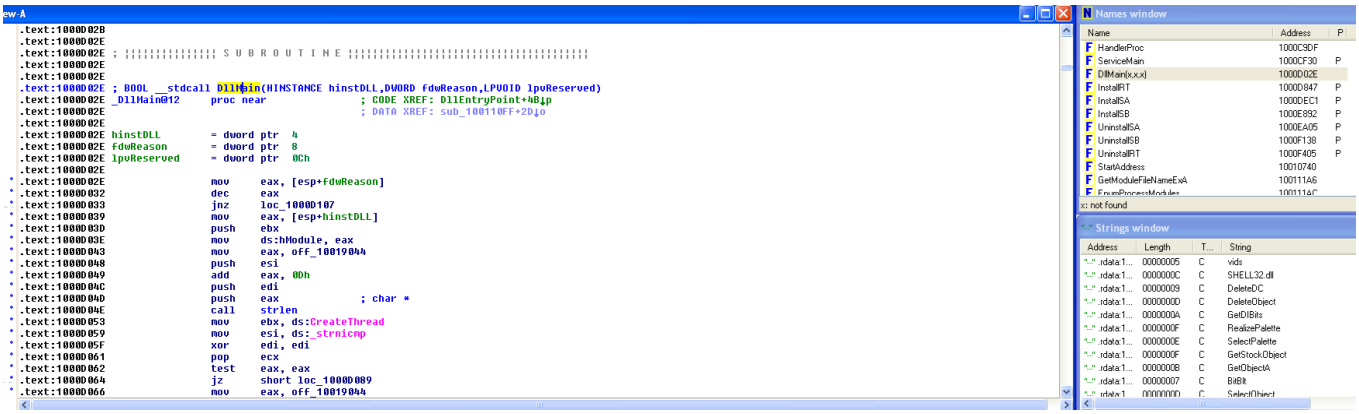
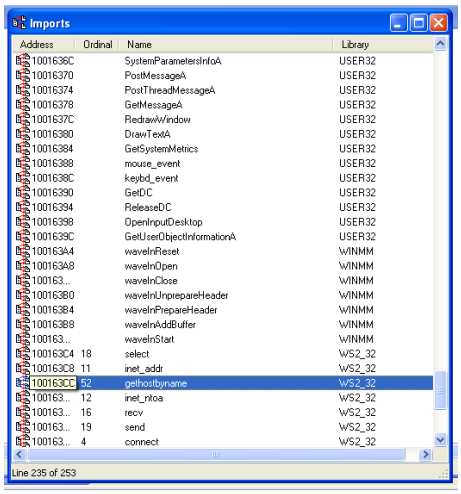
Note: We used IDA version 5.0 on Windows XP machine, and simultaneously used IDA version 8.3 on Windows 10 machine to see if there were any differences. The results in this report will mainly be from Windows XP as requested in the project, but any major distinctions will be highlighted throughout the report.

PART(1): MAL01.dll:

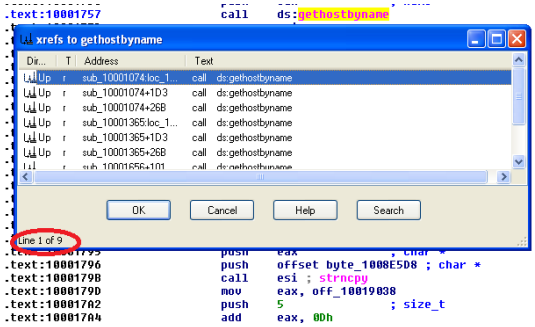
1. The address for DllMain: **0x1000D02E**



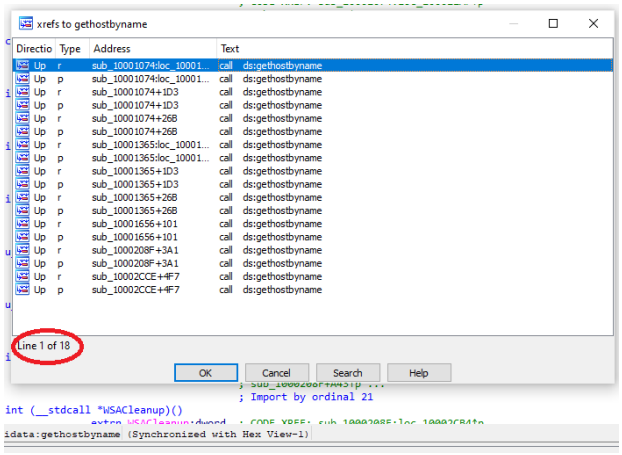
2. gethostbyname is located at **0x100163CC**



