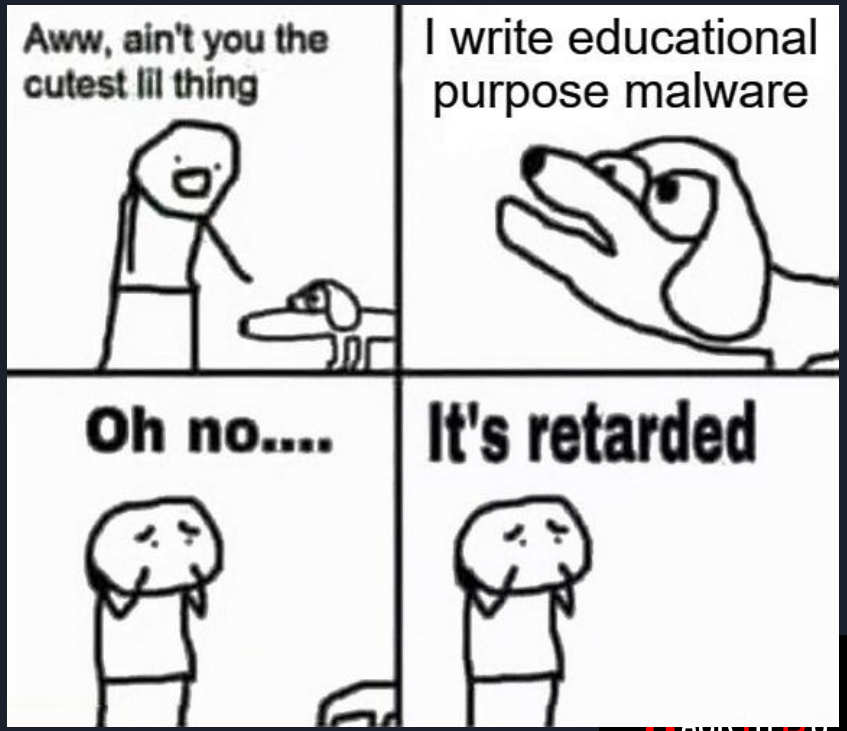
HACK IN BO®
Winter 2019 Edition
13ª EDIZIONE

Disclaimer

This presentation is not intended to teach to the bad guys how to write malware. There are already too many “education purpose projects” in GitHub, we don’t need another one :)

The goal of the presentation is to show how to analyze malicious code by considering how a malware author think.

But remember… CODING IS NOT A CRIME!



Disclaimer

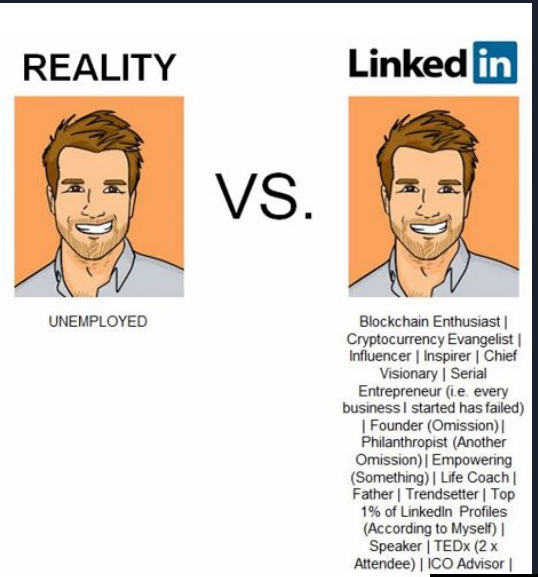
This presentation is not intended to teach the bad guys how to write malware. There are already too many “education purpose projects” in GitHub, we don’t need another one :)

The goal of the presentation is to show how to analyze malicious code by considering how a malware author think.

HACK IN BO
Winter 2019 Edition
13ª EDIZIONE

whoami.exe

We have more Cyber-Security guru on LinkedIn than IPv4 addresses



HACK IN BO Winter 2019 Edition
13ª EDIZIONE

whoami.exe

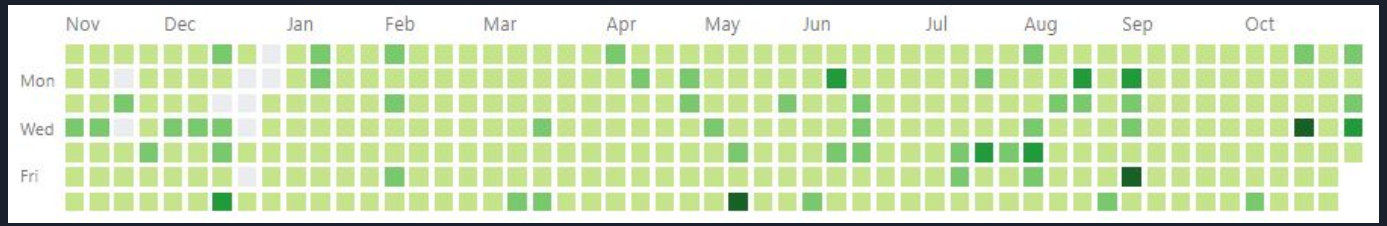
01 Fourth time attendee at HackInBo (three as speaker)

02 Senior Security Researcher CrowdStrike

03 Owasp Italy Board since 2006

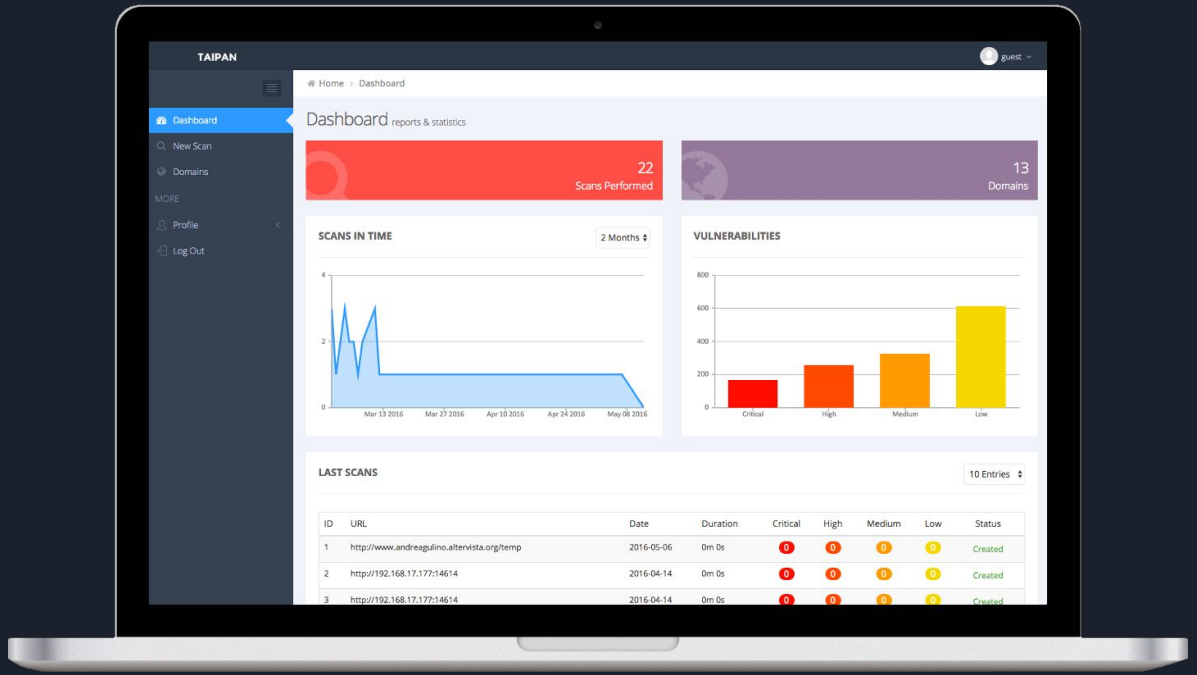
04 Phrack Author
http://www.phrack.org/papers/dotnet_instrumentation.html

