Threat hunting NET malware with YARA

Santiago Martin Pontiroli

Global Research and Analysis Team

Kaspersky Lab



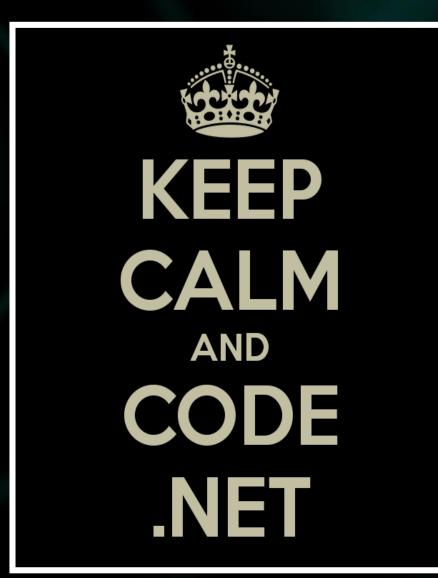
@spontiroli

GREAT

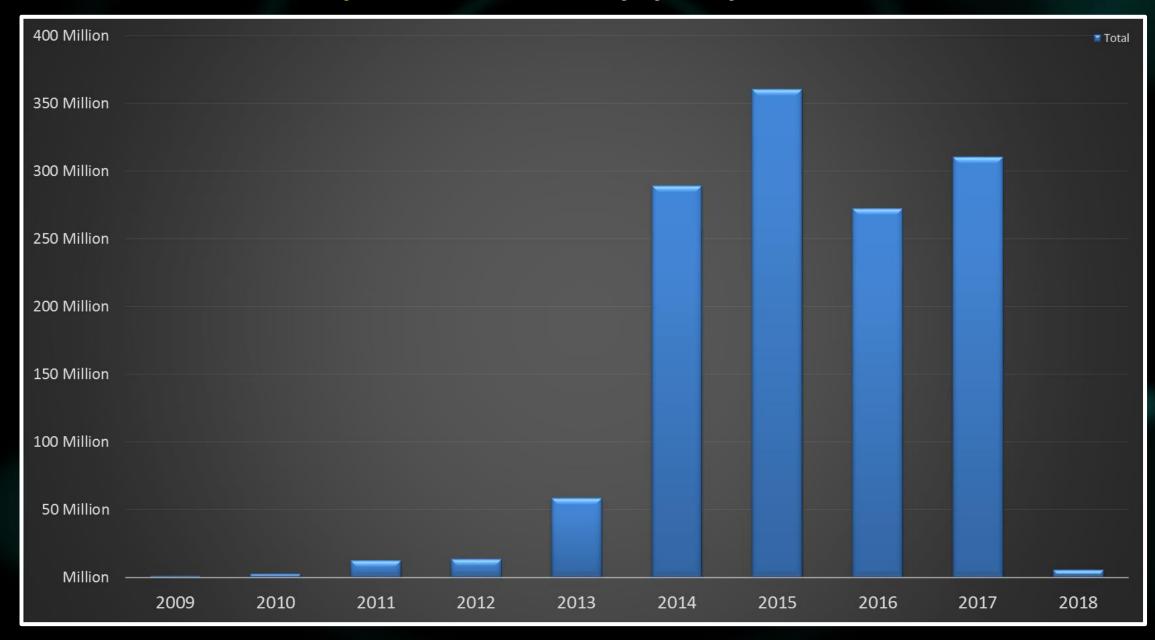
#BSIDESNYC KASPERSKY®

What's the problem with .NET malware?

