

IcedID aka #Bokbot Analysis with Ghidra.

By Dawid Golak

Published: 2019-06-25 · Archived: 2026-04-10 02:06:12 UTC



5 min read

Jun 25, 2019

A few days ago [@brad](#) published a post on the twitter about a resume-themed password-protected Word doc that was dropping IcedID (also known as #BokBot). The sample is available for download on the any.run <https://app.any.run/tasks/13d6d9f9-7033-4ce7-9ad4-76591f15274c/> service for further analysis.

BTW. any.run is the awesome sandbox which can speed up the initial analysis of malicious files.

The IcedID sample was packed and contains an interesting startup mechanism

File analysis:

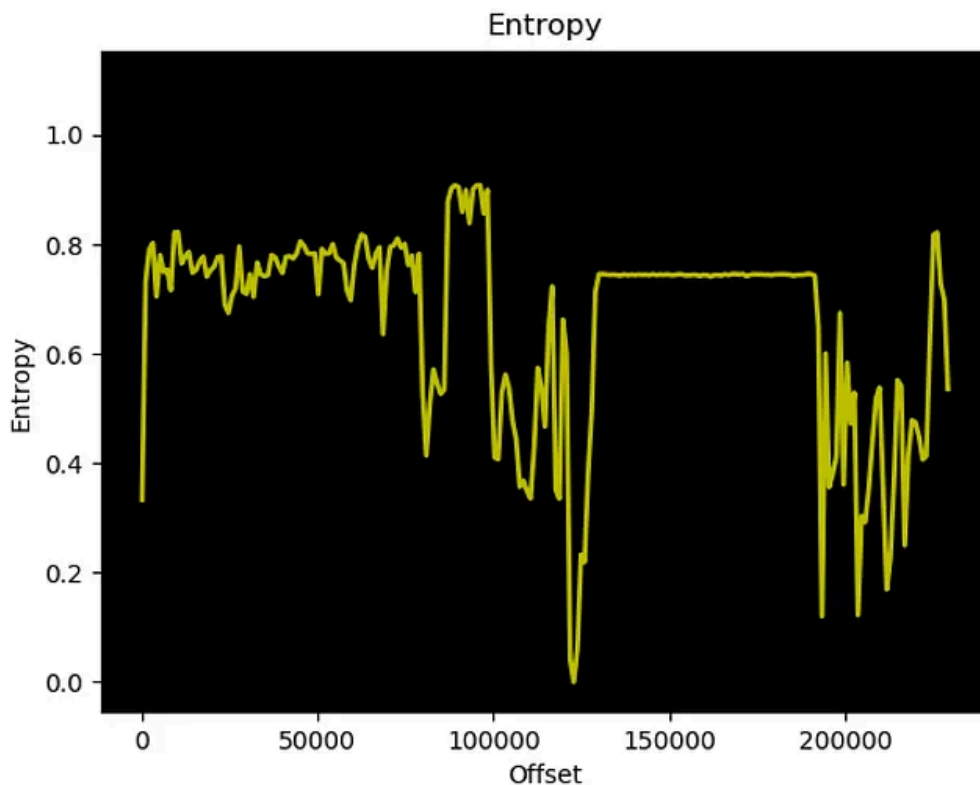
MD5:b05ea5fd73d25140cdb31f36789d9003

Filename:5.exe

sha256:d350d150f658a32c32984d7f879c6f3b3ddb6ba7918bfe22e19471a79a0cd490

At the first glance you can see, that the file is packed.

Press enter or click to view image in full size



After launching the sample in the debugger we can identify, that a standard packing mechanism is in use.

The first stage, or how to unpack IcedID?

If we want to unpack this file ourselves, then we can look at the OA Labs video tutorial, which is available [there](#) or follow the instruction below.