05 Passionate F# developer
https://github.com/sponsors/enkomio GitHub Sponsors



HACK IN BO Winter 2019 Edition 13ª EDIZIONE

whoami.exe

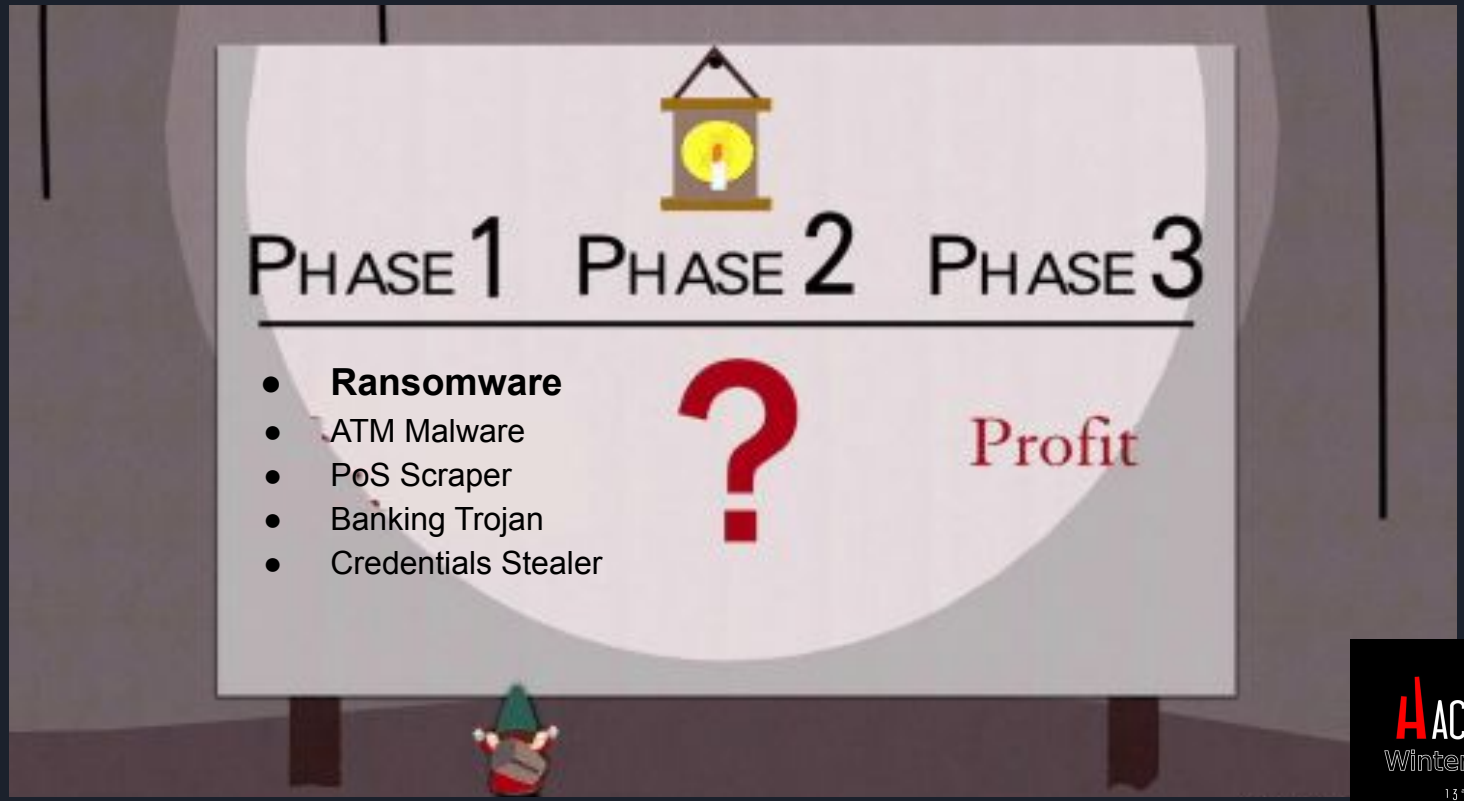


Taipan Web Vulnerability Scanner - https://taipansec.com

Cyber-Crime

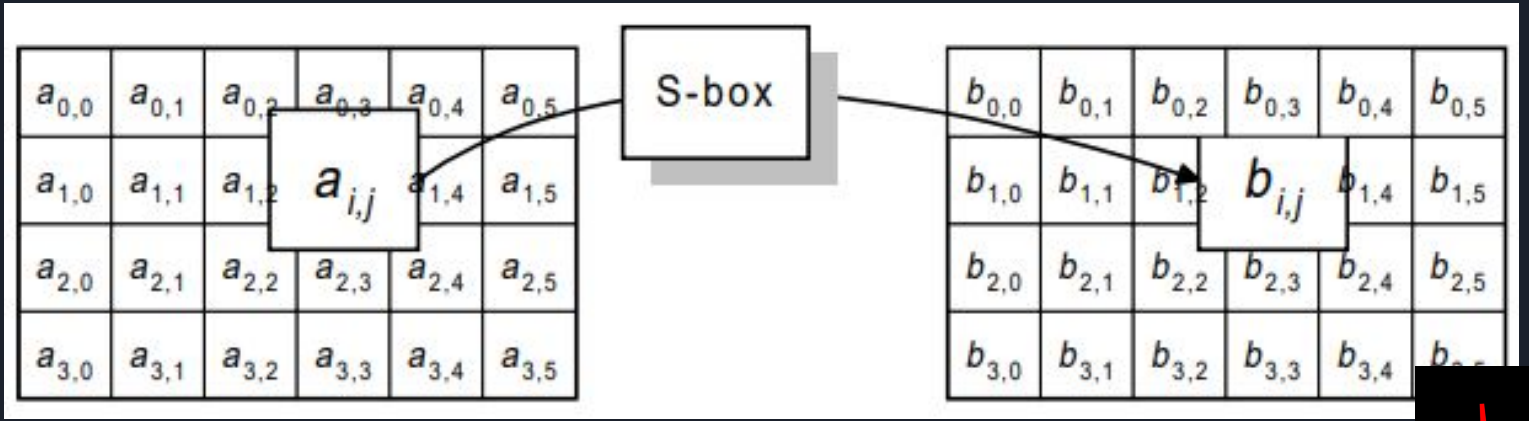
- We are not talking about amateur malware (skiddies writing a .NET RAT and posting it on HackForums)
- Professional cyber-criminal are very well organized:
  - They have a dedicated GIT repository
  - A testing botnet
  - A customer support platform (typically in form of Jabber chat)
  - A crypto service to evade AVs
  - They use a bulletproof hosting provider for their botnet
  - VPN service to hide his/her real IP
  - A distribution network (SPAM)
  - A mule network (monetization)

How to write a malware and make money



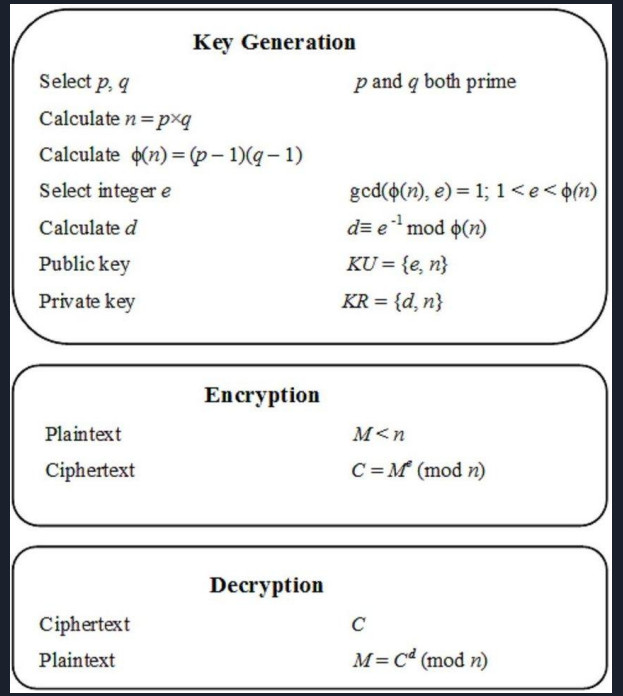
Reversing AES

Pretty easy if S-Box is not obfuscated, just use FindCrypt(2) IDA plugin to identify the code that use the S-Box



Reversing RSA

- **No hard coded constants but...**
- From Wikipedia:
  - *the most commonly chosen value for e is 216 + 1 = 65,537*
- So, if you find very weird math operations involving:
  - Two numbers
  - One of them is very big
  - The other number is 65537 (0x10001)
- Maybe you found an RSA encryption routine!



Key Generation

Select p, q	p and q both prime
Calculate $n = p \times q$	
Calculate $\phi(n) = (p-1)(q-1)$	
Select integer e	$\gcd(\phi(n), e) = 1;\ 1 < e < \phi(n)$
Calculate d	$d \equiv e^{-1} \bmod \phi(n)$
Public key	$KU = \{e, n\}$
Private key	$KR = \{d, n\}$

Encryption

Plaintext	$M < n$
Ciphertext	$C = M^e \pmod n$

Decryption

Ciphertext	C
Plaintext	$M = C^d \pmod n$

4. Choose an integer e such that $1 < e < \lambda(n)$ and $\gcd(e, \lambda(n)) = 1$; that is, e and $\lambda(n)$ are coprime.
   - e having a short bit-length and small Hamming weight results in more efficient encryption – the most commonly chosen value for e is $2^{16} + 1 = 65,537$. The smallest (and fastest) possible value for e has been shown to be less secure in some settings.[14]
   - e is released as part of the public key.

Reverse Engineering

What means being a *reverser*?