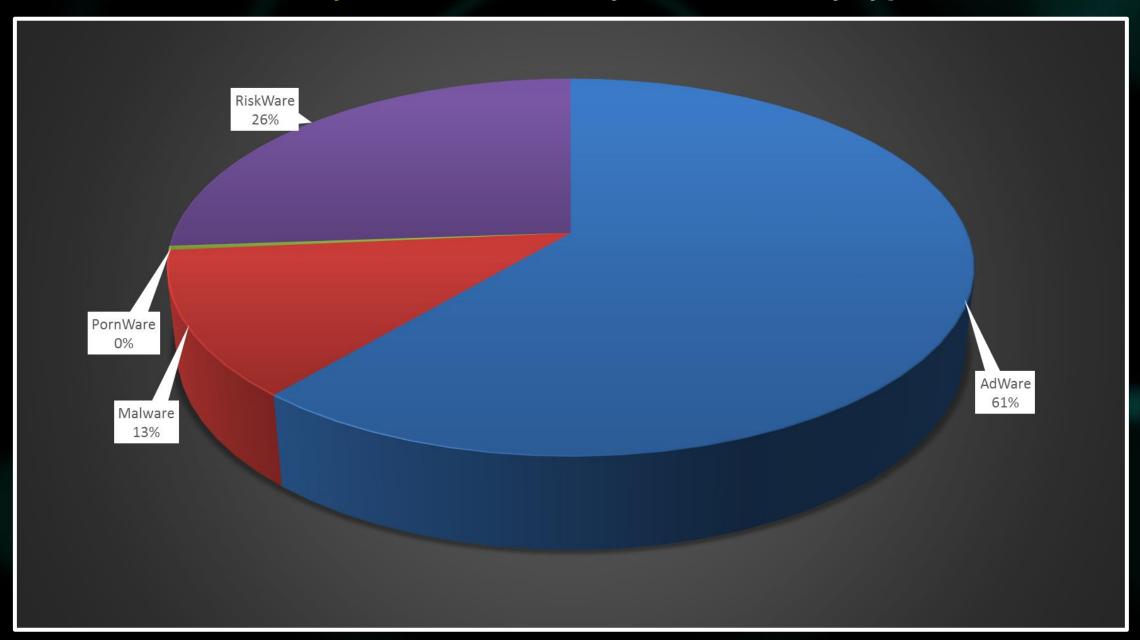
- Available by default in most Windows installations, it has become the de-facto standard for software development in Microsoft's family of operating systems.
- Vast amounts of ready-to-use functionality make .NET and PowerShell a deadly combination at the hands of cybercriminals.
- Since 2009 there has been a steady growth in the number of .NET malware, but it's still treated as other regular PEs by analysts.



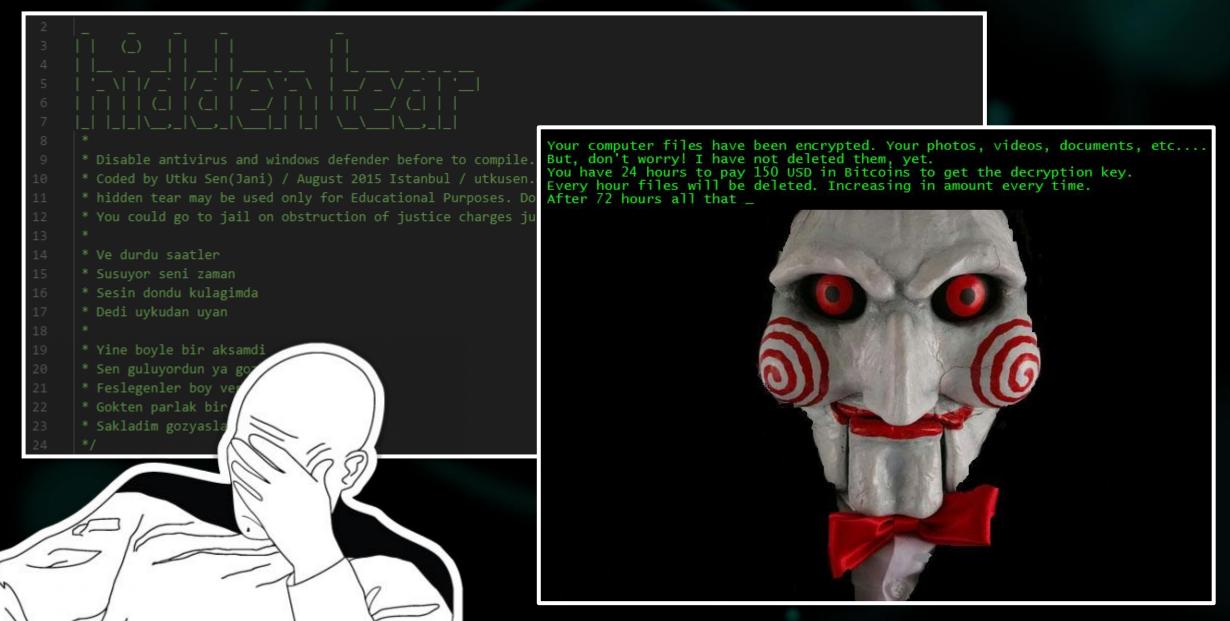
The .NET malware ecosystem as it is today, yearly detections