Briefly to unpack this sample:

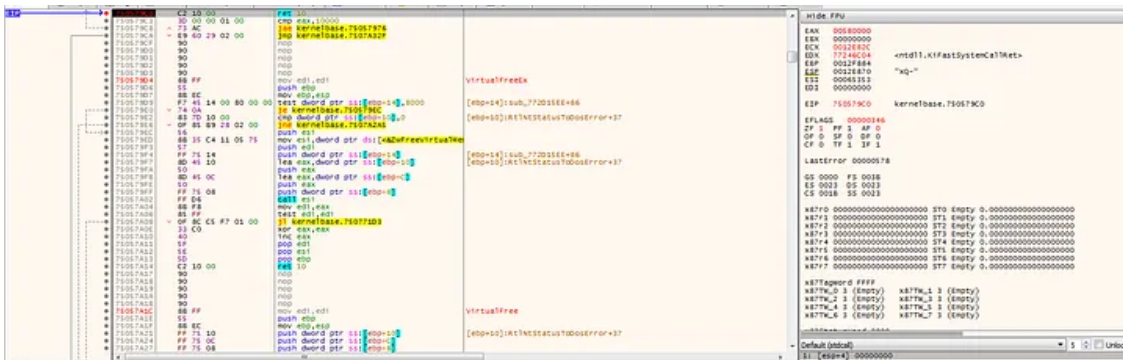
1. Set breakpoint on VirtualAlloc method

Press enter or click to view image in full size

Software name	Module name	State	Hit count	Log text	Condition	Full Command on hit	Comment
712AC6A	user32.dll,VirtualAlloc*	Enabled	1				

2. Launch sample and wait for the debugger to stop on this method, and then look at the allocated memory address (the address is returned by VirtualAlloc \$EAX).

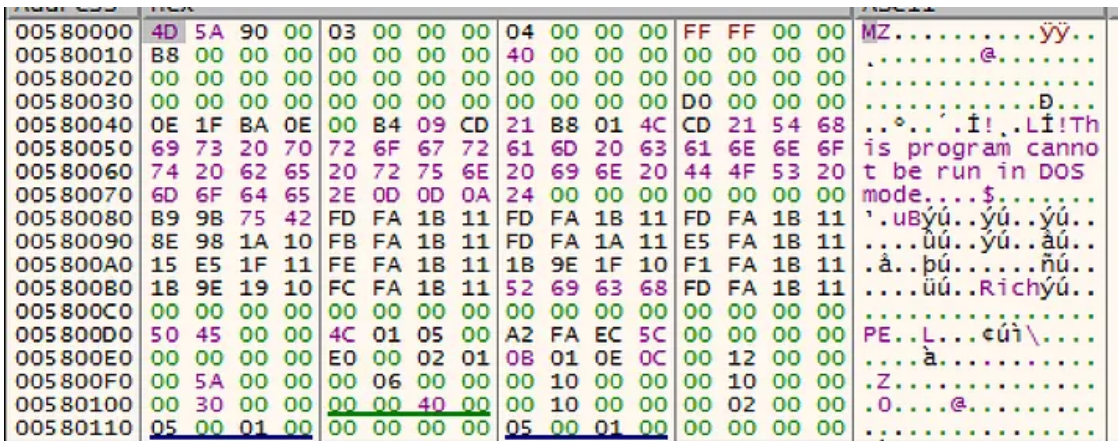
Press enter or click to view image in full size



Once we have this address, set a breakpoint on the initial bytes of the address and wait until the program write data to this address.