- Be able to code
- Knowledge about OS
- Knowledge about computer architecture
- Be able to read machine code

Reversing like a PRO

```
00406936 | 64:A1 30000000 | mov eax,dword ptr fs:[30]
0040693C | 8B40 0C        | mov eax,dword ptr ds:[eax+C]
0040693F | 8B40 0C        | mov eax,dword ptr ds:[eax+C]
00406942 | 8B00           | mov eax,dword ptr ds:[eax]
00406944 | 8B00           | mov eax,dword ptr ds:[eax]
00406946 | 8B40 18        | mov eax,dword ptr ds:[eax+18]
00406949 | C3             | ret
```

Move to EAX the value of FS[30]

Move to EAX the value at address EAX+C

Move to EAX the value at address EAX+C

Move to EAX the value at address EAX

Move to EAX the value at address EAX

Move to EAX the value at address EAX + 18

return

C0ngratz u r now an 31337 hax0r!!1

HACK IN BO Winter 2019 Edition 13ª EDIZIONE

Reversing like a ~~PRO~~ cat

```
00406936 | 64:A1 30000000 | mov eax,dword ptr fs:[30]
0040693C | 8B40 0C        | mov eax,dword ptr ds:[eax+C]
0040693F | 8B40 0C        | mov eax,dword ptr ds:[eax+C]
00406942 | 8B00           | mov eax,dword ptr ds:[eax]
00406944 | 8B00           | mov eax,dword ptr ds:[eax]
00406946 | 8B40 18        | mov eax,dword ptr ds:[eax+18]
00406949 | C3             | ret
```

Move to EAX the **PEB** address from **TEB**

Move to EAX the **Ldr** address

Move to EAX the **InLoadOrderModuleList** address

Move to EAX the **FLink** from LIST_ENTRY

Move to EAX the **FLink** from LIST_ENTRY

Move to EAX the **DllBase** of the library

Return the DllBase

Program name

ntdll.dll

kernel32.dll

This function resolves the base address of Kernel32. If you think that it's done in order to walk the EAT (Export Address Table) and to resolve the desider function address…

...

you are right! (more soon...)

One more Reversing exercise

```
0040C8F8 | 56                      | push esi
0040C8F9 | 8D04C5 88124000         | lea eax,dword ptr ds:[eax*8+401288]
0040C900 | 33C9                    | xor ecx,ecx
0040C902 | 33F6                    | xor esi,esi
0040C904 | 66:3B48 02              | cmp cx,word ptr ds:[eax+2]
0040C908 | 73 15                   | jae kpot2.0.40C91F
0040C90A | 8B50 04                 | mov edx,dword ptr ds:[eax+4]
0040C90D | 0FB7CE                  | movzx ecx,si
0040C910 | 8A140A                  | mov dl,byte ptr ds:[edx+ecx]
0040C913 | 3210                    | xor dl,byte ptr ds:[eax]
0040C915 | 46                      | inc esi
0040C916 | 881439                  | mov byte ptr ds:[ecx+edi],dl
0040C919 | 66:3B70 02              | cmp si,word ptr ds:[eax+2]
0040C91D | 72 EB                   | jb kpot2.0.40C90A
0040C91F | 0FB740 02               | movzx eax,word ptr ds:[eax+2]
0040C923 | C60438 00               | mov byte ptr ds:[eax+edi],0
0040C927 | 5E                      | pop esi
0040C928 | C3                      | ret
```

1 2 3 4 5 6 7

```
00401288  C3 00 13 00 94 35 40 00 A6 00 11 00 80 35 40 00  Ã....5@.¦....5@.
00401298  C3 00 10 00 6C 35 40 00 79 00 0F 00 5C 35 40 00  Ã...l5@.y...\5@.
004012A8  84 00 12 00 48 35 40 00 A8 00 13 00 34 35 40 00  ....H5@.¨...45@.
004012B8  70 00 13 00 20 35 40 00 8F 00 13 00 0C 35 40 00  p... 5@......5@.
004012C8  3E 00 1B 00                          34 40 00  >...ð4@.....Ô4@.
004012D8  FA 00 13 00 typedef struct data       34 40 00  ú...À4@.....¬4@.
004012E8  76 00 19 00 {                         34 40 00  v....4@.Ë....4@.
004012F8  67 00 0B 00     uint8_t key;          34 40 00  g...t4@.....d4@.
00401308  D2 00 04 00     uint16_t length;      34 40 00  Ò...\4@.-...T4@.
00401318  18 00 04 00     uint32_t buffer;      34 40 00  ....L4@.Ò...D4@.
00401328  EA 00 0D 00                           34 40 00  ê...44@.....$4@.
00401338  CB 00 08 00 } data_t;                 34 40 00  Ë....4@......4@.
00401348  20 00 08 00 00 34 40 00 40 00 04 00 F8 33 40 00   ....4@.@...ø3@.
00401358  1F 00 05 00 F0 33 40 00 10 00 04 00 E8 33 40 00  ....ð3@.....è3@.
00401368  5D 00 08 00 DC 33 40 00 3E 00 07 00 D4 33 40 00  ]...Ü3@.>...Ô3@.
00401378  85 00 13 00 C0 33 40 00 D3 00 0B 00 B4 33 40 00  ....À3@.Ó...´3@.
00401388  76 00 0B 00 A8 33 40 00 4C 00 08 00 9C 33 40 00  v...¨3@.L....3@.
```

Any idea?

Decompiler FTW!

- Decompilers (like Hex-Rays, Ghidra, ILSpy, ...) are able to translate machine-code in pseudo code like C or C#.
- This make the RCE task way easier!
- Unfortunately bad guys know this and they use obfuscators or other anti-analysis tricks to avoid decompilation

© Rolf Rolles: Automation Techniques in C++ Reverse Engineering

```
void __fastcall sub_17142D60(minsn_t *a1, minsn_t *a2)
{
  mop_t *v3; // rbp
  mop_t *v4; // rsi

  if ( a2 != a1 )
  {
    v3 = &a2->l;
    v4 = &a1->l;
    if ( &a2->l != &a1->l )
    {
      sub_17144EB0(&a1->l);
      sub_17142E10(v4, v3);
    }
    if ( &a2->r != &a1->r )
    {
      sub_17144EB0(&a1->r);
      sub_17142E10(&a1->r, &a2->r);
    }
    if ( &a2->d != &a1->d )
    {
      sub_17144EB0(&a1->d);
      sub_17142E10(&a1->d, &a2->d);
    }
    a1->ea = a2->ea;
    a1->opcode = a2->opcode;
    a1->iprops = a2->iprops;
  }
}
```

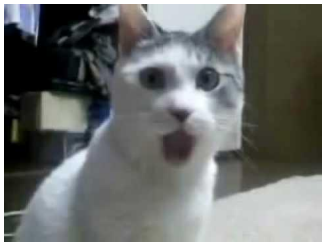
.NET decompilers

Original

```
class Program
{
    0 references
    static void Main(string[] args)
    {
        if (args.Length < 1)
        {
            Console.WriteLine("Please specify the program to extract resources.");
            return;
        }

        var filename = args[0];
        var assembly = Assembly.LoadFile(filename);

        var extractedResourceDirectory = "extractedResources";
        Directory.CreateDirectory(extractedResourceDirectory);

        foreach (var resourceName in assembly.GetManifestResourceNames())
        {
            var resourceDirectory = Path.Combine(extractedResourceDirectory, resourceName);
            Directory.CreateDirectory(resourceDirectory);

            var cleanResourceName = resourceName.Replace(".resources", String.Empty);
            var resourceManager = new ResourceManager(cleanResourceName, assembly);
            var assemblyName = assembly.GetName();

            var resourceSet = resourceManager.GetResourceSet(assemblyName.CultureInfo, true, true).OfType<DictionaryEn
            foreach (var dictionaryEntry in resourceSet)
            {
                var resKey = dictionaryEntry.Key.ToString();
                var resValue = dictionaryEntry.Value;
                var formatter = new BinaryFormatter();
                var memoryStream = new MemoryStream();
                formatter.Serialize(memoryStream, resValue);

                var base64Value = Convert.ToBase64String(memoryStream.GetBuffer());
                var resFilename = Path.Combine(resourceDirectory, resKey);
                File.WriteAllText(resFilename, base64Value);
            }
        }
    }
}
```

Decompiled

```
private static void Main(string[] args)
{
    if (args.Length < 1)
    {
        Console.WriteLine("Please specify the program to extract resources.");
    }
    else
    {
        string path = args[0];
        Assembly assembly = Assembly.LoadFile(path);
        string str2 = "extractedResources";
        Directory.CreateDirectory(str2);
        foreach (string str3 in assembly.GetManifestResourceNames())
        {
            string str4 = Path.Combine(str2, str3);
            Directory.CreateDirectory(str4);
            ResourceManager manager = new ResourceManager(str3.Replace(".resources", string.Empty), assembly);
            AssemblyName name = assembly.GetName();
            IEnumerable<DictionaryEntry> enumerable = manager.GetResourceSet(name.CultureInfo, true, true).OfType<DictionaryEntry>();
            foreach (DictionaryEntry entry in enumerable)
            {
                string str6 = entry.Key.ToString();
                object graph = entry.Value;
                BinaryFormatter formatter = new BinaryFormatter();
                MemoryStream serializationStream = new MemoryStream();
                formatter.Serialize(serializationStream, graph);
                string contents = Convert.ToBase64String(serializationStream.GetBuffer());
                File.WriteAllText(Path.Combine(str4, str6), contents);
            }
        }
    }
}
```

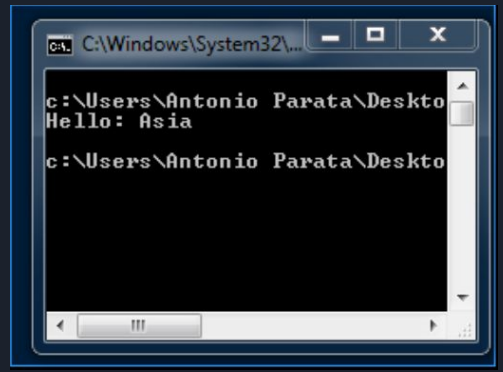


Breaking .NET decompilers

IL_0014: nop

IL_0015: ldarg.0 // pointer to ***this*** argument, this value is expected by instance methods