The .NET malware ecosystem as it is today, detections by type



From the most simple to targeted attacks, .NET malware is rising



YARA, the pattern matching swiss knife

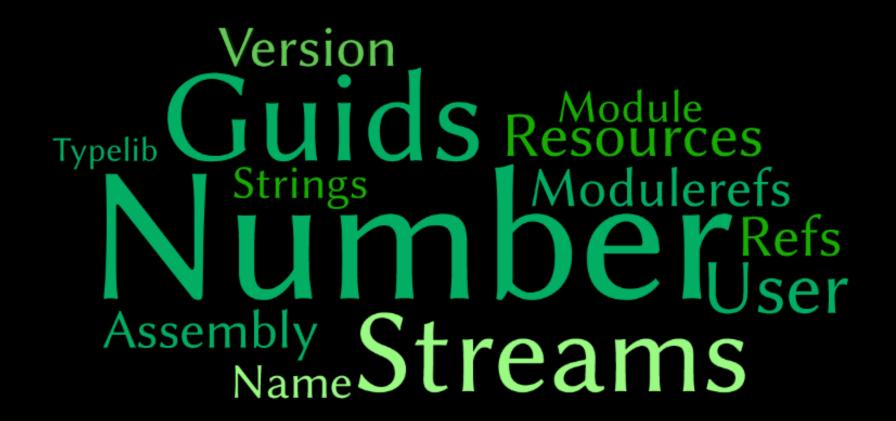
YARA helps malware researchers identify and classify malware samples by using descriptive patterns.

```
rule Derkziel
   meta:
        description = "Derkziel info stealer (Steam, Opera, Yandex, ...)"
        author = "The Malware Hunter"
        filetype = "pe"
       date = "2015-11"
       md5 = "f5956953b7a4acab2e6fa478c0015972"
        site = "https://zoo.mlw.re/samples/f5956953b7a4acab2e6fa478c0015972
       reference = "https://bhf.su/threads/137898/"
   strings:
       $drz = "{!}DRZ{!}"
       $ua = "User-Agent: Uploador"
       $steam = "SteamAppData.vdf"
       $login = "loginusers.vdf"
        $config = "config.vdf"
   condition:
        all of them
```

- You need a set of strings or a logical condition to build your rule.
- YARA contains several modules for extending its features and expressing more complex conditions.

YARA and .NET, getting better by the day

The dotnet module allows you to create more fine-grained rules for .NET files by using attributes and features of the .NET file format.



Why think about .NET malware differently?

In general, MSIL binaries are easier to reverse engineer than compiled code, however obfuscation makes some executables useless by reducing the amount of indicators we can gather from them statically.

Using additional information embedded into the binary then becomes useful for finding related samples in Virus Total and similar engines.

```
+ using ...
  [assembly: AssemblyVersion("1.0.0.0")]
  [assembly: Debuggable(DebuggableAttribute.DebuggingModes.IgnoreSymbolStoreSeque
  [assembly: AssemblyCompany("")]
  [assembly: AssemblyConfiguration("")]
  [assembly: AssemblyCopyright("Copyright @ 2014")]
  [assembly: AssemblyDescription("Your worst nightmare.")]
  [assembly: AssemblyFileVersion("1.0.0.0")]
  [assembly: AssemblyProduct("Locker")]
  [assembly: AssemblyTitle("CoinVault")]
  [assembly: AssemblyTrademark("")]
  [assembly: CompilationRelaxations(8)]
  [assembly: RuntimeCompatibility(WrapNonExceptionThrows = true)]
  [assembly: ComVisible(false)]
  [assembly: Guid("c91c210e-0d7f-4c15-b01d-7b51d00e7d77")]
  [module: ConfusedBy("Confuser v1.9.0.0")]
  [module: SuppressNdasm]
```

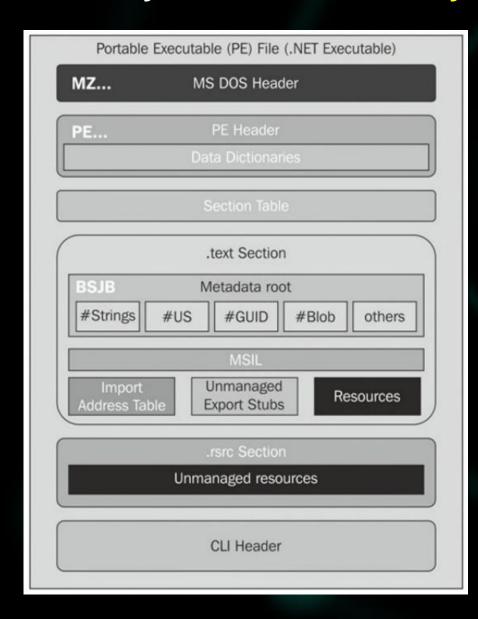
```
2 // foryou, Version=1.0.1.711, Culture=neutral, PublicKeyToken=null
 4 // Entry point: ZRmIcR.Program.Main
 6 using System;
 7 using System.Reflection;
 8 using System.Runtime.CompilerServices;
 9 using System.Security.Permissions;
  [assembly: AssemblyVersion("1.0.1.711")]
12 [assembly: AssemblyCompany("Malwarebytes Corporation"
  [assembly: AssemblyCopyright("@ Malwarebyt& Coeerporation. All rights reserved.")]
  [assembly: AssemblyDescription("Malwarebytes Anti-Malware")]
  [assembly: AssemblyFileVersion("1.0.1.711")]
   [assembly: AssemblyProduct("Malwargsegfasebytes Anti-Malware")]
   [assembly: AssemblyTitle(\msdfsdbam.exe")]
18 [assembly: AssemblyTrademark("")]
  [assembly: CompilationRelaxations(8)]
20 [assembly: RuntimeCompatibility(WrapNonExceptionThrows = true)]
  [assembly: SecurityPermission(SecurityAction.RequestMinimum, SkipVerification = true)]
```