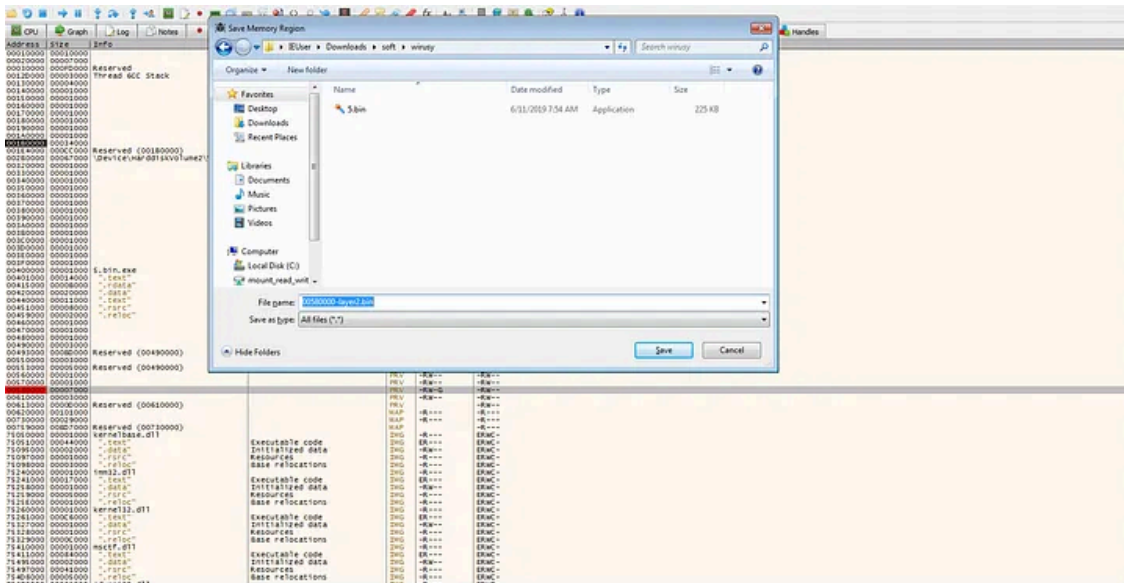
3. As we can see at this point, the initial bytes of the allocated memory area begin with 4D 5A (MZ). It means that there is the second stage of the sample.

Press enter or click to view image in full size



4. We save the memory to a file and now we have a copy of the second stage to analyze.

Press enter or click to view image in full size



The second stage

After unpacking the file and then previewing the generated pseudocode in #Ghidra, we can see the following flow.

Press enter or click to view image in full size

```

2 | undefined4 entry(int param_1,int param_2,char *param_3)
3 |
4 | {
5 |     char cVar1;
6 |     byte bVar2;
7 |     LPSTR pCVar3;
8 |     LPSTR value_of_arg1;
9 |     int iVar4;
10 |    char *magic_string;
11 |    uint uVar5;
12 |    uint uVar6;
13 |
14 |    pCVar3 = GetCommandLine();
15 |    value_of_arg1 = obtain_argv_q(pCVar3);
16 |    if (value_of_arg1 == (LPSTR)0x0) {
17 |        iVar4 = obtain_argv_param_p(pCVar3);
18 |        if ((iVar4 != 0) && (magic_string = FUN_004011be(), magic_string != (char *)0x0)) {
19 |            Sleep(1000);
20 |            create_self_process_with_additional_params(magic_string);
21 |        }
22 |    }
23 |    else {
24 |        create_svchost_process();
25 |    }
26 |    ExitProcess(0);

```

In line №15, the argument from parameter “-q=” is retrieved.

If this parameter is not present, the code beginning on line 17 is started. Further parameters are checked.

Why does he do it ?

The process under debugger, creates the a new process which is not debugged. It is one of the way to escape from the debugger.

Press enter or click to view image in full size



We have several options to debug the new process.

Ps. For instance I looked at what parameters it is run “CreateProcessA” and again execute the malware with the additional option under the debugger.

“C:\Users\admin\AppData\Local\Temp\5.exe” -q=[int]”

for example:

“C:\Users\admin\AppData\Local\Temp\5.exe” -q=412588568”

Get Dawid Golak’s stories in your inbox

Join Medium for free to get updates from this writer.

Remember me for faster sign in

Once these argument checks have passed the next interested function to analyze is FUN_004011be().

