

Static Malware Analysis

CS 483: Digital Forensics

Lesson 8

Lesson Learning Objectives

- Hands-on reverse engineering using IDA Pro
- Pros and Cons of static analysis

Underpinning this lesson are program analysis techniques:

- 1. Control Flow (data flow, dependence, etc.)**
- 2. Conditional Constructs (if, switch, etc.)**
- 3. Loop structures (nested, etc.)**
- 4. Data structures (stacks, etc.)**

Tools are useful, but understanding the techniques are essential

Why Static Analysis?

- Can find precise location of vulnerabilities or capabilities
- Low-cost & low-threat
- Malware cannot evade if static
- Augments other sophisticated techniques (can help trouble shoot)

Why ^{not} Static Analysis?

- Time consuming
- Obfuscation and packing
- No run-time information
- Tedious
- Frustrating
- Makes me angry

Binary Dissassembly

Linear Sweep

1. Go through .text sections of the executable code and disassemble everything **sequentially**
2. Start at the first byte then decode each subsequent byte **until an illegal instruction is reached**
3. Susceptible to mistakes or intentionally placed bombs (**who would do such a thing!?**)

Objdump, WinDbg

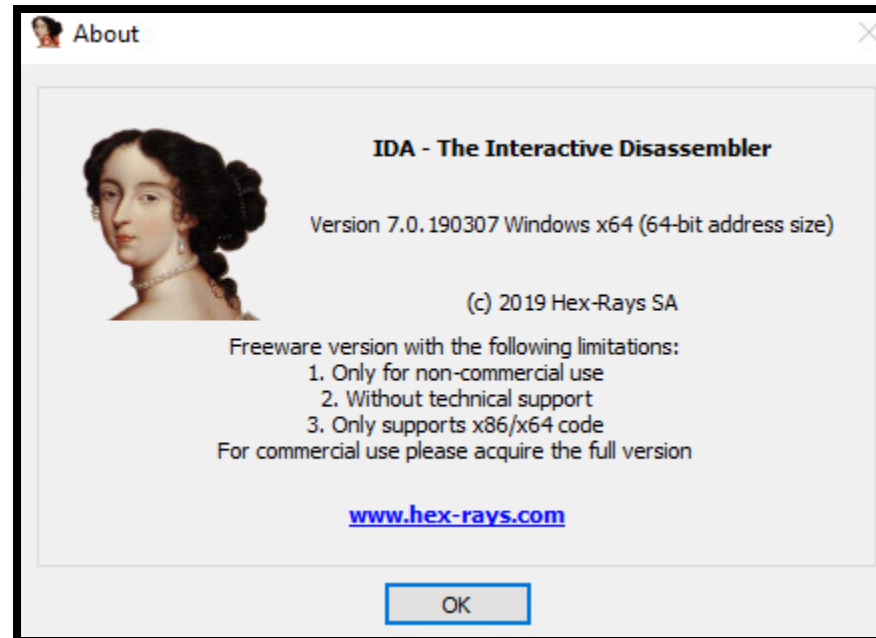
Recursive Descent

1. Each **jump and branch instructions** are followed
2. It's a **linear sweep** until a branch is encountered, then we follow it and **linear sweep** again
3. Susceptible to indirect jumps

Olly, IDA, Ghidra

Interactive DisAssembler (a tool)

- When IDA Pro starts it will ask you to start a new disassembly or open a previous one
- The defaults are almost always correct...unless you are dealing with nasty malware



Greencat (APT 1)