IL_0016: call instance void ConsoleApplication.SimpleClass::SayHello()



Assemble

Decompile

IL_0014: br.s IL_0017

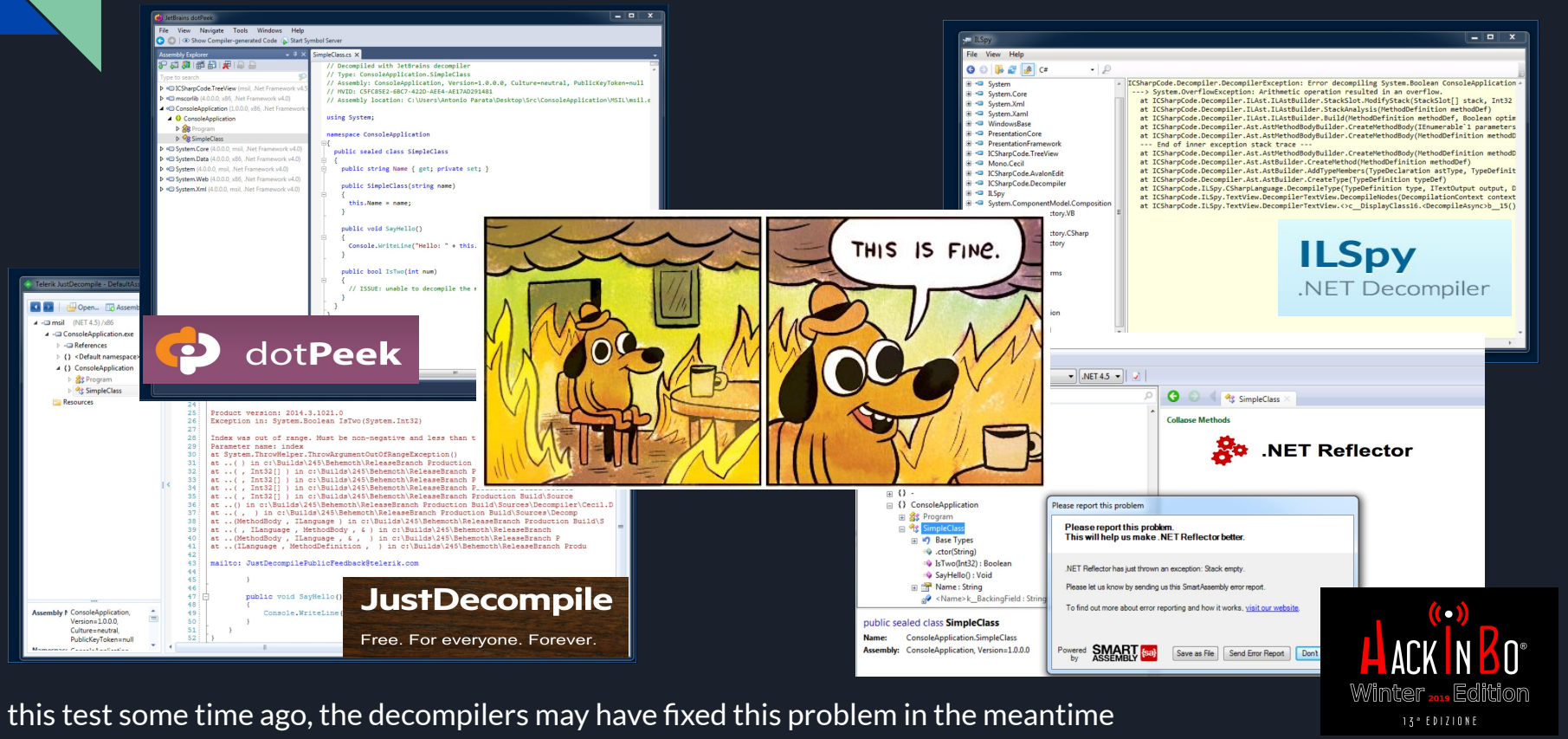
IL_0015: ldarg.0 // remove the push of the ***this*** argument and add a jump in order to avoid the call

IL_0016: call instance void ConsoleApplication.SimpleClass::SayHello()

IL_0017: nop

HACK IN BO Winter 2019 Edition 13ª EDIZIONE

Breaking .NET decompilers



I did this test some time ago, the decompilers may have fixed this problem in the meantime

Anti-analysis - IDA Hex-Rays decompiler

```
.586
.model flat,stdcall
.stack 4096
.DATA
.CODE
main PROC
        push ebp
        mov ebp, esp

        push 5
        call secret_algo
        sub esp, 4

        push 5
        call secret_algo_obf
        sub esp, 4

        mov esp, ebp
        pop ebp
        ret
main ENDP
END main
```

```
secret_algo PROC
        push ebp
        mov ebp, esp
        rdtsc
        cmp edx, 0
        ja real_code
        jmp dword ptr [switch_table + edx * type dword]
case0:
        mov eax, 0
        jmp real_code
case1:
        mov eax, 1
        jmp real_code
real_code:
        ; start secret algo code
        mov edx, dword ptr [ebp+8]
        xor edx, 0C0D3CA05h
        mov eax, edx
        mov esp, ebp
        pop ebp
        ret
switch_table:
        dword case0
        dword case1
secret_algo ENDP
```

```
secret_algo_obf PROC
        push ebp
        mov ebp, esp
        rdtsc
        cmp edx, 0
        ja real_code
        jmp dword ptr [real_code + edx * type dword]
case0:
        mov eax, 0
        jmp real_code
case1:
        mov eax, 1
        jmp real_code
real_code:
        ; start secret algo code
        mov edx, dword ptr [ebp+8]
        xor edx, 0C0D3CA05h
        mov eax, edx
        mov esp, ebp
        pop ebp
        ret
secret_algo_obf ENDP
```

HACK IN BO® Winter 2019 Edition 13ª EDIZIONE

Anti-analysis - IDA Hex-Rays decompiler

```
secret_algo PROC
        push ebp
        mov ebp, esp
        rdtsc
        cmp edx, 0
        ja real_code
        jmp dword ptr [switch_table + edx * type dword]
case0:
        mov eax, 0
        jmp real_code
case1:
        mov eax, 1
        jmp real_code
real_code:
        ; start secret algo code
        mov edx, dword ptr [ebp+8]
        xor edx, 0C0D3CA05h
        mov eax, edx
        mov esp, ebp
        pop ebp
        ret
switch_table:
        dword case0
        dword case1
secret_algo ENDP
```

```
.text:00401030 sub_401030      proc near               ; CODE XREF: .text:0040100A↑j
.text:00401030                                         ; start_0+5↓p
.text:00401030
.text:00401030 arg_0           = dword ptr  8
.text:00401030
.text:00401030                 push    ebp
.text:00401031                 mov     ebp, esp
.text:00401033                 rdtsc
.text:00401035                 cmp     edx, 0          ; switch 1 cases
.text:00401038                 ja      short def_40103A ; jumptable 0040103A default case
.text:0040103A                 jmp     jpt_40103A[edx*4] ; switch jump
.text:00401041 ; ---------------------------------------------------------------------------
.text:00401041
.text:00401041 loc_401041:                             ; CODE XREF: sub_401030+A↑j
.text:00401041                                         ; DATA XREF: .text:jpt_40103A↓o
.text:00401041                 mov     eax, 0          ; jumptable 0040103A case 0
.text:00401046                 jmp     short def_40103A ; jumptable 0040103A default case
.text:00401048 ; ---------------------------------------------------------------------------
.text:00401048
.text:00401048 loc_401048:                             ; DATA XREF: .text:00401062↓o
.text:00401048                 mov     eax, 1
.text:0040104D                 jmp     short $+2       ; jumptable 0040103A default case
.text:0040104F ; ---------------------------------------------------------------------------
.text:0040104F
.text:0040104F def_40103A:                             ; CODE XREF: sub_401030+8↑j
.text:0040104F                                         ; sub_401030+16↑j ...
.text:0040104F                 mov     edx, [ebp+arg_0] ; jumptable 0040103A default case
.text:00401052                 xor     edx, 0C0D3CA05h
.text:00401058                 mov     eax, edx
.text:0040105A                 mov     esp, ebp
.text:0040105C                 pop     ebp
.text:0040105D                 retn
.text:0040105D sub_401030      endp
.text:0040105D
.text:0040105D ; ---------------------------------------------------------------------------
.text:0040105E jpt_40103A      dd offset loc_401041    ; DATA XREF: sub_401030+A↑r
.text:0040105E                                         ; jump table for switch statement
.text:00401062                 dd offset loc_401048
.text:00401066 ; ---------------------------------------------------------------------------
```