3. On WinXP: 9 xrefs showed up in the xrefs window (type r: read):

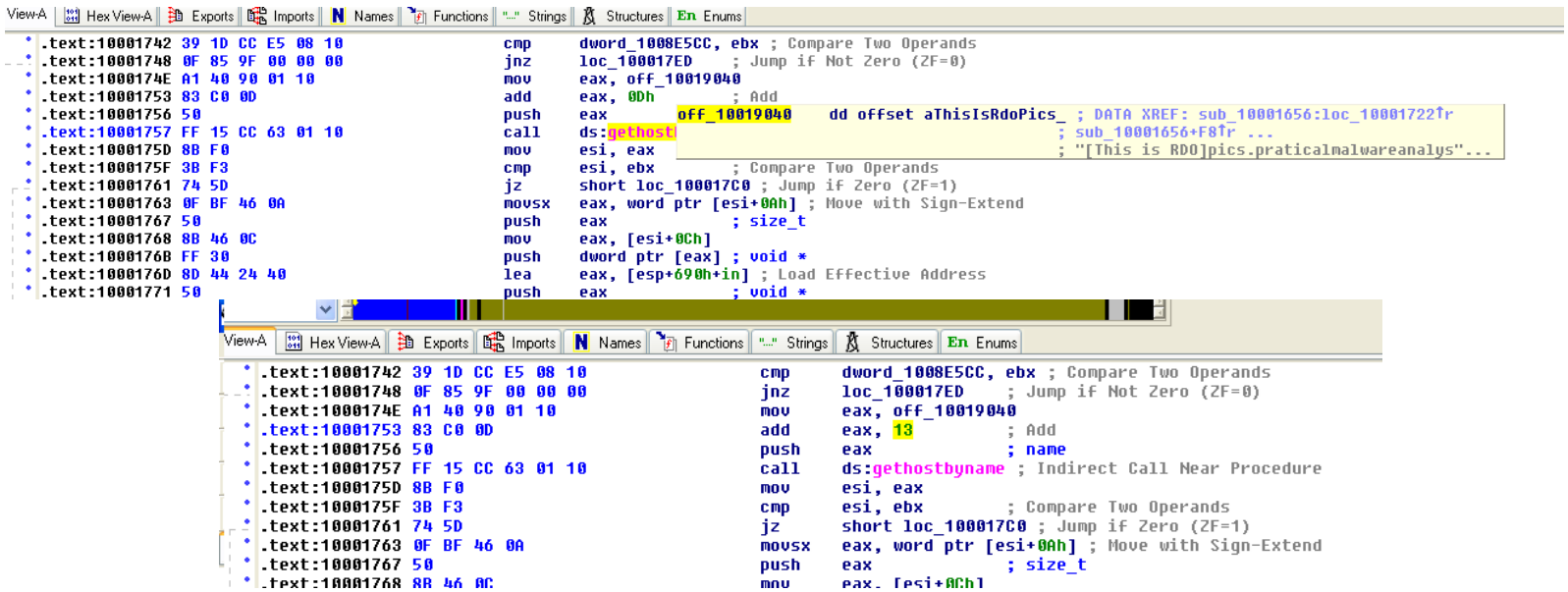


However, on Windows 10, IDA 8.3 showed 18 xrefs (9 of type r: read, and 9 of type p: near (intrasegment) calls):

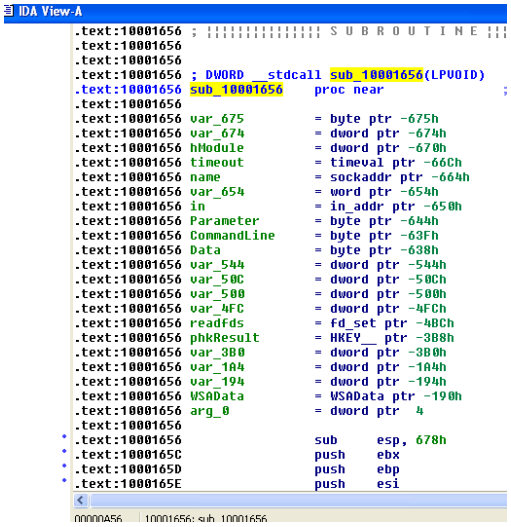


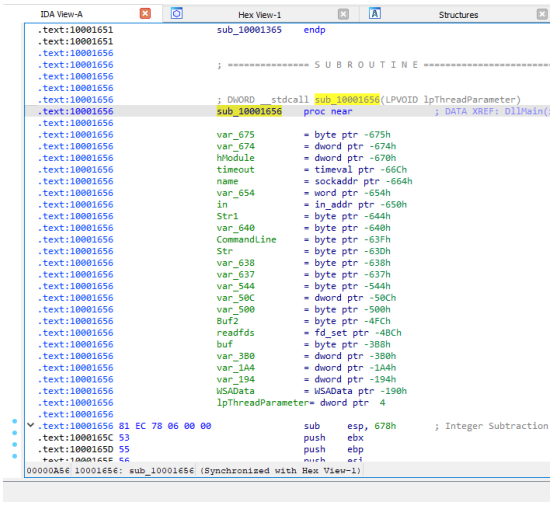
After looking into the IDA documentation, it seems like the type attributes **p** for near (intrasegment) calls and **P** for far (intersegment) calls were introduced in IDA version 6.5, which is why they didn't show up on our WinXP machine. As we can see from the screenshot above, these p type xrefs refer to the same addresses as the r types, it's there to distinguish between near and far calls.

4. As we can see from the screenshot, gethostbyname takes one parameter: the contents of eax, at **off_10019040**, which points to a variable aThisIsRdoPicsP which contains "[This is RD0]pics.practicalmalwareanalysis.com". This is moved into eax, and then 0Dh (0Dh = 13 decimal) is added to eax, this will move the pointer 13 characters inside the contents, which will skip "[This is RD0]" , so the DNS request will be made exactly for: **pics.practicalmalwareanalysis.com**.



5. Number of local variables differed drastically in IDA 5.0 (***left***) vs. IDA 8.3 (***right***):

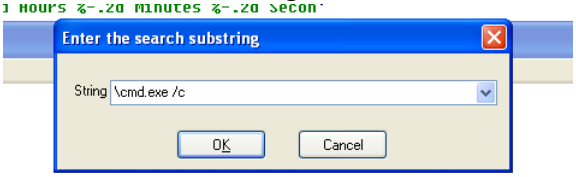




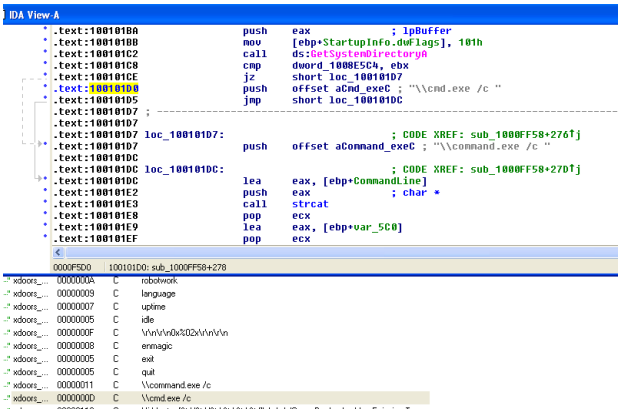
Local variables have negative offsets: IDA 5.0 on WinXp recognized 20 variables, while IDA 8.3 was able to recognize 23 in total. Some variables in 5.0 were also renamed in 8.3 (e.g. Parameter ⟶ Str1).

6. Parameters have positive offsets, only 1 parameter was recognized.

7. After searching for "\cmd.exe /c" in the strings window:

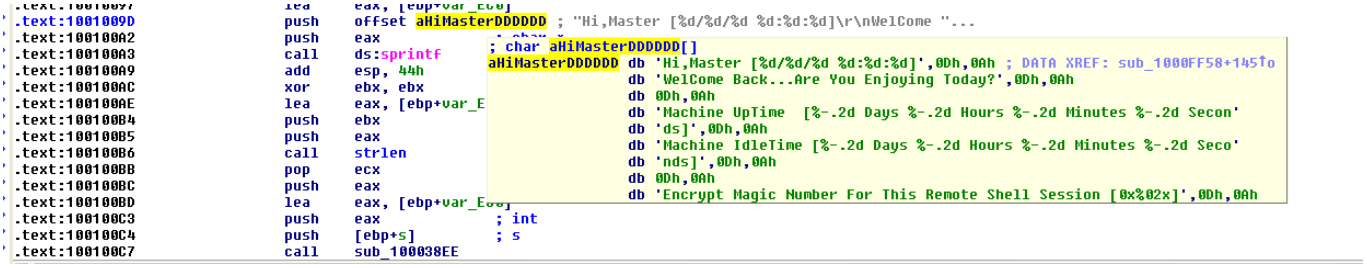


This is stored as aCmdExecS, within subroutine sub_1000FF58 at the offset **0x100101D0**:

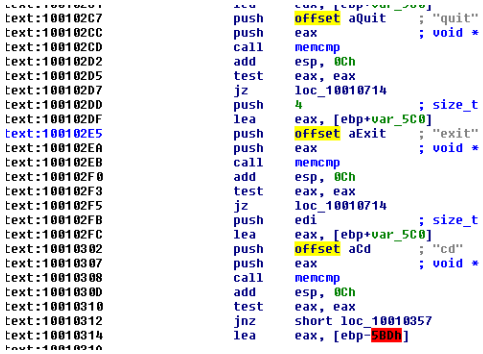


8. cmd.exe will be opened, then /c carries out the command specified then terminates it, so we will look for something it will possibly execute. The beginning of the subroutine doesn't contain anything suspicious, until we come across an offset called

aHiMasterDDDDDDD at 0x1001009D: this contains a long list of strings that have to do with system time (Machine Uptime, Machine IdleTime), and the last line contains “Encrypt Magic Number For This Remote Shell Session”:



The subroutine contains some other offsets (aQuit, aExit), these are part of any command-line execution.

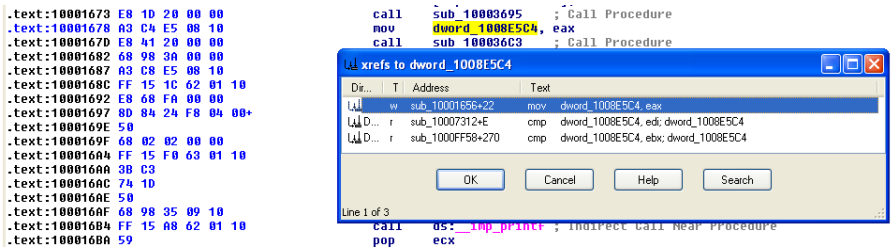


However, some interesting offsets also appear, that are not actually part of any ordinary command-line execution:

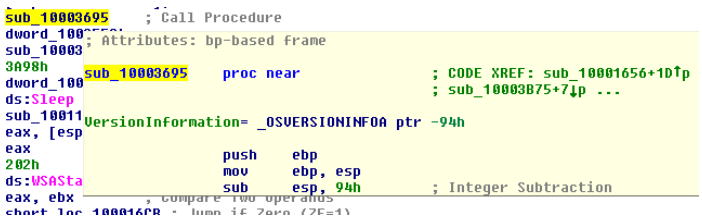
```
:ext:100105B2                   lea     eax, [ebp+var_6C0]
:ext:100105B8                   push    offset aIexplore_exe ; "iexplore.
:ext:100105BD                   push    eax             ; char *
:ext:100105BE                   call    strcpy
:ext:100105C3                   pop     ecx
:ext:100105C4                   pop     ecx
:ext:100105C5
:ext:100105C5 loc_100105C5:                             ; CODE XREF: sub_
:ext:100105C5                                           ; sub_1000FF58+65
:ext:100105C5                   push    [ebp+s]         ; s
:ext:100105C8                   lea     eax, [ebp+var_6C0]
:ext:100105CE                   push    eax             ; char *
:ext:100105CF                   call    sub_1000D5B0
:ext:100105D4                   pop     ecx
:ext:100105D5                   jmp     loc_100103F6
```

aRobotwork, aInject, aIexplore_exe: all of these are examples of added functions. aInject indicates process injection: process injection is basically running code in the context of another process, which may allow access to the process's memory, system/network resources, and possibly elevated privileges. Execution via process injection may also evade detection from security products since the execution is masked under a legitimate process. (1)

9. After looking at the xrefs of dword_1008E5C4, we can see it is referenced at sub_10001656 type w (write), with eax:

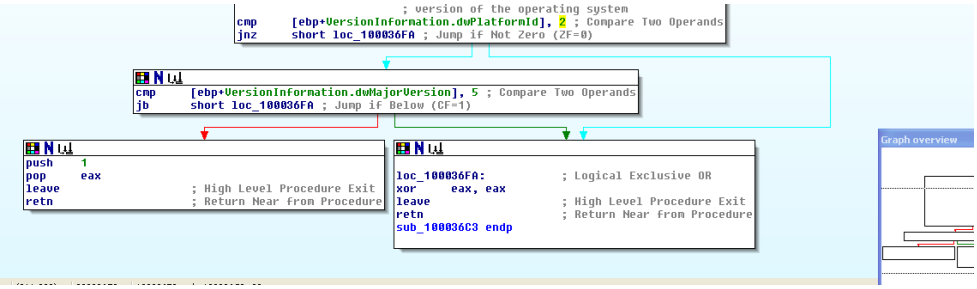


Right before that, there is a call to sub_1003695:

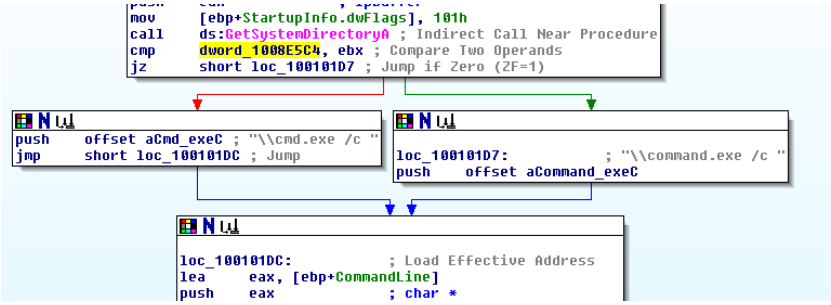


As we can see from the screenshot, this procedure checks the system information, using the structure OSVERSIONINFOA, which contains OS version information. According to Microsoft, this structure is used with the GetVersionEx function, to decide if the OS used is Windows 7, Windows Server 2008, Windows Vista, Windows Server 2003, Windows XP, or Windows 2000. (2)

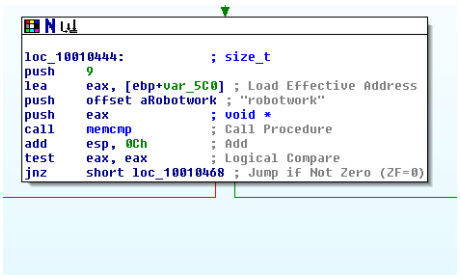
Switching to graph mode for easier understanding, a comparison occurs between VersionInformation.dwPlatfromId and 2, and according to Microsoft's platform IDs, it is checking if the OS is Windows NT or later. (3) :