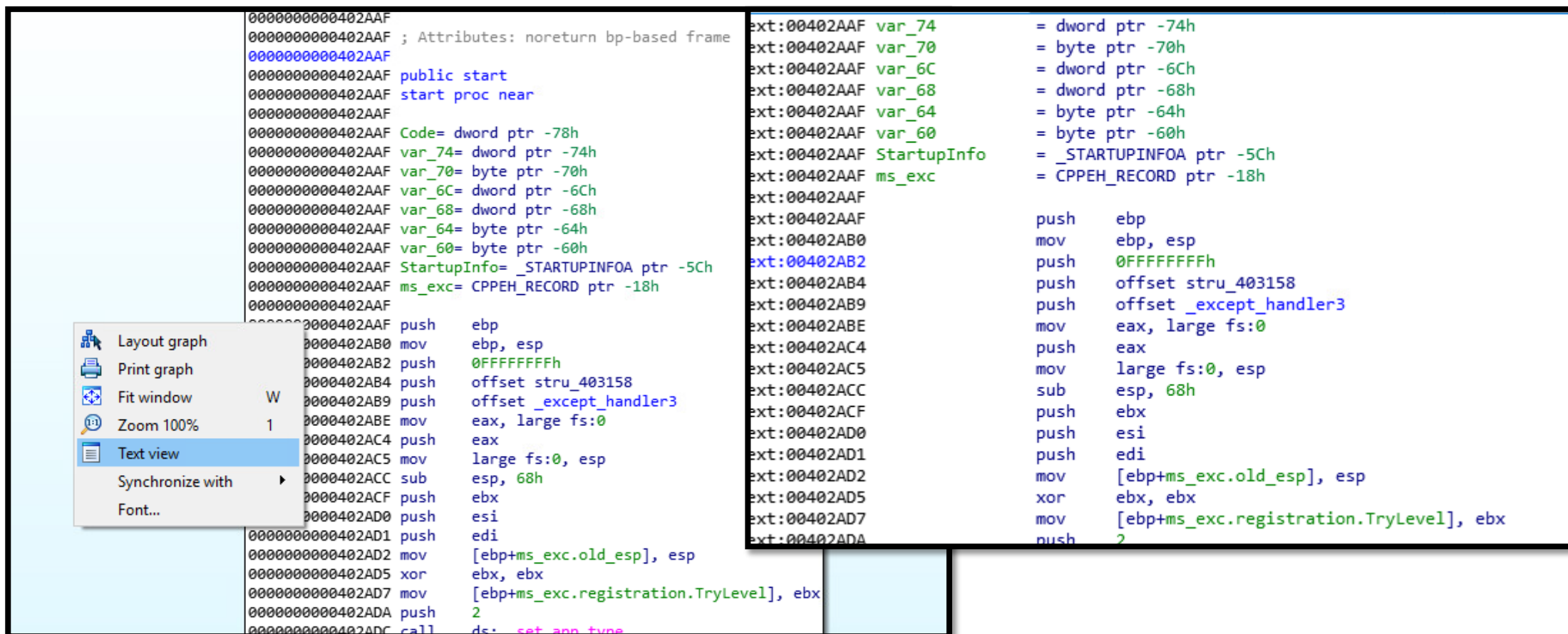
<https://github.com/fullerj/PMA/raw/main/webc2-greencat-2.7z>

It's “infected”, so be careful.



IDA will open in CFG view

- Right-click and select Text View for flat disassembled code (or Space bar)