Press enter or click to view image in full size

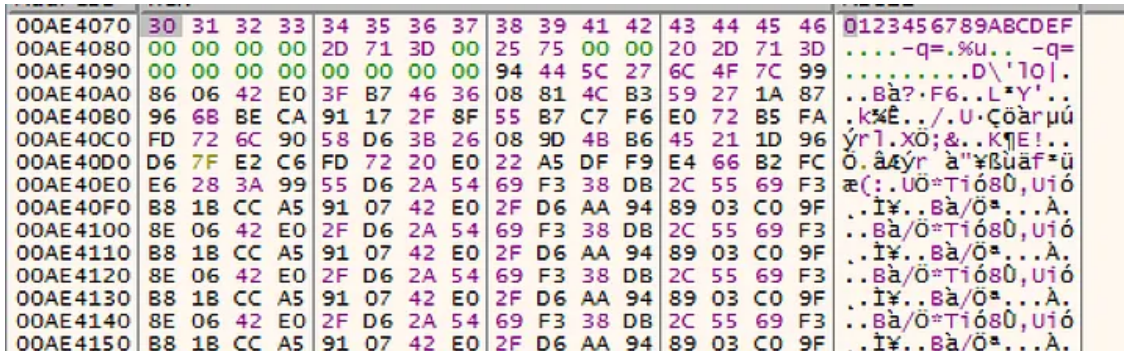

```

004011f1 83 e0 0f      AND     EAX,0xf
004011f4 46          INC     ESI
004011f5 8a 80 70      MOV     AL,byte ptr [s_0123456789ABCDEF_00404070
          40 40 00
004011fb 88 41 ff      MOV     byte ptr [ECX + -0x1],AL
    
```

AND EAX,0xf, the retrieved is lower bits of \$EAX register

The AL register value indicates the element number from the second data set **“0123456789ABCDEF”**

Press enter or click to view image in full size



This algorithm has been replicated in python below.

```

key=file('data3.txt').read()
tab=file('data2.txt').read()
ind = 0
res = []
p=""
for i in key:
    first_poz = ord(i)>>4
    second_poz = ord(i)&0xf
    res.append(tab[first_poz])
    res.append(tab[second_poz])
    ind+=1
if ind > 1107: # while (uVar5 < 0x454);
    break
for i in res:
    p=p+i
print(p)
    
```

An example of the output is shown below.

Press enter or click to view image in full size


```

Decompile: obtain_argv_q - (350000-layer-2.bin)
1
2 LPSTR __cdecl obtain_argv_q(LPCSTR param_1)
3
4 {
5     LPSTR argv_1;
6     int iVar1;
7     LPSTR lpBuffer;
8     CHAR value_of_argv1 [32];
9
10    argv_1 = StrStrA(param_1, "-q=");
11    if (argv_1 != (LPSTR)0x0) {
12        iVar1 = StrToIntA(argv_1 + 3);
13        wsprintfA(value_of_argv1, "%u", iVar1);
14        lpBuffer = (LPSTR)allocate_heap_memory(0x8e8);
15        argv_1 = lpBuffer;
16        if (lpBuffer != (LPSTR)0x0) {
17            argv_1 = (LPSTR)GetEnvironmentVariableA(value_of_argv1, lpBuffer, 0x8e8);
18            if (argv_1 != (LPSTR)0x0) {
19                argv_1 = (LPSTR)FUN_0040104c((int)lpBuffer, 0x454, &CHAR_??_00403000);
20            }
21            heap_free_memory(lpBuffer);
22        }
23    }
24    return argv_1;
25 }
26

```

The next step is to run the process again with the parameters that were checked in the above conditions (-q=[int]). In the second start of the process, after passing the conditions, we come to the function.

create_self_process_with_additional_params() (org. FUN_0040124a()), which launches a new process of svchost.exe

Press enter or click to view image in full size

			make_svchost.exe_string	XREF[1]:
0040122d	8b 44 24 04	MOV	EAX, dword ptr [ESP + param_1]	
			svchost.exe	
00401231	c7 00 5c	ds	C7, "\0\\svc", C7, "@", 04, "host", C7, "@\b.exe"	
	73 76 63			
	c7 40 04 ...			
00401245	c6 40 0c 00	MOV	byte ptr [EAX + 0xc], 0x0	
00401249	c3	RET		

Anti-debugging, or GetNativeSystemInfo.

Before the creation of the svchost process the malware uses an anti-debugging method. Malware uses a known method to this end called "GetNativeSystemInfo" (User32.dll).