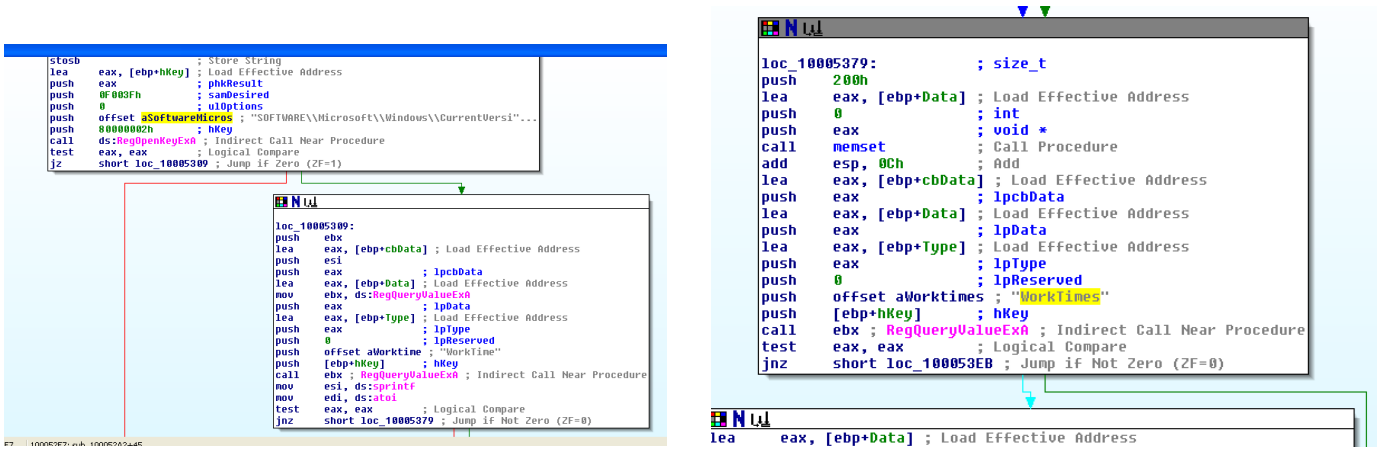
Back to the original comparison, dword_1008E5C4 at 0x100101C8 will decide if \cmd.exe /c is pushed: if the OS is Windows NT or later, it will be pushed, if not then it is \command.exe /c. So this is used to choose the correct command prompt to use based on the OS:



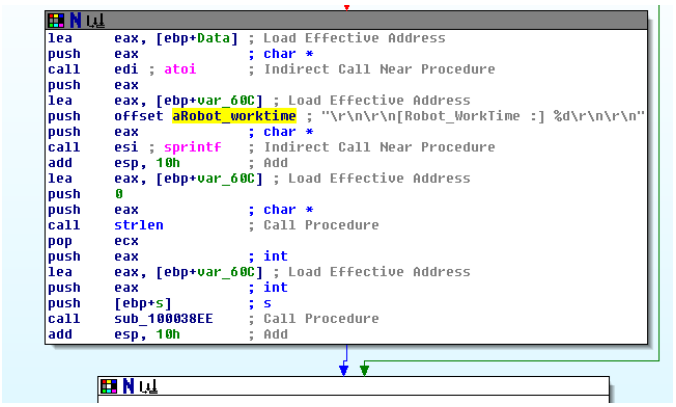
10. If the string comparison is successful, it returns 0, so the jump (jnz) will not execute and it will follow the red path:



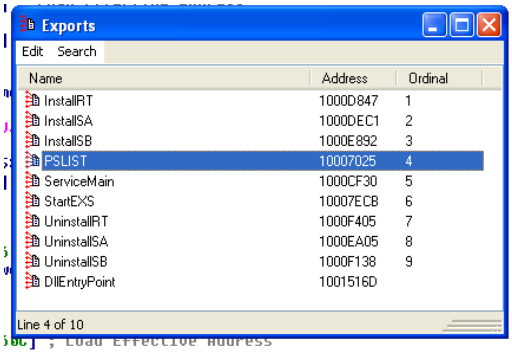
The red path leads to a new subroutine sub_100052A2, which has registry keys SOFTWARE\Microsoft\Windows\CurrentVersion: WorkTime and WorkTimes. This function is looking for values within these keys using RegValueExA:



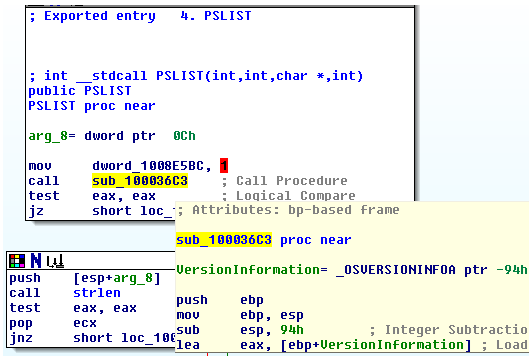
These registry keys WorkTime and WorkTimes are requested and their values are displayed as part of aRobotWorktimes offset addresses ([Robot_WorkTimes:] %d].



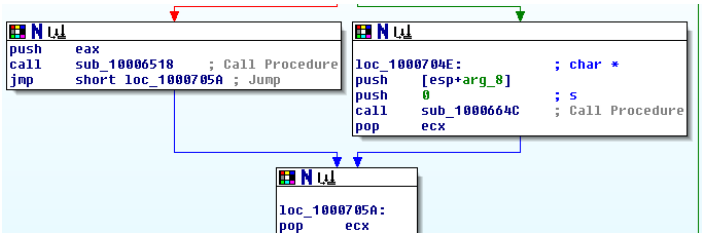
11. The exports window shows us the address of PSLIST:



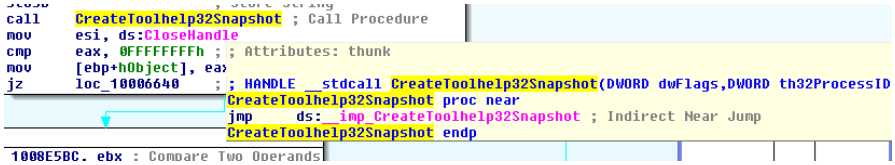
After navigating there, we can see that there are three subroutines associated with PSLIST. The first one has to do with OS information, similar to the one we saw before, but this one also checks if dwMajorVersion is 5 (dwMajorVersion includes major and minor version numbers and information about product suites [(4)]):



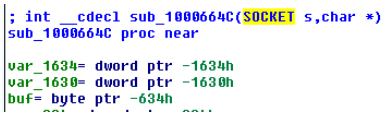
Depending on whether or not the dwMajorVersion is 5, it will call one of the other two subroutines: sub_10006518 or sub_1000664C:



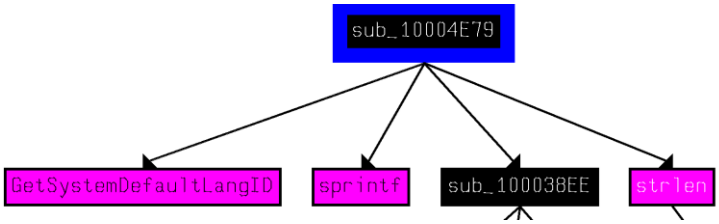
Both of these use CreateToolhelp32Snapshot to take a snapshot of the associated information, and then they execute commands to query information about the running processes IDs, names and number of threads used:



The difference between them is that sub_1000664C also includes SOCKET s, to send the output:



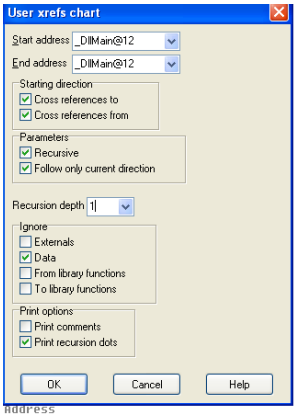
12. Graphing the xrefs from sub_10004E79:

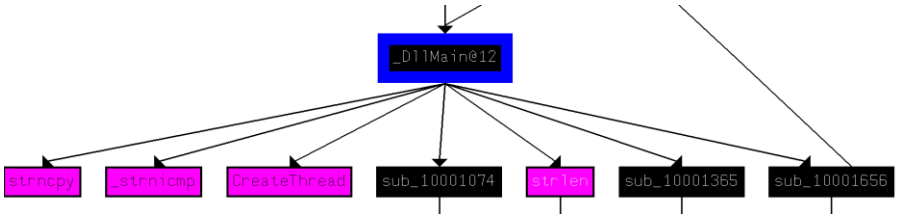


GetSystemDefaultLangId determines system language, the rest of the functions do some string operations. We can also see the offset aLanguageId0xX in the subroutine:

Based on this information, we can rename sub_10004E79 to **getSystemLanguage**.

13. To see the functions that are directly called, we used the user xrefs chart and set the start and end addresses to _DllMain@12:





A total of 4 API functions are *directly* called. At a depth of 2, a total of 32 API functions are called (gethostbyname, CreateThread, WSAStartup, send, socket, connect, LoadLibraryA).

14. a. move ax, off_10019020: aThisIsCti30 is moved into eax, which has "[This is CTI]30".
b. add eax, 0Dh: the pointer is moved 13 characters along eax (0Dh=13), leaving eax with "30".
c. call atoi: converts this string into an integer.
d. imul eax, 3E8h: eax is multiplied by 1 second (3E8h = 1000 = 1 second), which equals 30 seconds.
eax is then pushed, so the Sleep function will sleep for 30 seconds:

```
mov     eax, off_10019020
add     eax, 13             ; Add
push    eax                 ;off_10019020    dd offset unk_100192AC  ;
call    ds:atoi             ;                                        ;
imul    eax, 1000           ; Signed Multiply
pop     ecx
push    eax                 ; dwMilliseconds
call    ds:Sleep            ; Indirect Call Near Procedure
xor     ebp, ebp            ; Logical Exclusive OR
jmp     loc_100010B4        ; Jump
```

15. The three parameters af, type, and protocol:

```
mov     edi, eax
cmp     edi, 0; SOCKET __stdcall socket(int af,int type,int protocol)
jnz     short                 extrn socket:dword      ; DATA XREF: sub_10001656+ABTr
call    ds:WSA                                        ; sub_1000208F+3F4Tr ...
push    eax
push    offset aSocketGetlaste ; "socket() GetLastError reports %d\n"
call    ds:__imp_printf ; Indirect Call Near Procedure
pop     ecx
pop     ecx
```

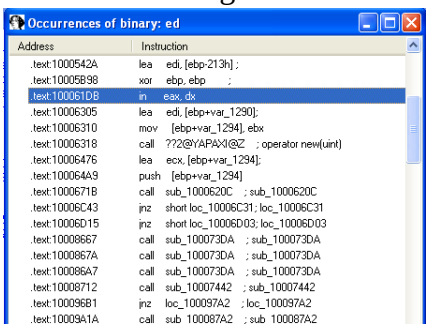
The values for these parameters were pushed to the stack just before; af=2, type=1 and protocol=6:

```
                                ; sub_10001
push    6                       ; protocol
push    1                       ; type
push    2                       ; af
call    ds:socket               ; Indirect
mov     edi, eax
cmp     edi, 0FFFFFFFFh ; Compare T
jnz     short loc_10001722 ; Jump i
call    ds:WSAGetLastError ; Indire
push    eax
push    offset aSocketGetlaste ; "s
call    ds:__imp_printf ; Indirect
```

16. After looking at Microsoft's socket documentation [5], **protocol** value 6 corresponds to IPROTO_TCP (TCP protocol), **type** 1 means the type of socket used is SOCK_STREAM often used with IPv4 address family, and **af** (address family specification) of value 2 corresponds to AF_INET (IPv4). So we can determine it is TCP IPv4. Using the convert to symbolic constant utility to rename the parameters:

```
push    IPPROTO_TCP      ; protocol
push    SOCK_STREAM      ; type
push    AF_INET          ; af
call    ds:socket        ; Indirect Call Near Procedure
mov     edi, eax
cmp     edi, 0FFFFFFFFh ; Compare Two Operands
jnz     short loc_10001722 ; Jump if Not Zero (ZF=0)
call    ds:WSAGetLastError ; Indirect Call Near Procedure
push    eax
push    offset aSocketGetlaste ; "socket() GetLastError rep
call    ds:__imp_printf ; Indirect Call Near Procedure
pop     ecx
pop     ecx
```

17. Searching for the occurrence of 0xed, we can see this in instruction is used with the string VMXh to determine if this malware is running inside VMWare:



This is located in sub_10006196, taking a look at the xrefs of this function, we see that there are 3 xrefs: all of these contain **aFoundVirtualMa**, which indicates this malware will stop installing if it detects the existence of a virtual machine ("Found Virtual Machine,Install Cancel."):

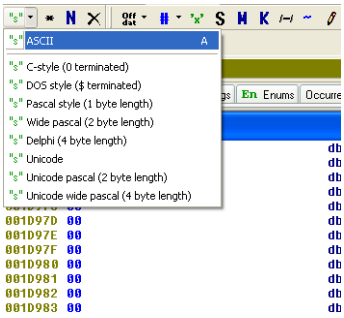
```
call    sub_10006196    ; Call Procedure
test    al, al          ; Logical Compare
jz      short loc_1000DF08 ; Jump if Zero (ZF=1)

                        ; CODE XREF: InstallSA+1E↑j
push    offset unk_1008E5F0 ; char *
call    sub_10003592    ; Call Procedure
mov     [esp+8+var_8], offset aFoundVirtualMa ; "Found Virtual Machine,Install Cancel."
call    sub_10003592    ; Call Procedure
pop     ecx
call    sub_10005567    ; Call Procedure
jmp     short loc_1000DF1E ; Jump
```