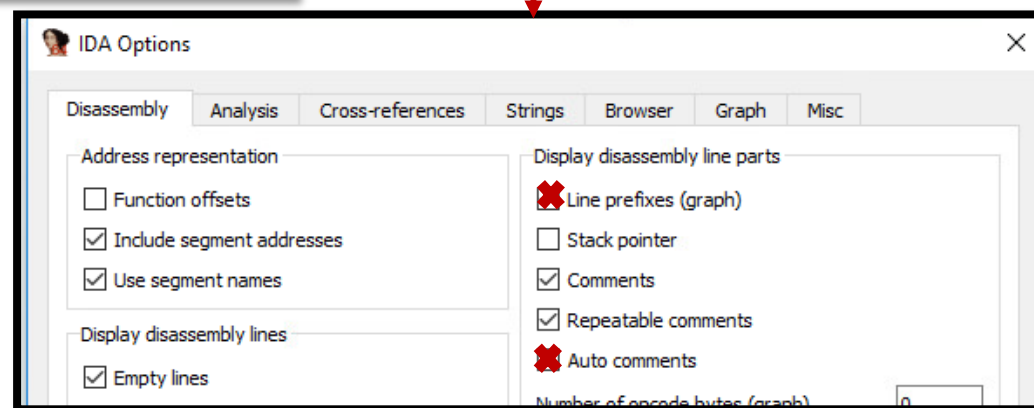
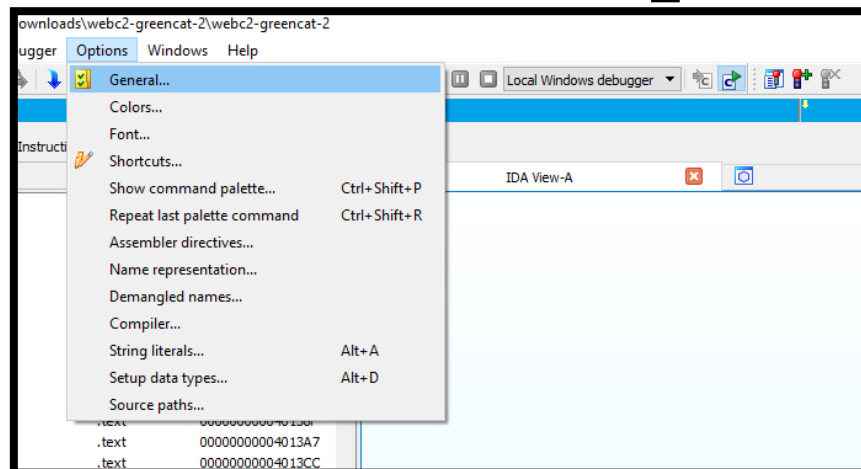
Pro tip #1: Add line prefixes

```
public start
start proc near

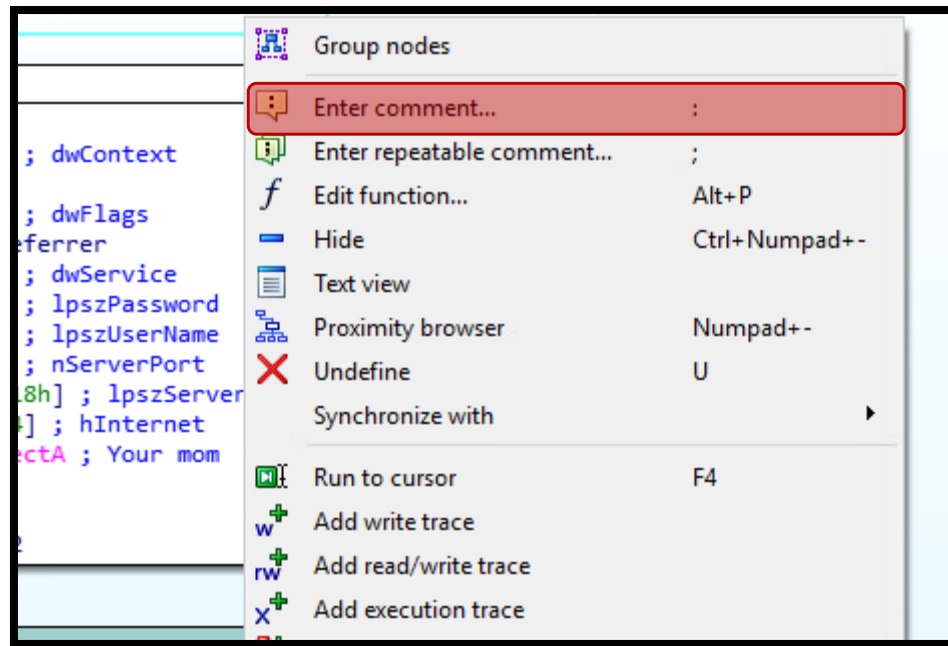
Code= dword ptr -78h
var_74= dword ptr -74h
var_70= byte ptr -70h
var_6C= dword ptr -6Ch
var_68= dword ptr -68h
var_64= byte ptr -64h
var_60= byte ptr -60h
StartupInfo= _STARTUPINFOA ptr -5Ch
ms_exc= CPPEH_RECORD ptr -18h
```

```
push    ebp
mov     ebp, esp
push    0FFFFFFFh
push    offset stru_403158
push    offset _except_handler3
mov     eax, large fs:0
push    eax
mov     large fs:0, esp
sub     esp, 68h
push    ebx
push    esi
push    edi
mov     [ebp+ms_exc], edi
```