Anti-analysis - IDA Hex-Rays decompiler

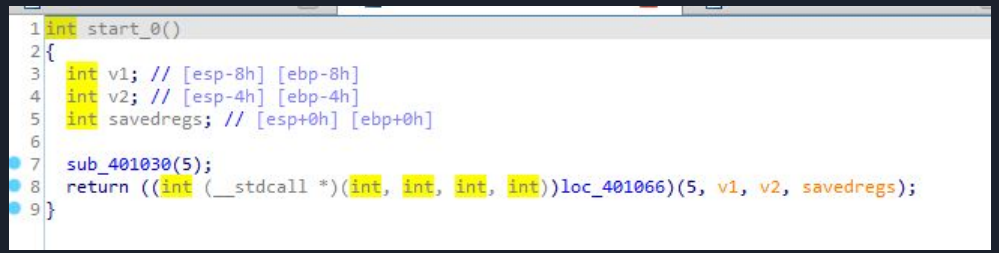
```
secret_algo_obf PROC
        push ebp
        mov ebp, esp
        rdtsc
        cmp edx, 0
        ja real_code
        jmp dword ptr [real_code + edx * type dword]
case0:
        mov eax, 0
        jmp real_code
case1:
        mov eax, 1
        jmp real_code
real_code:
        ; start secret algo code
        mov edx, dword ptr [ebp+8]
        xor edx, 0C0D3CA05h
        mov eax, edx
        mov esp, ebp
        pop ebp
        ret
secret_algo_obf ENDP
```

```
.text:00401066                                  ; start_0+F↓p
.text:00401066                  push    ebp
.text:00401067                  mov     ebp, esp
.text:00401069                  rdtsc
.text:0040106B                  cmp     edx, 0          ; switch 0 cases
.text:0040106E                  ja      short near ptr def_401070 ; jumptable 00401070 default case
.text:00401070                  jmp     def_401070[edx*4] ; switch jump
.text:00401070 ; ---------------------------------------------------------------------------
.text:00401077                  db 0B8h
.text:00401078                  dd 0
.text:0040107C                  dd 1B80EEBh, 0EB000000h, 2B807h, 0EB0000h
.text:0040108C def_401070       dd 8108558Bh            ; CODE XREF: .text:0040106E↑j
.text:0040108C                                          ; DATA XREF: .text:00401070↑r
.text:0040108C                                          ; jumptable 00401070 default case
.text:00401090                  dd 0D3CA05F2h, 8BC28BC0h
.text:00401098                  db 0E5h, 5Dh, 0C3h
.text:0040109B
.text:0040109B ; =============== S U B R O U T I N E =======================================
```

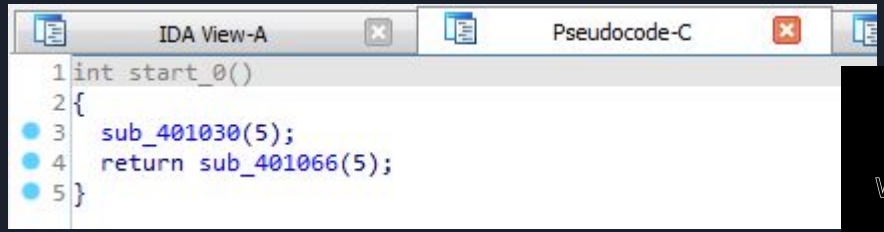
???

Anti-analysis - IDA Hex-Rays decompiler

```
.586
.model flat,stdcall
.stack 4096
.DATA
.CODE
main PROC
        push ebp
        mov ebp, esp

        push 5
        call secret_algo
        sub esp, 4

        push 5
        call secret_algo_obf
        sub esp, 4

        mov esp, ebp
        pop ebp
        ret
main ENDP
END main
```



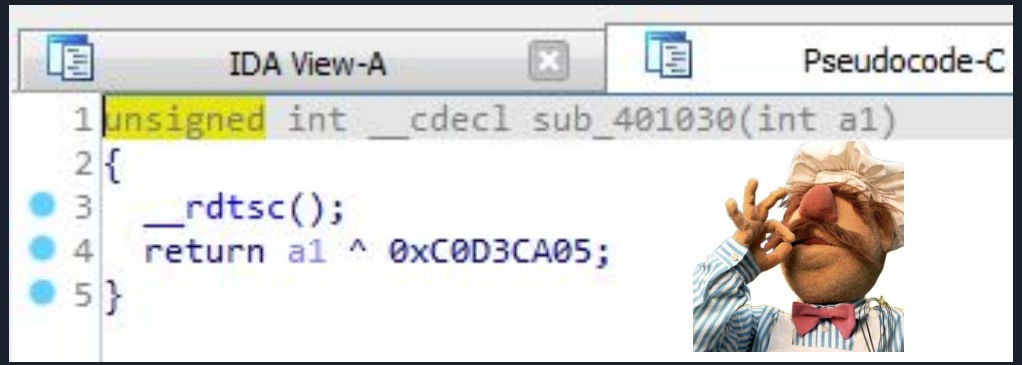
Let's give IDA some love and re-define the data as code and create a function



HACK IN BO Winter 2019 Edition 13ª EDIZIONE

Anti-analysis - IDA Hex-Rays decompiler

```
secret_algo PROC
        push ebp
        mov ebp, esp
        rdtsc
        cmp edx, 0
        ja real_code
        jmp dword ptr [switch_table + edx * type dword]
case0:
        mov eax, 0
        jmp real_code
case1:
        mov eax, 1
        jmp real_code
real_code:
        ; start secret algo code
        mov edx, dword ptr [ebp+8]
        xor edx, 0C0D3CA05h
        mov eax, edx
        mov esp, ebp
        pop ebp
        ret
switch_table:
        dword case0
        dword case1
secret_algo ENDP
```



Anti-analysis - IDA Hex-Rays decompiler

```
secret_algo_obf PROC
        push ebp
        mov ebp, esp
        rdtsc
        cmp edx, 0
        ja real_code
        jmp dword ptr [real_code + edx * type dword]
case0:
        mov eax, 0
        jmp real_code
case1:
        mov eax, 1
        jmp real_code
real_code:
        ; start secret algo code
        mov edx, dword ptr [ebp+8]
        xor edx, 0C0D3CA05h
        mov eax, edx
        mov esp, ebp
        pop ebp
        ret
secret_algo_obf ENDP
```


```
401070: switch analysis failed: switch information is incomplete or incorrect
401070: switch analysis failed: switch information is incomplete or incorrect
401070: switch analysis failed: switch information is incomplete or incorrect
401070: switch analysis failed: switch information is incomplete or incorrect
401070: switch analysis failed: switch information is incomplete or incorrect
401070: switch analysis failed: switch information is incomplete or incorrect
401070: switch analysis failed: switch information is incomplete or incorrect
401070: switch analysis failed: switch information is incomplete or incorrect
401070: switch analysis failed: switch information is incomplete or incorrect
401070: switch analysis failed: switch information is incomplete
401070: switch analysis failed: switch information is incomplete
```

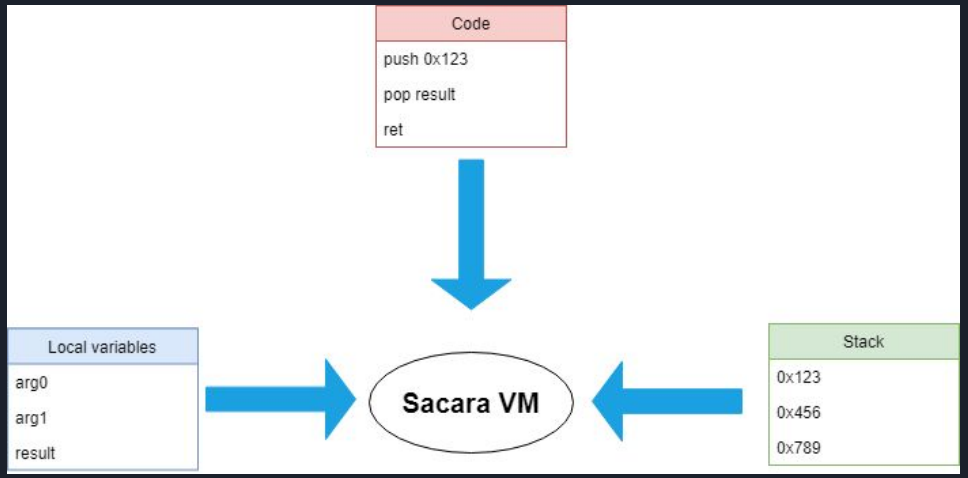
*

HACK IN BO® Winter 2019 Edition 13ª EDIZIONE

* This problem is well know to IDA devs: https://www.hex-rays.com/products/decompiler/manual/failures.shtml#06

VM based obfuscation

- One of the most difficult task in Reverse Engineering is to understand how the underline computer architecture works (instruction set, calling convention, memory layout, compiler characteristics, used Libs, …)
- We are very used to INTEL arch on Windows OS, but what about a new unknown architecture? This is the basic concept of VM base protection
- A personal experiment, Sacara: https://github.com/enkomio/sacara