18. It seems like the beginning of some kind of encrypted message (using Hex view to see the full message):

```
1001D980  00 00 00 00 00 00 00 00  2D 31 3A 3A 27 75 3C 26  ........-1::'u<&
1001D990  75 21 3D 3C 26 75 37 34  36 3E 31 3A 3A 27 79 75  u!=<&u746>1::'yu
1001D9A0  26 21 27 3C 3B 32 75 31  30 36 3A 31 30 31 75 33  &!'<;2u106:101u3
1001D9B0  3A 27 75 05 27 34 36 21  3C 36 34 39 75 18 34 39  :'u.'46!<6490.49
1001D9C0  22 34 27 30 75 14 3B 34  39 2C 26 3C 26 75 19 34  "4'0u¶;49,&<&u↓4
1001D9D0  37 75 6F 7C 64 67 66 61  00 00 00 00 00 00 00 00  7uo|dgfa........
```

PART(2): MAL02.py:

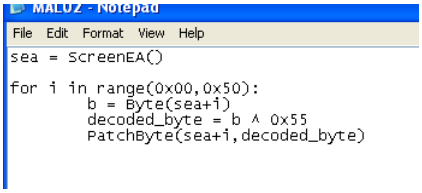
We can turn this data into a single ASCII string using IDA's convert to ASCII functionality, by clicking on this icon or simply hitting "A" on our keyboard:



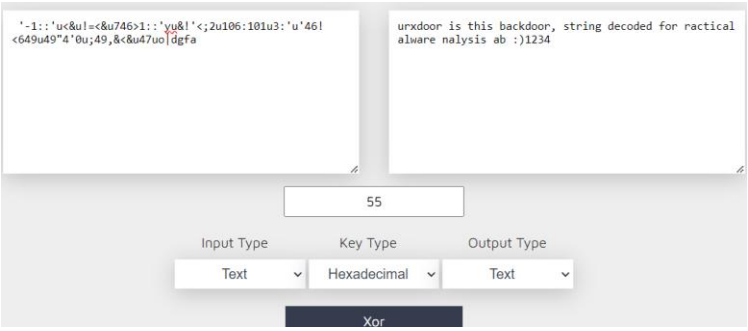
After converting, the results will look like this:

```
2D 31 3A 3A 27 75+a1UUU7461Yu2u10 db '-1::',27h,'u<&u!=<&u746>1::',27h,'yu&!',27h,'<;2u106:101u3:',27h,'u'
05                                db   5
27 34 36 21 3C 36+a466649u        db 27h,'46!<649u'
18                                db  18h
34 39 22 34 27 30+a4940u          db '49"4',27h,'0u'
14                                db  14h
3B 34 39 2C 26 3C+a49U            db ';49,&<&u'
19                                db  19h
34 37 75 6F 7C 64+a47uoDgfa       db '47uo|dgfa',0
```

We don’t have the IDA python plugin, but we will take a look at the script to try and understand what it does:



This script seems to perform an xor 55h, for 50h bytes from the current position (1001D988), we will use md5decrypt.net to decode the string we found above, using 55 as the XOR key:



“urxdoor is this backdoor, string decoded for ractical alware nalysis ab ☺1234”

As we can see, the text isn’t completely readable, but we can translate it to “Your door is this backdoor, string decoded for practical malware analysis lab”, indicating the use of a backdoor in this malware.

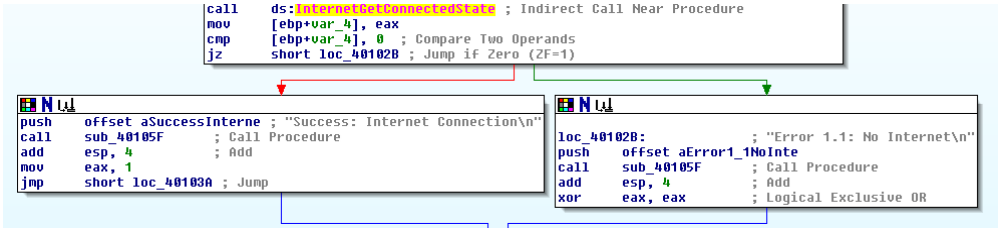
PART(3): MAL03.exe:

1. main contains subroutine sub_401000:

```
                        ; int __cdecl main(int argc,const char **argv,const char *envp)
ext:00401040            _main           proc near               ; CODE XREF: start+AF↓p
ext:00401040
ext:00401040            var_4           = dword ptr -4
ext:00401040            argc            = dword ptr  8
ext:00401040            argv            = dword ptr  0Ch
ext:00401040            envp            = dword ptr  10h
ext:00401040
ext:00401040 55                         push    ebp
ext:00401041 8B EC                      mov     ebp, esp
ext:00401043 51                         push    ecx
ext:00401044 E8 B7 FF FF FF             call    sub_401000      ; Call Procedure
ext:00401049 89 45 FC                   mov     [ebp+var_4], eax
ext:0040104C 83 7D FC 00 ; ||||||||||||||| S U B R O U T I N E |||||||||||||||||||||||||||||||||||||||
ext:00401050 75 04                      ; Attributes: bp-based frame
ext:00401052 33 C0
ext:00401054 EB 05                      sub_401000      proc near               ; CODE XREF: _main+4↓p
ext:00401056
ext:00401056                            var_4           = dword ptr -4
ext:00401056 B8 01 00 00
ext:0040105B                            push    ebp
ext:0040105B                            mov     ebp, esp
ext:0040105B 8B E5                      mov     esp, ebp
ext:0040105D 5D                         pop     ebp
```

After taking a look at this subroutine, we can see it has an external API call InternetGetConnectedState: according to Microsoft, this function returns TRUE if the system is connected to the internet, and FALSE if it isn't [6]:

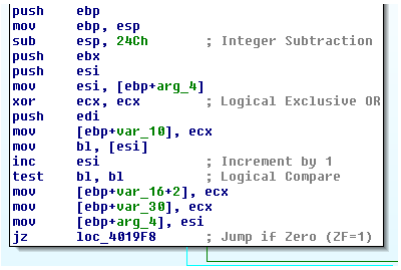
```
sub_401000      proc near               ; CODE XREF: _main+4↓p

var_4           = dword ptr -4

                push    ebp
                mov     ebp, esp
                push    ecx
                push    0               ; dwReserved
                push    0               ; lpdwFlags
                call    ds:InternetGetConnectedState ; Indirect Call Near Procedure
                mov     [ebp+var_4], eax
                cmp     [ebp+var_4], 0  ; Compare Two Operands
                jz      short loc_40102B ; Jump if Zero (ZF=1)
                push    offset aSuccessInterne ; "Success: Internet Connection\n"
                call    sub_40105F      ; Call Procedure
                add     esp, 4          ; Add
                mov     eax, 1
                jmp     short loc_40103A ; Jump
```