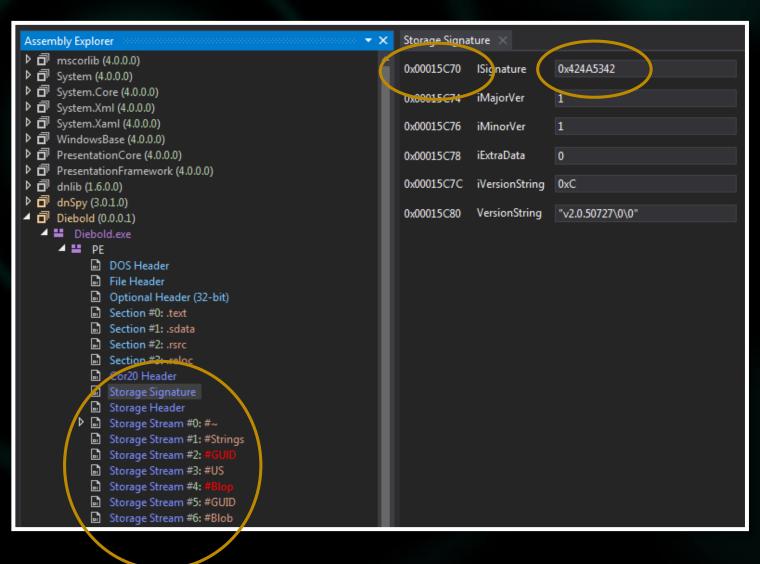
Anatomy of a .NET Assembly, from MZ to BSJB

4D	5A	90	00	03	00	00	00	04	00	00	00	FF	FF	00	00	MZ ♥ ♦ ÿÿ
В8	00	00	80	00	00	00	00	40	00	00	00	00	00	00	00	. @
90	00	00	00	00	00	00	00	99	00	00	00	00	00	00	00	
99	00	00	99	00	00	00	99	99	00	99	00	80	00	00	00	€
ØE	1 F	ВА	ØE	00	В4	09	CD	21	В8	01	4 C	CD	21	54	68	ມື♥ºມີ ´oÍ!,@LÍ!Th
69	73	20	70	72	6F	67	72	61	6D	20	63	61	6E	6E	6F	is program canno
74	20	62	65	20	72	75	6E	20	69	6E	20	44	4F	53	20	t be run in DOS
6D	6F	64	65	2E	0D	0D	0A	24	00	00	00	00	00	00	00	mode.♪♪ ≡ \$
50	45	00	00	4 C	01	04	99	5C	57	20	58	00	00	00	99	PE L⊕♦ \W,X
00	00	00	00	Ε0	00	ØE	01	0B	01	06	00	00	ØA	03	00	à 2000⊕ ≘ V
99	0A	00	00	00	00	00	99	BE	29	03	00	00	20	00	99	≥ ¾)♥
99	40	03	99	99	00	40	99	99	20	99	99	00	02	00	99	@♥ @ •
04	00	00	00	99	00	00	99	94	00	99	00	00	00	00	99	* * * * * * * * * * * * * * * * * * *
99	Α0	03	00	00	04	00	99	99	00	00	00	02	00	40	85	∀ ♦ ⊕
99			99					99		10						Problem Broken
99			99					99		99						•
70	29	03	99	4B	99	99	99	99	60	03	99	F8	02	99	99	p)♥ K `♥ ø@
99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	99	
99	80	03	99	9C	99	99	99	2E	29	03	99	1 C	99	99	99	€♥ ♀ .)♥ ∟
99			99			99		99		99						
99			99			99	99	99		99				99		
99			99			99		99		99				99		
99			99					98		99				99		■ н
99			00					2E		65						.text
C4			00					99		03						Äo♥ €♥ ♦
99			00					99		99						
2E			61					E8		99						.sdata è⊕ @♥
99		99			ØE		99	99		99		99		99		+ <i>₽</i> ¥
99			00			99		2E		73	72		99	99		@ À.rsrc
F8			00				00	99		00				03		Ø 9 `♥ ♦ \$♥
99		00		00		00		00		00			00	00		@ @
2E			6C					9C	00	00	99		80	03		.reloc ♀ €♥
99		00			16			99	00	00	00	00	00	00		9
99	99	99	00	40	99	99	42	99	99	00	99	99	99	99	99	@ B

```
4A 42 01 00 01 00
                          00 00 00 00 00 00 00 06
                                                   BSJB@ @
76 32 2E 39 2E 35 30 37
                          32 37 00 00 00 00 07 00
8C 00 00 00 7C 49 00 00
50 53 00 00 23 53 74 72
58 9D 00 00 01 00 00 00
                                                           #US å
59 9D 00 00 8C 03 00 00
                          23 55 53 00 E5 A0 00
01 00 00 00 23 42 6C 6F
                                                        #Blop
                          70 00 00 00 E6 A0 00 00
                                                        #GUID
20 00 00 00 23 47 55 49
                          44 00 00 00 06 A1 00 00
                                                        #Blob
30 10 00 00 23 42 6C 6F
                          62 00 00 00 00 00 00
02 00 00 01 57 F7 A2 3D
                                                       @W÷⊄=0•
00 00 00 00 02 00 00 00
                          B8 00 00 00 2C 00 00 00
79 00 00 00 14 03 00 00
                          50 01 00 00 01 00 00 00
4F 01 00 00 31 00 00 00
07 00 00 00 43 00 00 00
                          01 00 00 00 02 00 00
03 00 00 00 06 00 00 00
                          03 00 00 00 1B 00 00
09 00 00 00 02 00 00 00
                          07 00 00 00 04 00 00
14 00 00 00 01 00 00 00
                          00 00 69 02 01 00 00 00
FF FF 00 00 00 00 02 00
29 00 06 00 58 00 5D 00
6A 00 89 00 06 00 9B 00
06 00 BB 00 CF 00 1F 00
5D 00 06 00 F9 00 29 00
                          06 00 17 01 89 00 06 00
2E 01 89 00 06 00 4B 01
06 00 83 01 89 00 06 00
CB 01 06 00 EA 01 CB 01
10 02 89 00 06 00 2B 02
                          89 00 5F 00 A4 02 00 00
5B 00 B7 02 00 00 06 00
4F 03 06 00 78 03 5D 00
                          0E 00 B0 03 B5 03 12 00
                          56 05 06 00 91 06 5D 00
62 04 6C 04 06 00 34 05
                          19 07 5D 00 16 00 58 00
06 00 FC 06 5D 00 06 00
5D 00 06 00 DA 08 5D 00
                          06 00 E0 08 5D 00 06 00
E5 08 F0 08 06 00 FA 08
                          5D 00 06 00 04 09 F0 08
```