VM based obfuscation

Example: decrypt a buffer

Src: https://github.com/enkomio/sacara/blob/master/Src/Examples/LoadEncryptedAssembly/Encryption.cs

```
proc main
  push buffer
  push buffer_length
  push key
  push key_length
  push 4
  push de_encrypt
  call
  halt
endp
```

```
/*
This method accept:
1 - the length of the password
2 - a pointer to the password to use
3 - the lengh of the buffer
4 - a pointer to the buffer
*/
proc de_encrypt
  pop key_length
  pop key
  pop buffer_length
  pop buffer
  push 0
  pop buffer_index
  push 0
  pop key_index
  push 0
  pop buffer_char
  push 0
  pop key_char
```

```
encryption_loop:
  /* read the character from the buffer */
  push buffer_index
  push buffer
  add
  nread
  pop buffer_char
  /* read the character from the key buffer */
  push key_index
  push key
  add
  nread
  pop key_char
  /* do XOR and save the result on the stack */
  push key_char
  push buffer_char
  xor
  /* write back the result */
  push buffer_index
  push buffer
  add
  nwrite
```

```
/* increase counter */
push 1
push key_index
add
pop key_index
push 1
push buffer_index
add
pop buffer_index
/* check if I have
push key
_index
mp
push check_for_comp
jumpifl
round_key:
  push 0
  pop key_index
```

```
check_for_completation:
  push buffer_length
  push buffer_index
```

```
/* do XOR and save the result on the stack */
push key_char
push buffer_char
xor
```

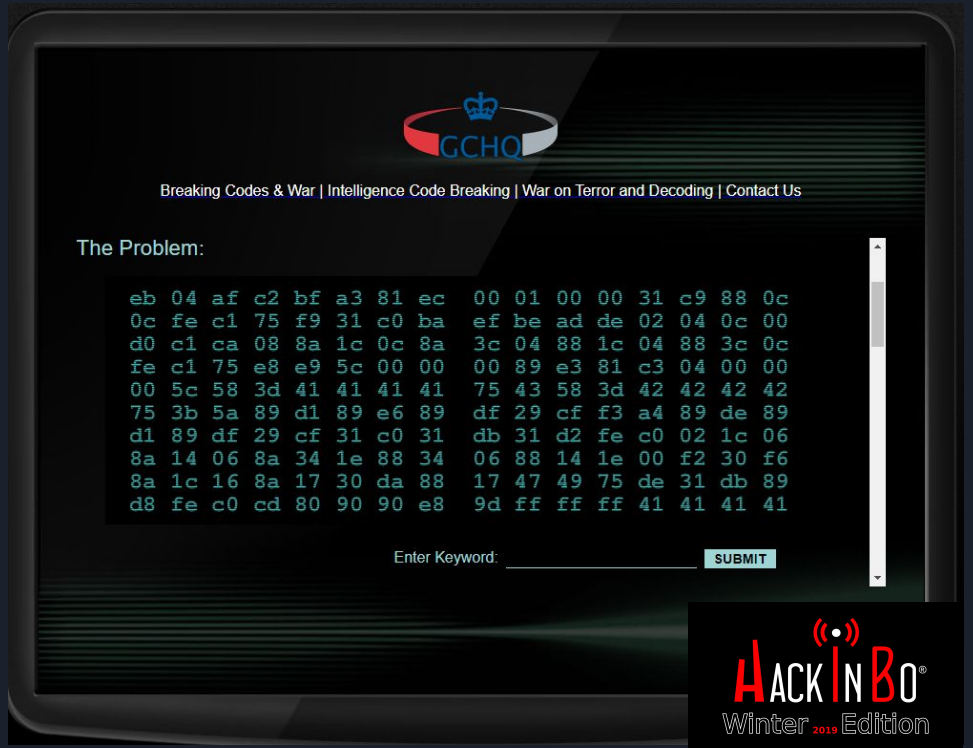
VS

```
xor eax, ebx
```

\+ Encrypted Opcode
\+ Anti-tampering
\+ ...

Reverse Engineering != Reading Assembly

- Doing Reverse Engineering doesn't always imply to read Assembly
- Sometimes it is easier to just try to get rid of the data by looking for patterns
- Some interesting links:
  - https://www.canyoucrackit.co.uk/codeexplained.html
  - http://blog.pi3.com.pl/?p=213
- If you want a more fresh challenge and you like more NSA, here is another one:
  - https://codebreaker.ltsnet.net/challenge



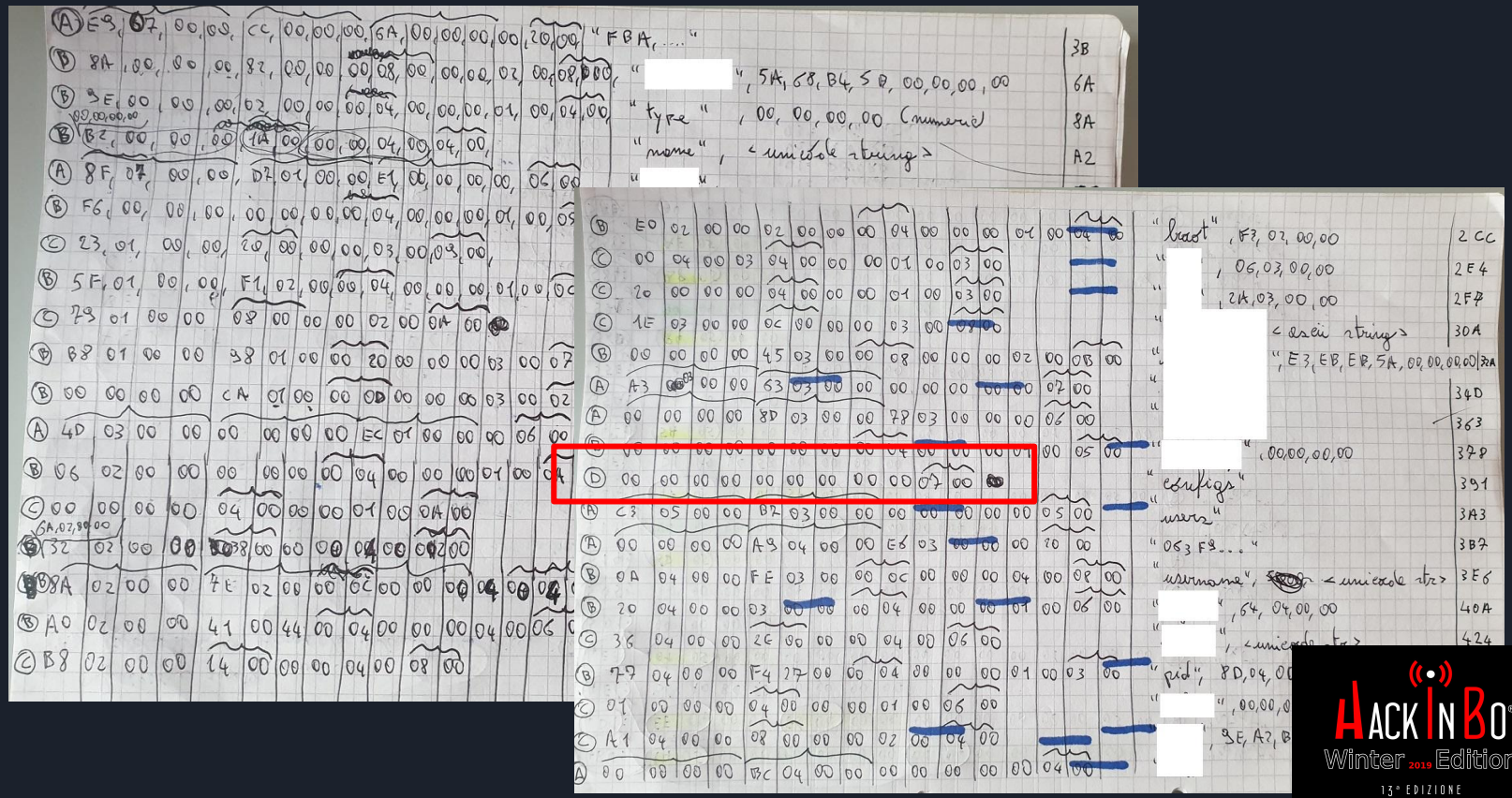
Reverse Engineering != Reading Assembly

- A real world case
  - File containing information about compromised computers
  - Malware written in C++, the code that read and update the file wasn't easy to understand and difficult to trigger
  - File seems to be in plain text (no encryption)