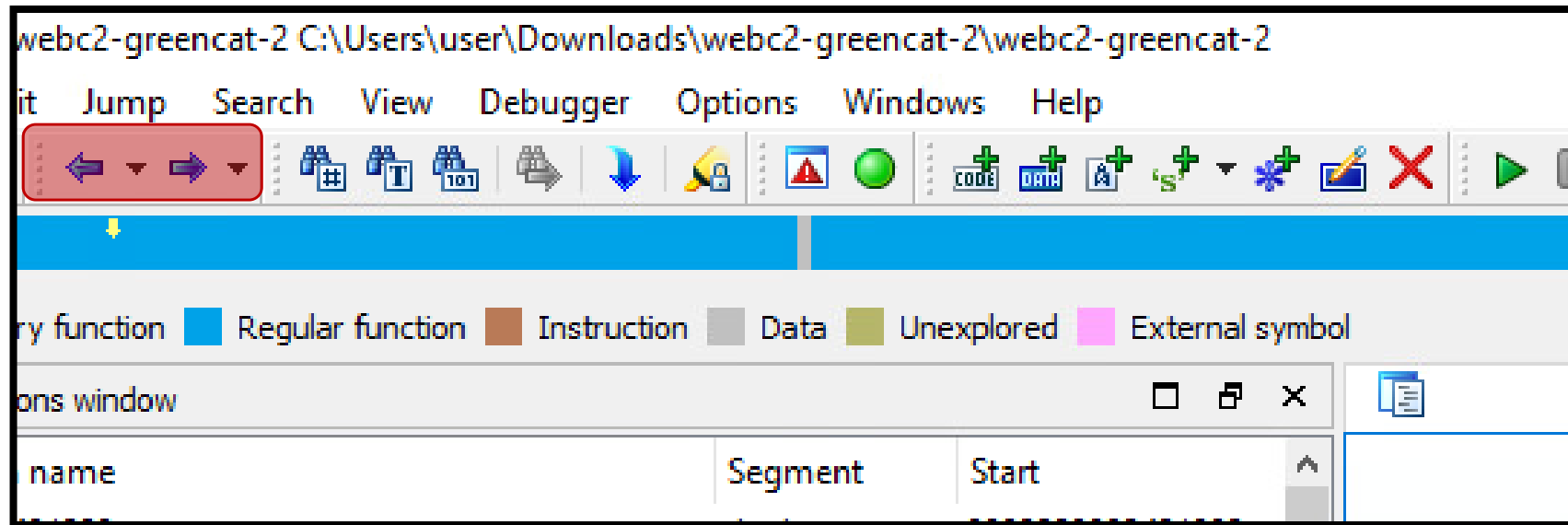
```
0000000000402AAF public start
0000000000402AAF start proc near
0000000000402AAF Code= dword ptr -78h
0000000000402AAF var_74= dword ptr -74h
0000000000402AAF var_70= byte ptr -70h
0000000000402AAF var_6C= dword ptr -6Ch
0000000000402AAF var_68= dword ptr -68h
0000000000402AAF var_64= byte ptr -64h
0000000000402AAF var_60= byte ptr -60h
0000000000402AAF StartupInfo= _STARTUPINFOA ptr -5Ch
0000000000402AAF ms_exc= CPPEH_RECORD ptr -18h
0000000000402AAF
0000000000402AAF push    ebp
0000000000402AB0 mov     ebp, esp
0000000000402AB2 push    0FFFFFFFh
0000000000402AB4 push    offset stru_403158
0000000000402AB9 push    offset _except_handler3
0000000000402ABE mov     eax, large fs:0
0000000000402AC4 push    eax
0000000000402AC5 mov     large fs:0, esp
0000000000402ACC sub     esp, 68h
0000000000402ACF push    ebx
0000000000402AD0 push    esi
0000000000402AD1 push    edi
```