Anatomy of a .NET Assembly, plenty of metadata to use





What are we looking at?

.NET holds metadata information in a number of streams, each in a different format.

- #Strings: an array of ascii strings, these are referenced by MetaData Tables
- #US: an array of unicode strings, these are referenced directly by code instructions
- #Blob: contains data referenced by MetaData Tables
- #GUID: contains 128 bit long unique identifiers
- #~: contains the MetaData Tables



Let's use YARA!

How to write YARA rules for .NET binaries and don't lose your mind...

```
rule create_yara_rule_dont_go_insane {
     meta:
         description = "You will need some tools if you want to save time"
         conference = "BSides NYC"
         author = "Santiago Pontiroli"
         date = "2018-01-20"
         version = "1.0"
10
     strings:
         $str1 = "YARA" ascii
         $str2 = "CFF Explorer" ascii
         $str3 = "dnSpy" ascii
         $str4 = "FAR Manager" ascii
     condition:
         all of them
```

Having PE Studio and Visual Studio Code won't hurt either!

Using **GUIDs** the right way

```
        0x00015D78
        Generation
        0
        UInt16

        0x00015D7A
        Name
        0x269
        Diebold.exe (#Strings Heap Offset)

        0x00015D7C
        Mvid
        1
        a6a39fb3-17d1-421b-823c-0de3765fac81 (#GUID Heap Index)

        0x00015D7E
        EncId
        0
        #GUID Heap Index

        0x00015D80
        EncBaseId
        0xFFFF
        #GUID Heap Index
```

```
using System;
using System.Diagnostics;
using System.Reflection;
using System.Runtime.CompilerServices;
using System.Runtime.InteropServices;
[assembly: AssemblyVersion("0 0 0 1")]
[assembly: ComVisible(false)]
[assembly: Guid("dc804d65-c6cd-45ef-a299-bcf8b69a11ea")]
[assembly: AssemblyCopyright("Copyright @ 2015")]
[assembly: AssemblyProduct("Diebold")]
[assembly: AssemblyKeyName("")]
[assembly: CompilationRelaxations(8)]
[assembly: SuppressIldasm]
[assembly: AssemblyDelaySign(false)]
[assembly: Debuggable(DebuggableAttribute.DebuggingModes.IgnoreSymbolStoreSequencePoints)]
[assembly: RuntimeCompatibility(WrapNonExceptionThrows = true)]
[assembly: AssemblyConfiguration("")]
[assembly: AssemblyCompany("")]
[assembly: AssemblyTrademark("")]
[assembly: AssemblyFileVersion("0.0.0.1")]
[assembly: AssemblyTitle("Diebold")]
[assembly: AssemblyDescription("")]
```

A TypeLib is GUID but not vice versa.