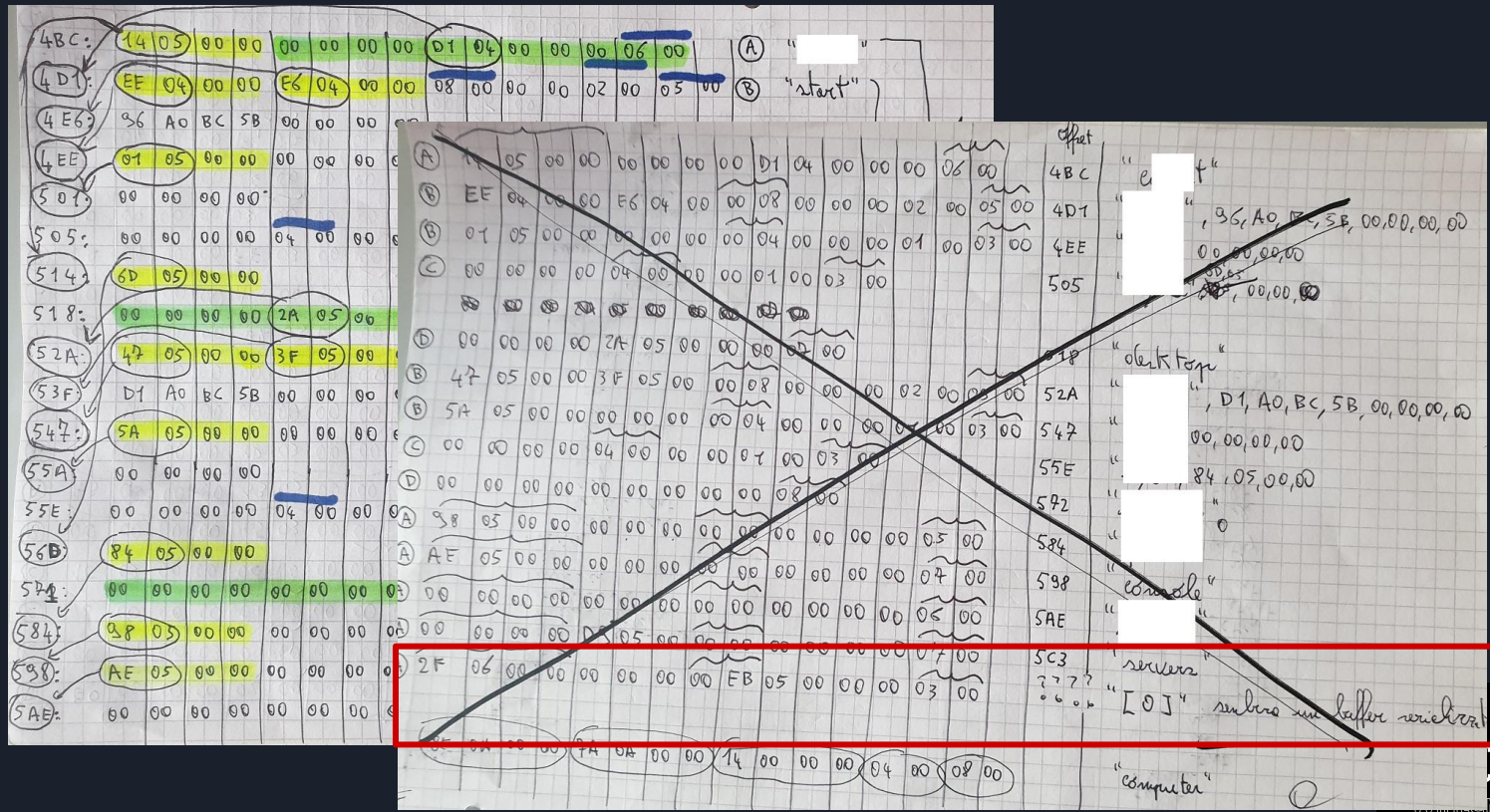
- Initial bytes

```
0x73 0x65 0x63 0x72 0x03 0x00 0x00 0x00 0x18 0x00 0x00 0x00 0x00 0x00 0x00 0x00
0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x27 0x00 0x00 0x00
0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x3B 0x00 0x00 0x00 0x00
0x00 0x00 0x00 0x00 0x05 0x00 0x73 0x34 0x74 0x61 0x6E 0x02 0x01 0x00 0x00 0xB3
0x00 0x00 0x00 0x6A 0x00 0x00 0x00 0x00 0x20 0x00 0x68 0x74 0x74 0x70 0x73 0x3A
0x2F 0x2F 0x77 0x77 0x77 0x2E 0x74 0x61 0x69 0x70 0x61 0x6E 0x73 0x65 0x63 0x2E
0x63 0x6F 0x6D 0x2F 0x69 0x6E 0x64 0x65 0x78 0x20 0x8A 0x00 0x00 0x00 0x82 0x00
0x00 0x00 0x08 0x00 0x00 0x00 0x02 0x00 0x08 0x00 0x70 0x72 0x6F 0x65 0x64 0x69
0x74 0x69 0xF5 0x7B 0xE9 0x5B 0x00 0x00 0x00 0x00 0x9E 0x00 0x00 0x00 0x02 0x00
0x00 0x00 0x04 0x00 0x00 0x00 0x01 0x00 0x04 0x00 0x63 0x6F 0x6F 0x6C 0xC6 0x00
0x00 0x00 0x00 0x00 0x00 0x00 0x04 0x00 0x00 0x00 0x01 0x00 0x05 0x00 0x69 0x74
0x61 0x6C 0x79 0x2A 0x07 0x00 0x00 0x00 0x00 0x00 0x00 0xCB 0x04 0x00 0x00 0x00
0x04 0x00 0x69 0x6E 0x66 0x6F 0xE6 0x00 0x00 0x00 0xD8 0x00 0x00 0x00 0x0E 0x00
0x00 0x00 0x03 0x00 0x02 0x00 0x48 0x61 0x63 0x6B 0x49 0x
0x39 0x31 0x31 0x30 0x31 0x39 0xEE 0x07 0x00 0x00 0xFA 0x
```

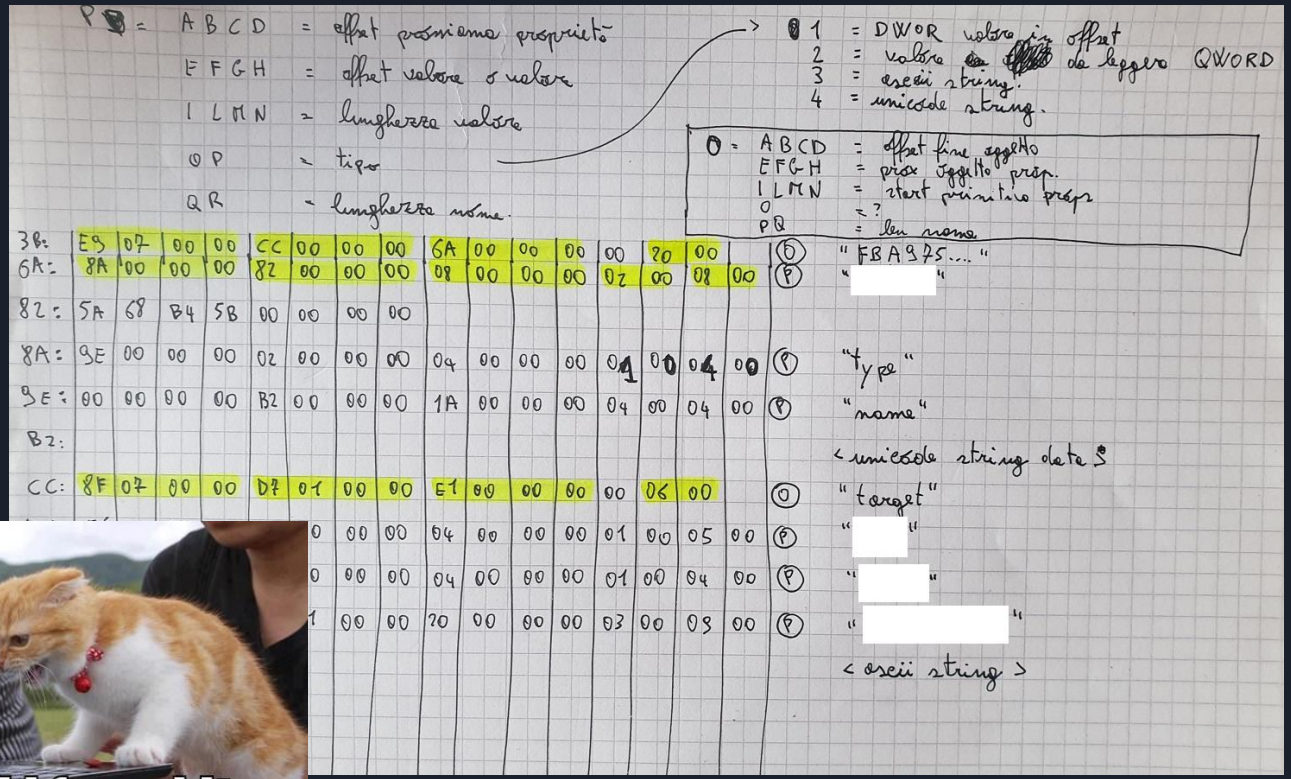
Reverse Engineering != Reading Assembly



Reverse Engineering != Reading Assembly



Reverse Engineering != Reading Assembly



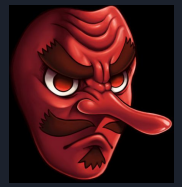
HACK IN BO
Winter 2019 Edition
13ª EDIZIONE

Sojobo a B2R2 emulator

- Sojobo emulates the B2R2 IR in order to provide an environment where you can emulate the execution of a binary. You can download it from:

  ★ Star https://github.com/enkomio/Sojobo

- At the current state it supports:
  - Intel architecture X86 32 bit
  - Window Process
  - A limited API set
- Tengu is a command line debugger like tool based on Sojobo
  - Same command switches as **windbg**
  - It allows to save snapshot
  - It emulates main Windows functions