Pro tip #2: Right-click and select comments or press “:”



Pro tip #3: Wait, where was I?



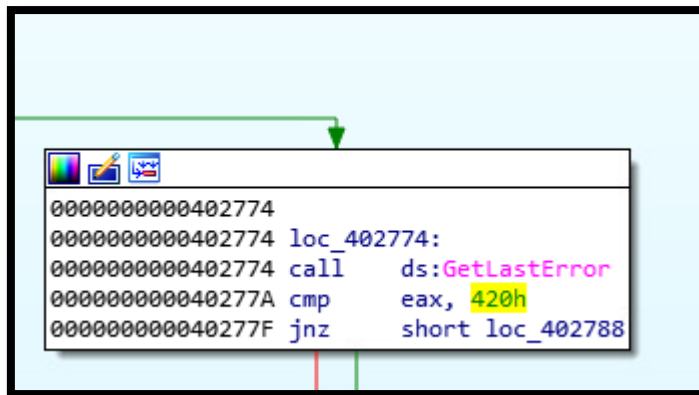
Pro tip #4: Rename Symbolic Constants

- IDA's FLIRT signatures know the arguments for common APIs

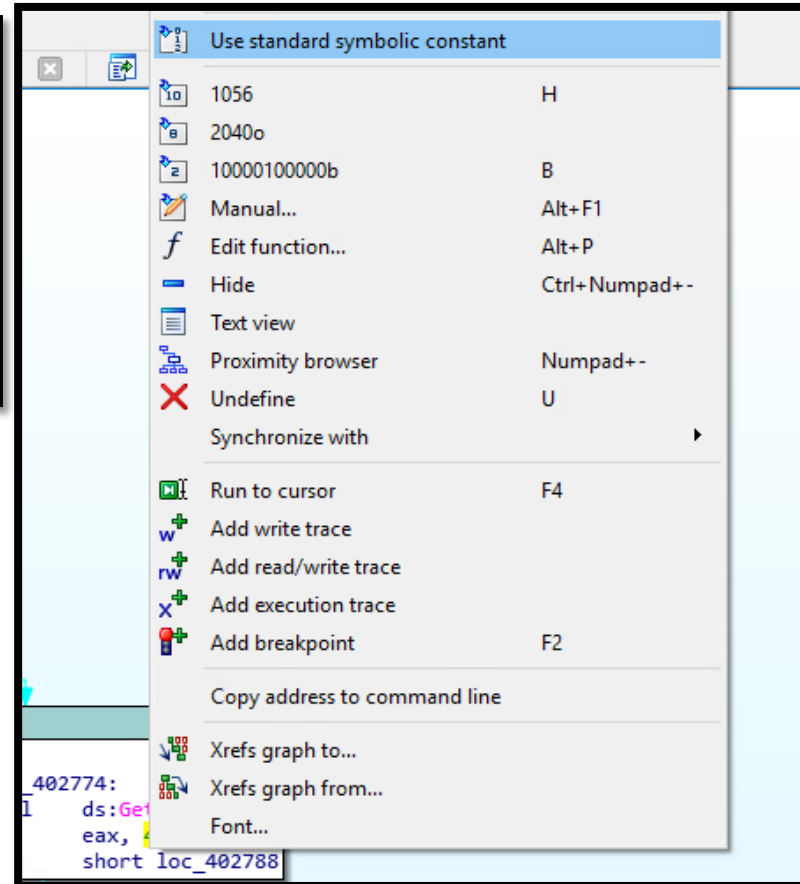
```
00000000004010C0
00000000004010C0 push    ecx
00000000004010C1 push    ebx
00000000004010C2 push    ebp
00000000004010C3 push    esi
00000000004010C4 xor     ebp, ebp
00000000004010C6 push    edi
00000000004010C7 push    ebp          ; dwFlags
00000000004010C8 push    ebp          ; lpszProxyBypass
00000000004010C9 mov     esi, ecx
00000000004010CB push    ebp          ; lpszProxy
00000000004010CC push    ebp          ; dwAccessType
00000000004010CD push    dword ptr [esi+1Ch] ; lpszAgent
00000000004010D0 call    ds:InternetOpenA
00000000004010D6 cmp     eax, ebp
00000000004010D8 mov     [esi+4], eax
00000000004010DB jz      loc_401162
```