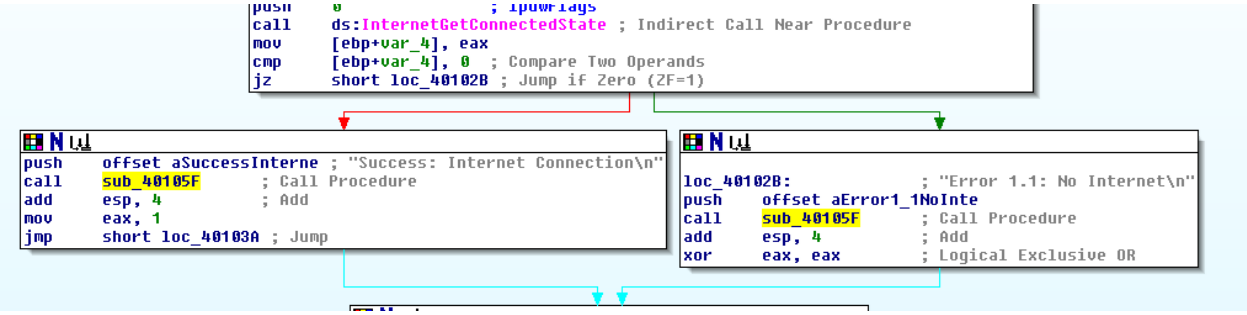
2. sub_40105F:

```
sub_40105F proc near

arg_0= dword ptr  0Ch
arg_4= dword ptr  10h

push    ebx
push    esi
mov     esi, offset unk_407098
push    edi
push    esi
call    __stbuf         ; Call Procedure
mov     edi, eax
lea     eax, [esp+8+arg_4] ; Load Effective Address
push    eax             ; int
push    [esp+0Ch+arg_0] ; int
push    esi             ; FILE *
call    sub_401282      ; Call Procedure
push    esi
push    edi
mov     ebx, eax
call    __ftbuf         ; Call Procedure
add     esp, 18h        ; Add
mov     eax, ebx
pop     edi
pop     esi
pop     ebx
retn                    ; Return Near from Procedure
sub_40105F endp
```

First, a call to a function __stbuf is made with 3 parameters, two of them integers and one of them is referencing a file (push esi ; FILE *). Then another subroutine is called: **sub_401282**, which performs some calculations and logical operations:



Finally, a third function __ftbuf is called which has two parameters: the reference to the file and the result of the first function call. By looking at the xrefs for this subroutine:



We can see this subroutine is called twice: if the result of the comparison above it is zero (green), it will be called and will print “Error 1.1: No Internet” and go to the next line. If it’s not zero (red), it will print “Success: Internet Connection” and go to the next line. This subroutine is a printf() function.

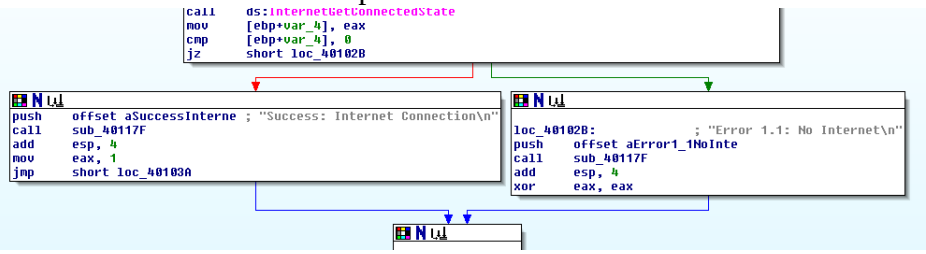
3. The purpose of this program is to check if there is an internet connection and print a statement as mentioned above, based on the results.

PART(4): MAL04.exe:

1. The first subroutine called by main sub_401000:

```
; int __cdecl main(int argc,const char **argv,const char *envp
_main           proc near               ; CODE XREF: start+AF

var_8           = byte ptr -8
var_4           = dword ptr -4
argc            = dword ptr  8
argv            = dword ptr  0Ch
envp            = dword ptr  10h

                push    ebp
                mov     ebp, esp
                sub     esp, 8
                call    sub_401000
                mov     [ebp+var_4], eax
                cmp     [ebp+var_4], 0
                jnz     short loc_401148
                xor     eax, eax
                jmp     short loc_40117B
```

This subroutine performs the same function as MAL03.exe, using InternetGetConnectedState and then prints if it is connected or not:

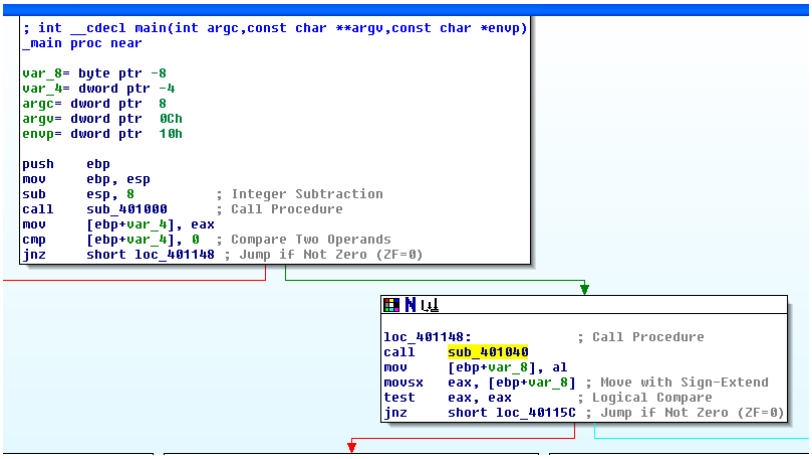


2. Same as the 2nd question in MAL03.exe:sub_40117F calls __stbuf, **sub_4013A2** and __ftbuf:

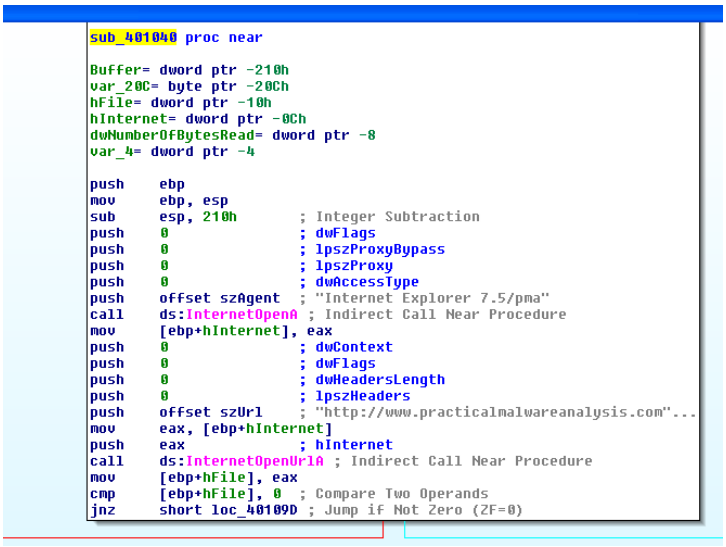
```
text:0040117F sub_40117F      proc near               ; CODE XREF: sub_4
text:0040117F                                         ; sub_401000+30Tp
text:0040117F
text:0040117F arg_0           = dword ptr  0Ch
text:0040117F arg_4           = dword ptr  10h
text:0040117F
text:0040117F                 push    ebx
text:00401180                 push    esi
text:00401181                 mov     esi, offset unk_407160
text:00401186                 push    edi
text:00401187                 push    esi
text:00401188                 call    __stbuf
text:0040118D                 mov     edi, eax
text:0040118F                 lea     eax, [esp+8+arg_4]
text:00401193                 push    eax             ; int
text:00401194                 push    [esp+0Ch+arg_0] ; int
text:00401198                 push    esi             ; FILE *
text:00401199                 call    sub_4013A2
text:0040119E                 push    esi
text:0040119F                 push    edi
text:004011A0                 mov     ebx, eax
text:004011A2                 call    __ftbuf
text:004011A7                 add     esp, 18h
text:004011AA                 mov     eax, ebx
text:004011AC                 pop     edi
text:004011AD                 pop     esi
text:004011AE                 pop     ebx
text:004011AF                 retn
text:004011AF sub_40117F      endp
text:004011AF
```

sub_4013A2 also contains some calculations and logical operations. Same as the previous one: checks if there is an internet connection.

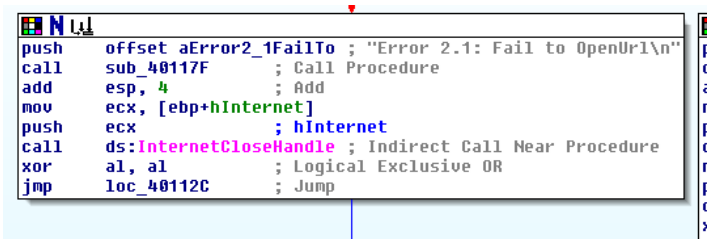
3. The second subroutine sub_401040 :



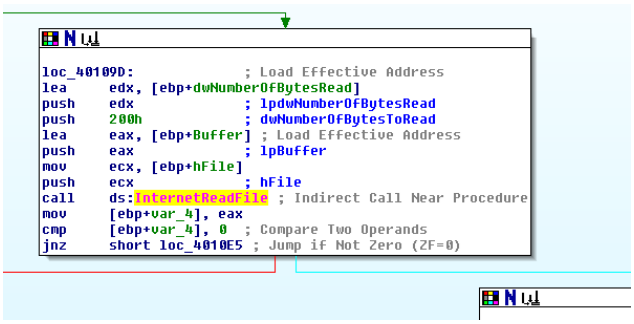
This subroutine can be reached if the comparison equals 0. By taking a look at this subroutine:



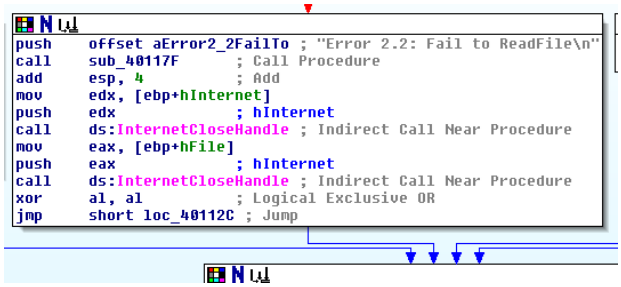
It contains two API calls: InternetOpenA, which is used to initiate an internet connection [7] and InternetOpenUrlA, which opens a resource specified by a complete FTP or HTTP URL [8], according to Microsoft. If this function is not given permission, the screenshot below shows that it will print “Error 2.1: Fail to OpenUrl” and move to the next line. We also see some strings: szAgent contains “Internet Explorer 7.5/pma” and szUrl contains “http://www.practicalmalwareanalysis.com/cc.htm”.



If the jump is taken, it contains another API: InternetReadFile, which will read data from a handle opened by the InternetOpenUrl function previously used, if the request is successful: [9]

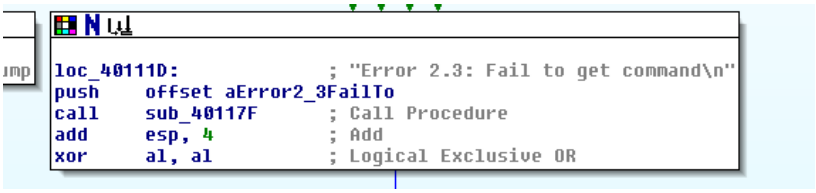


If the InternetReadFile function is unsuccessful, it will print “Error 2.2: Fail to ReadFile” and move to the next line:



We can conclude that this subroutine does a few things: first, it makes a query to "http://www.practicalmalwareanalysis.com/cc.htm" using the InternetOpenUrlA function. If the URL can be opened, the webpage is then read with the InternetReadFile function. This will attempt to read a command from the URL, if it’s successful, the subroutine will execute some comparisons, and some jumps and at the end, the string '<!--' is checked. Basically, this function makes a query to the website in order to receive commands to know what to do next. In order to read those commands, the file must start with '<!--'.

If these characters are not found, it will print: “Error 2.3: Fail to get command” and move to the next line:



InternetCloseHandle function is called after any of these error messages, to close any handles.

4. The code construct used in this subroutine is a string. This string is compared character to character using a set of ifs.

5. Network-based indicators: as can be seen above, URL related to the InternetOpenA and InternetOpenUrlA calls: Internet Explorer 7.5/pma and http://www.practicalmalwareanalysis.com/cc.htm.

6. This malware first checks if there is an internet connection, prints the subsequent message. If there is internet connection, it will attempt to download and read files.

Resources:

(1) Process Injection Part 1: The Theory

(2) OSVERSIONINFOA structure (winnt.h)

(3) PlatformID Enum

(4) OSVERSIONINFOEXA structure (winnt.h)

(5) socket function (winsock2.h)

(6) InternetGetConnectedState function (wininet.h)

(7) InternetOpenA function (wininet.h)

(8) InternetOpenURLA function (wininet.h)

(9) InternetReadFile function (wininet.h)