- The TypeLib ID is a GUID generated by Visual Studio on the creation of a new project by default.
 - The Module Version ID, or MVID, is a GUID that can be used to distinguish various versions of a .NET module. This value is generated at build time, resulting in a new GUID for each unique build.

Inspecting the assembly and its references

RID	Token	Offset	HashAlgId	MajorVersion	MinorVersion	BuildNumber	RevisionNumber	Flags	PublicKey	Name	Locale	into
1	0x20000001	0x0001A538	0x8004	0	0	0	1	0	0	1	0	Diebold
2	0x20000002	0x0001A54E	0x8004	1	0	0	0	0	0	0x244	0	032086e4-2252-4002-868b

0x00015C70	lSignature	0x424A5342
0x00015C74	iMajorVer	1
0x00015C76	iMinorVer	1
0x00015C78	iExtraData	0
0x00015C7C	iVersionString	0xC
0x00015C80	VersionString	"v2.0.50727\0\0"

	Token	Offset	MajorVersion	MinorVersion	BuildNumber	RevisionNumber	Flags	PublicKeyOrToke	Name	Locale	HashValue	Info
1	0x23000001	0x0001A564	2	0	0	0	0	0xA	0x49	0	0	mscorlib
2	0x23000002	0x0001A578	2	0	0	0	0	0xBB	0x34F	0	0	System.ServiceProcess
3	0x23000003	0x0001A58C	2	0	0	0	0	0xA	0x3B5	0	0	System.Windows.Forms
4	0x23000004	0x0001A5A0	2	0	0	0	0	0xBB	0x46C	0	0	System.Configuration.Insta
5	0x23000005	0x0001A5B4	65535	65535	65535	65535	0	0xA	0x49	0	0xFFFF	mscorlib
6	0x23000006	0x0001A5C8	2	0	0	0	0	0xA	0x5D	0	0	System
7	0x23000007	0x0001A5DC	2	0	0	0	0	0xBB	0x1A71	0	0	System.Drawing

Stop, syntax time!

```
1   import "dotnet"
2
3   rule mw_latam_plou
4   {
5   meta:
6    description =
7    filetype = "MS
8    author = "Sans
9    conference = '
10    date = "2018-6"
11    version = "1.6"
12
```

```
condition:
13
         uint16(0) == 0x5A4D and
15
         filesize < 320KB and
16
         //GUID and TypeLib
18
         dotnet.guids[0] == "a6a39fb3-17d1-421b-823c-0de3765fac81" and
19
         dotnet.typelib == "dc804d65-c6cd-45ef-a299-bcf8b69a11ea" and
20
21
         //Assembly (many more properties available)
22
         dotnet.assembly.name == "Diebold" and
23
         dotnet.version == "v2.0.50727" and
24
25
         //Assembly References (let's use one as an example)
         dotnet.assembly_refs[0].name == "mscorlib" and
26
         dotnet.assembly_refs[0].version.major == 2 and
28
         dotnet.assembly_refs[0].version.minor == 0 and
29
         dotnet.assembly_refs[0].version.build_number == 0 and
         dotnet.assembly_refs[0].version.revision_number == 0 and
         //Using YARA's regular expressions for the token
         dotnet.assembly_refs[0].public_key_or_token == "\xb7z\\V\x194\xe0\x89"
```

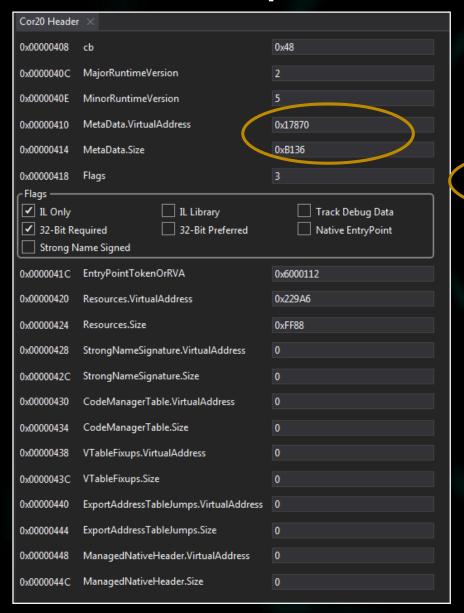
Getting more information from the god of wealth, modules and resources