Press enter or click to view image in full size

```

Decompile: FUN_00401706 - (350000-layer-2.bin)
1
2 uint FUN_00401706(void)
3
4 {
5     short local_28 [18];
6
7     FUN_00401e58((undefined *)local_28,0x24);
8     GetNativeSystemInfo((LPSYSTEM_INFO)local_28);
9     return (uint)(local_28[0] == 9);
10 }
11

```

The method is called in a function FUN_00401706(), which is called in FUN_004015a9()

Press enter or click to view image in full size

```

Decompile: FUN_004015a9 - (350000-layer-2.bin)
1
2 /* WARNING: Globals starting with '_' overlap smaller symbols at the same address */
3
4 void FUN_004015a9(void)
5
6 {
7     uint uVar1;
8     int *piVar2;
9     uint uVar3;
10    uint uVar4;
11    uint uVar5;
12    uint uVar6;
13    uint uVar7;
14    uint uVar8;
15    uint uVar9;
16    uint uVar10;
17    uint uVar11;
18    int iVar12;
19    longlong lVar13;
20
21    _DAT_00403098 = 0;
22    uVar1 = FUN_00401706();
23    if (uVar1 != 0) {
24        lVar13 = empty_fun();
25        iVar12 = (int)((ulonglong)lVar13 >> 0x20);
26        uVar1 = (uint)lVar13;
27        if ((lVar13 != 0) && (piVar2 = (int *)allocate_heap_memory2(uVar1,iVar12), piVar2 != (int *)0x0)
28            ) {
29            uVar3 = FUN_0040142d(piVar2,uVar1,iVar12,0xed46dfd2,(int *)&DAT_004030e0,(int)&DAT_00403120);
30            uVar4 = FUN_0040142d(piVar2,uVar1,iVar12,0x434c7242,(int *)&DAT_004030e8,(int)&DAT_00403126);
31            uVar5 = FUN_0040142d(piVar2,uVar1,iVar12,0xd510f438,(int *)&DAT_004030a8,(int)&DAT_004030f6);
32            uVar6 = FUN_0040142d(piVar2,uVar1,iVar12,0xbe06b948,(int *)&DAT_004030b8,(int)&DAT_00403102);
33            uVar7 = FUN_0040142d(piVar2,uVar1,iVar12,0x3d3fb58a,(int *)&DAT_004030d0,(int)&DAT_00403114);
34            uVar8 = FUN_0040142d(piVar2,uVar1,iVar12,0xc33bb247,(int *)&DAT_004030a0,(int)&DAT_004030f0);
35            uVar9 = FUN_0040142d(piVar2,uVar1,iVar12,0xe4f137ca,(int *)&DAT_004030d8,(int)&DAT_0040311a);

```

According to the documentation [MSDN](#) we can obtain SYSTEM_INFO structure

Press enter or click to view image in full size

SYSTEM_INFO structure

Contains information about the current computer system. This includes the architecture and type of the processor, the number of processors in the system, the page size, and other such information.

Syntax

```
typedef struct _SYSTEM_INFO {
    union {
        DWORD dwOemId;
        struct {
            WORD wProcessorArchitecture;
            WORD wReserved;
        };
    };
    DWORD dwPageSize;
    LPVOID lpMinimumApplicationAddress;
    LPVOID lpMaximumApplicationAddress;
    DWORD_PTR dwActiveProcessorMask;
    DWORD dwNumberOfProcessors;
    DWORD dwProcessorType;
    DWORD dwAllocationGranularity;
    WORD wProcessorLevel;
    WORD wProcessorRevision;
} SYSTEM_INFO;
```

The code shows that the returned value is compared with 0x9 which means the comparison of the processor architecture.

PROCESSOR_ARCHITECTURE_AMD64 (AMD or Intel) = 9

Press enter or click to view image in full size

wProcessorArchitecture

The processor architecture of the installed operating system. This member can be one of the following values.

Value	Meaning
PROCESSOR_ARCHITECTURE_AMD64 9	x64 (AMD or Intel)
PROCESSOR_ARCHITECTURE_ARM 5	ARM
PROCESSOR_ARCHITECTURE_ARM64 12	ARM64
PROCESSOR_ARCHITECTURE_IA64 6	Intel Itanium-based
PROCESSOR_ARCHITECTURE_INTEL 0	x86
PROCESSOR_ARCHITECTURE_UNKNOWN 0xffff	Unknown architecture.

Once the malware executes the svchost process it injects a final stage of itself into the process. I will cover this technique and continue out analysis in Part Two.

Source: <https://medium.com/@dawid.golak/icedid-aka-bokbot-analysis-with-ghidra-560e3eccb766>