Sojobo a B2R2 emulator

```
// emulate a malware and take snapshot at a given address
let sandbox = new Win32Sandbox()
let snapshotManager = new SnapshotManager(sandbox)
sandbox.Load(malwareFile)

// setup handlers
sandbox.BeforeEmulation.Add(fun proc ->
        if 0x401061 = proc.ProgramCounter.As<Int32>() then
                snapshotManager.TakeSnaphot()
)

// run the sample
sandbox.Run()
```

Case Study: KPOT v2

- KPOT v2 is an information stealer malware sold on underground forums
- A description about the malware is provided by the author

```
KPOT v2.0 update:
Soft:
1.1) Added the ability to grabbing files across the entire disk and over the network.
1.2) The storage structure in the grabber was revised. Now all the files are divided into folders as they were in the directory fro
m which the collection was.
2) Added to the RDP collection from the user folder for all users from which it is possible to collect.
3) Reworked collection from Windows storage (Credentials and Protected Storage). Now collects all the data pack without filtering
on any particular, i.e. if the software meets data of an unknown type without encryption, it will collect it in its pure form, if t
hey
will be encrypted, it will collect, but will not benefit from them.
4) Added collection of programs in the system information. Gathers the name and version of the installed program.
Both x64 and x86 programs are compiled.
5) Added Outlook collection from the registry for all users from which it is possible to collect.
6) Improved resolv .bit domains. All the workpieces I found at the time of adding dns for a resolver, as well as the dotbit proxy,
were added.
...
Current price: $ 85
Installation of the admin: $ 25 (the guide has been redone, now the installation is described much more clearly).
```

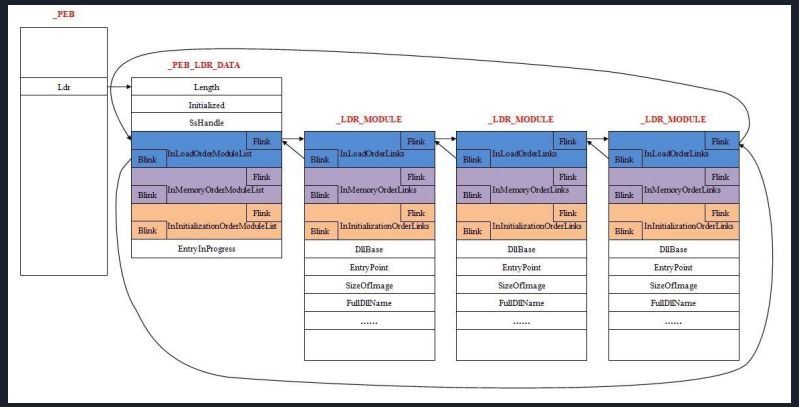
*

* Source: https://www.proofpoint.com/us/threat-insight/post/new-kpot-v20-stealer-brings-zero-persistence-and-memory-features-silently-steal

KPOT function resolution algorithm

Steps to resolve a function pointer:

1. Walk **TEB->PEB->Ldr** to get the base address for Kernel and ntdll. Resolve *LoadLibraryA* by walking Kernel32 EAT. Use *LoadLibraryA* to load the desired DLLs
2. Store the DLL base address and other info in a structure composed by the following items: *<base address, number of functions to lookup, function array>*
3. Parse PE and walk EAT. For each exported function compute the *MurmurHash* hash and search for this value in the above array. If found store the pointer.

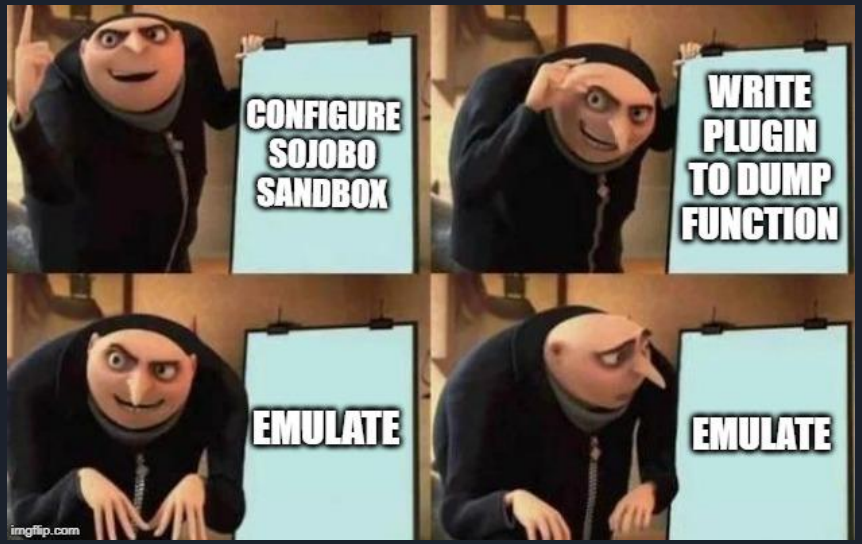



```
0018FB54  3C 00 00 00 00 00 1B 75 B0 FC 18 00 0B 00 00 00  <.....u°ü.....
0018FB64  00 00 46 6C 08 FD 18 00 0A 00 00 00 00 00 2B 75  ..Fl.ý........+u
0018FB74  58 FD 18 00 0A 00 00 00 00 00 EA 76 38 FE 18 00  Xý........êv8þ..
0018FB84  06 00 00 00 00 00 5F 77 B8 FE 18 00 04 00 00 00  ......_w¸þ......
```

- DLL Base address
- Array of hash to search for
- Number of hash in array

Goal: We want to know which are the functions that are resolved by the malware

- Sample SHA-256 : 67f8302a2fd28d15f62d6d20d748bfe350334e5353cbdef112bd1f8231b5599d
- By knowing which are the used functions we can have a better picture of the malware functionalities. Let's emulate the previous steps in Sojobo.



Goal: We want to know which are the functions that are resolved by the malware

At Step 1 we have the biggest problem. We need to have a valid PEB structure to correctly emulate the execution. The *Ldr* field is one of the most difficult to represent since it contains a linked list via LIST_ENTRY structure.

At lower level it is easy to manage LIST_ENTRY, but how to represent it at a high level language like F#? Possible solution:

- LIST_ENTRY can point to any kind of data, it is a nice use case for using inheritance!
  - We can't do this if we consider LIST_ENTRY like a struct. Struct cannot be inherited by definition.
- Then consider LIST_ENTRY as a class
  - We can't do this, since it is treated like a structure (it occupy 8 bytes in x86, since it has 2 pointers). If we define it like a class we will have a pointer during serialization (4 bytes and not 8).
- Treat it as a struct and consider the pointed object like a generic Object class
  - Goodbye deserialization → Impossible to know during deserialization which Object type we have to create
- ...

Goal: We want to know which are the functions that are resolved by the malware

- Writing Binary Analysis tools it's not an easy task :)

Conclusion

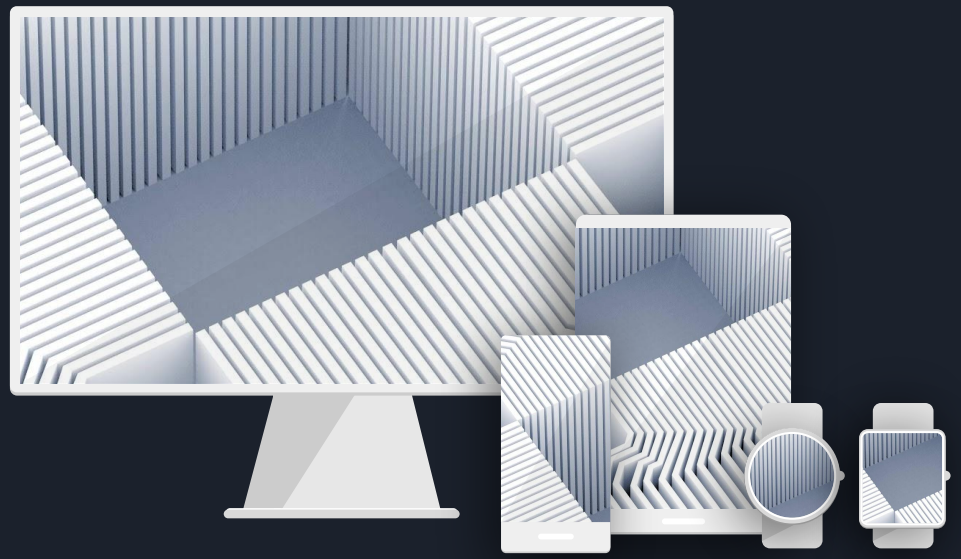
- Effective malware can be very complex
- Effective anti-analysis techniques can slower the reverse engineering process
  - Anti-VM
  - Anti-Debugging
  - VM based protection
- Some implementation choices can further slow the analysis
  - Usage of rarely used compression algorithms
  - Usage of external lib for crypto instead of relying on Windows Crypto API
- There are many tools that can help to analyze malware, not only debuggers and disassemblers :)
  - In order to be proficient with them is necessary to have some basic/medium knowledge about reverse engineering

Thank you!

Twitter: s4tan

GitHub: https://github.com/sponsors/enkomio

Contact: aparata@gmail.com



HACK IN BO
Winter 2019 Edition
13ª EDIZIONE