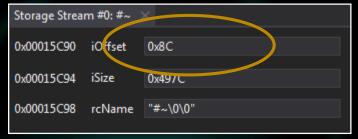
1A Mod	uleRef (6) ×				
RID	Token	Offset	Name	Info	
1	0x1A000001	0x0001A418	0x158A	user32.dll	
2	0x1A000002	0x0001A41A	0x15E9	kernel32.dll	
3	0x1A000003	0x0001A41C	0x235F	gdi32.dll	
4	0x1A000004	0x0001A41E	0x2C55	kernel32	
5	0x1A000005	0x0001A420	0x2D7B	advapi32.dll	
6	0x1A000006	0x0001A422	0x3133	shell32.dll	

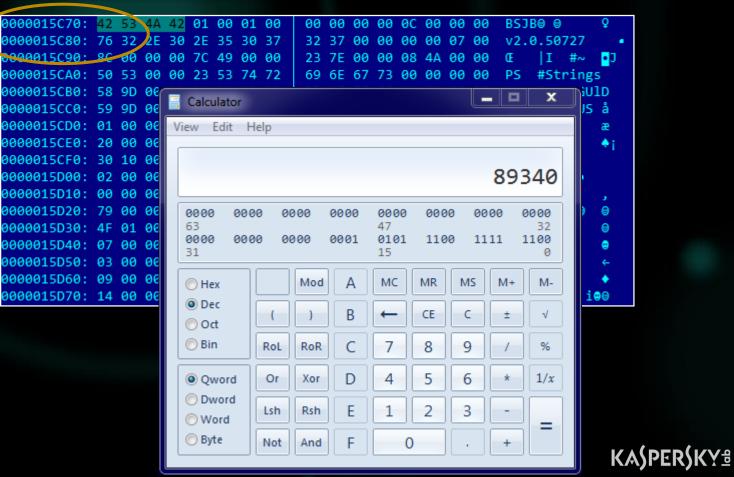
Stop, syntax time!

```
//Module References
         dotnet.number_of_modulerefs == 6 and
35
         dotnet.modulerefs[0] == "user32.dll" and
         dotnet.modulerefs[1] == "kernel32.dll" and
         dotnet.modulerefs[2] == "gdi32.dll" and
38
         dotnet.modulerefs[3] == "kernel32" and
39
         dotnet.modulerefs[4] == "advapi32.dll" and
41
         dotnet.modulerefs[5] == "shell32.dll" and
42
43
         //Resources
         dotnet.number_of_resources == 4 and
44
45
         //Offset for resource "vPPWagCMyNQp9yf6ae.HeX84Grn56V6JSxcQF" is 20DAA
46
         dotnet.resources[0].offset == 134570 and
         dotnet.resources[0].length == 55532 and
48
         dotnet.resources[0].name == "vPPWagCMyNQp9yf6ae.HeX84Grn56V6JSxcQF"
49
50
```

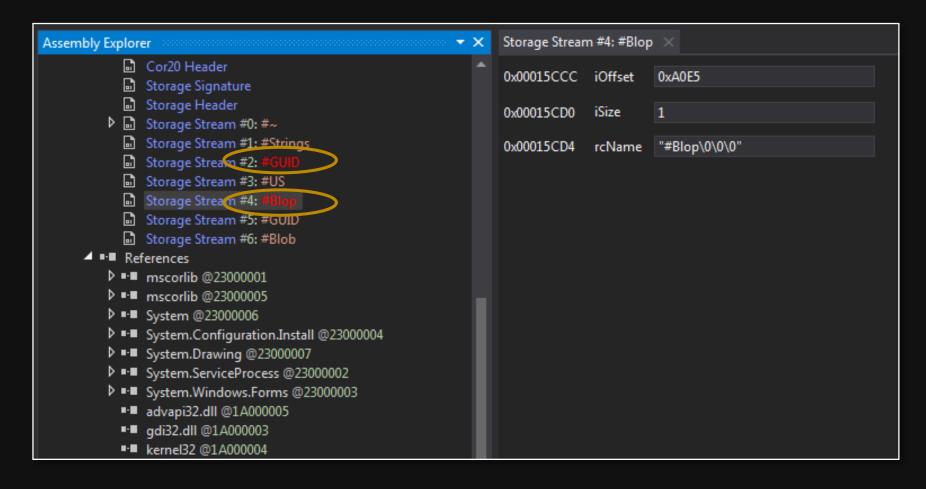
You must be shapeless, formless, like .NET streams







Not all streams are created equal



Finding interesting indicators in the binary

GUID != GUID

BLOB!=BLOP

Stop, syntax time!

```
50
         //Streams
         dotnet.number_of_streams == 7 and
52
         //GUIlD
53
         dotnet.streams[2].name == "#GU1D" and
         dotnet.streams[2].offset == 129480 and
         dotnet.streams[2].size == 1 and
56
58
         //BLOP
59
         dotnet.streams[4].name == "#Blop" and
60
         dotnet.streams[4].offset == 130389 and
61
         dotnet.streams[4].size == 1
62
```

Let's build our "final" YARA rule

```
condition:
              13
     import
                        uint16(0) == 0x5A4D and
                        filesize < 320KB and
     rule mw
                        //Example, using YARA's dotnet module
     meta:
                        dotnet.typelib == "dc804d65-c6cd-45ef-a299-bcf8b69a11ea" and
          des
                        dotnet.module name == "Diebold.exe" and
          fi]
                        dotnet.assembly.name == "Diebold" and
          aut
                        dotnet.version == "v2.0.50727" and
          cor
10
          dat
                        //Streams
          ver
                        dotnet.number of streams == 7 and
                        dotnet.streams[2].name == "#GU1D" and
                        dotnet.streams[2].offset == 129480 and
                        dotnet.streams[2].size == 1 and
                        dotnet.streams[4].name == "#Blop" and
                        dotnet.streams[4].offset == 130389 and
              29
              30
                        dotnet.streams[4].size == 1
              31
              32
```

It sounds like too much work, but YARA always likes to help

```
λ yara32.exe -r ploutus.yar .\BSidesNYC\ -D
dotnet
                                                                    name = "System.Drawing"
        number_of_constants = UNDEFINED
                                                   number of resources = 4
        constants
        typelib = "dc804d65-c6cd-45ef-a299-bcf
                                                   resources
                                                            [0]
        number of user strings = 17
                                                                    offset = 134570
        user strings
                [0] = "\x00"
                                                                    length = 55532
                [1] = "D\x00i\x00e\x00b\x00o\x0
                                                                    name = "vPPWagCMyNQp9yf6ae.HeX84Grn56V6JSxcQF"
00c\x00e\x00s\x00\x00"
                                                           [1]
                [2] = "S\x00y\x00s\x00t\x00e\x
                                                                    offset = 190106
00.\x000\x00,\x00 \x00C\x00u\x001\x00t\x00u\x0
                                                                    length = 4602
0T\x00o\x00k\x00e\x00n\x00=\x00b\x007\x007\x00
                                                                    name = "79QqHuks9wgCgBDn80.1vUDpFAoXtnxZW0f28"
                [3] = "S\x00y\x00s\x00t\x00e\x
                                                            [2]
00h\x00y\x00.\x00A\x00e\x00s\x00C\x00r\x00y\x0
                                                                    offset = 194712
                [4] = "R\x00K\x00a\x00w\x00w\x
                                                                    length = 5058
00M\x00f\x00Q\x001\x00U\x00u\x00t\x00a\x00W\x0
                                                                    name = "QNeRWrGiNRsROcgCJX.uDLYHATXunXUToMggM"
                [5] = "Q\x00N\x00e\x00R\x00W\x
                                                           [3]
00X\x00u\x00n\x00X\x00U\x00T\x00o\x00M\x00g\x0
                                                                    offset = 199774
                [6] = "{\x001\x001\x001\x}
                                                                    length = 208
0"
                                                                    name = "RKawwHdUCKccExTuH5.NhiP3K7MfQ1UutaWu4"
                [7] = "G\x00e\x00t\x00D\x00e\x
                                                   number of guids = 2
00t\x00e\x00r\x00\x00"
                                                   guids
                [8] = "_\x00_\x00\x00"
                [9] = \text{"m} \times 00 \times 00p \times 00t \times 00r \times 00p
                                                           [0] = "a6a39fb3-17d1-421b-823c-0de3765fac81"
```

The D flag will help you debug your rules, yara32.exe -r ploutus.yar \BSidesNYC\-D

Do you really need all that information?

- Are you trying to find new samples of a known threat or something similar and still unknown?
- What is the lowest amount of information you can use to identify the sample?
- If it only detects that sample, the rule is rarely useful (except incident response cases).
- If it detects too much and the number of false positives hinders your investigation it's time to go back to the drawing board.



Additional resources, becoming a YARA ninja

- How to write simple but sound YARA rules (Florian Roth)
 - https://www.bsk-consulting.de/2015/02/16/write-simplesound-yara-rules/
- Using .NET GUIDs to help hunt for malware (Brian Wallace)
 - https://www.virusbulletin.com/virusbulletin/2015/06/usingnet-guids-help-hunt-malware
- YARA dotnet module documentation
 - https://yara.readthedocs.io/en/v3.7.0/modules/dotnet.html
- YARA dotnet module source code
 - https://github.com/VirusTotal/yara/blob/master/libyara/modul es/dotnet.c

A special thank you goes to Wesley Shields (@wxs) and Victor Alvarez (@plusivc)





Thank you! Questions?

Santiago Martin Pontiroli

Global Research and Analysis Team

Kaspersky Lab



@spontiroli

KASPERSKY®