Pro tip #4: Rename Symbolic Constants

- IDA's FLIRT signatures know the arguments for common APIs
- IDA also knows symbolic names for defined constants! Just tell IDA what to look for



```
0000000000402774  
0000000000402774 loc_402774:  
0000000000402774 call    ds:GetLastError  
000000000040277A cmp     eax, 420h  
000000000040277F jnz     short loc_402788
```



Pro tip #4: Rename Symbolic Constants

- IDA's FLIRT signatures know the arguments for common APIs
- IDA also knows symbolic names for defined constants! Just tell IDA what to look for

0000000000402774
0000000000402774 loc_402774:
0000000000402774 call ds:GetLastError
000000000040277A cmp eax, 420h
000000000040277F jnz short loc_402788

Use standard symbolic constant

1056	H
2040o	
10000100000b	B
Manual...	Alt+F1
Edit function...	Alt+P
Hide	Ctrl+Numpad+-
Text view	
Proximity browser	Numpad+-

ERROR_SERVICE_ALREADY_RUNNING

1056 (0x420)

An instance of the service is already running.

Symbol name	Value	Type library
CV_IA64_IntR32	00000420	MS SDK (Windows XP)
DISPID_EVMETH_ONOBJECTCONTENTSROLLED	00000420	MS SDK (Windows XP)
DISPID_IHTMLDOCUMENT2_OPEN	00000420	MS SDK (Windows XP)
DISPID_IHTMLLEVENTOBJ5_ISSESSION	00000420	MS SDK (Windows XP)
DISPID_ONOBJECTCONTENTSROLLED	00000420	MS SDK (Windows XP)
DISPID_WMPMETADATA_TEXT_DESCRIPTION	00000420	MS SDK (Windows XP)
ERROR_SERVICE_ALREADY_RUNNING	00000420	MS SDK (Windows XP)
MSG_FTS_JUMP_HASH	00000420	MS SDK (Windows XP)
SDP_ST_INT128	00000420	MS SDK (Windows XP)
TBM_SETBUDDY	00000420	MS SDK (Windows XP)
TB_SETBITMAPSIZE	00000420	MS SDK (Windows XP)
TTM_SETTITLE	00000420	MS SDK (Windows XP)

Pro tip #4: Rename Symbolic Constants

- IDA's FLIRT signatures know the arguments for common APIs
- IDA also knows symbolic names for defined constants! Just tell IDA what to look for

The screenshot illustrates the process of renaming a symbolic constant in IDA Pro. The main window displays assembly code at location `loc_402774`:

```
0000000000402774 loc_402774:  
0000000000402774 call ds:GetL  
000000000040277A cmp eax, 42  
000000000040277F jnz short 1
```

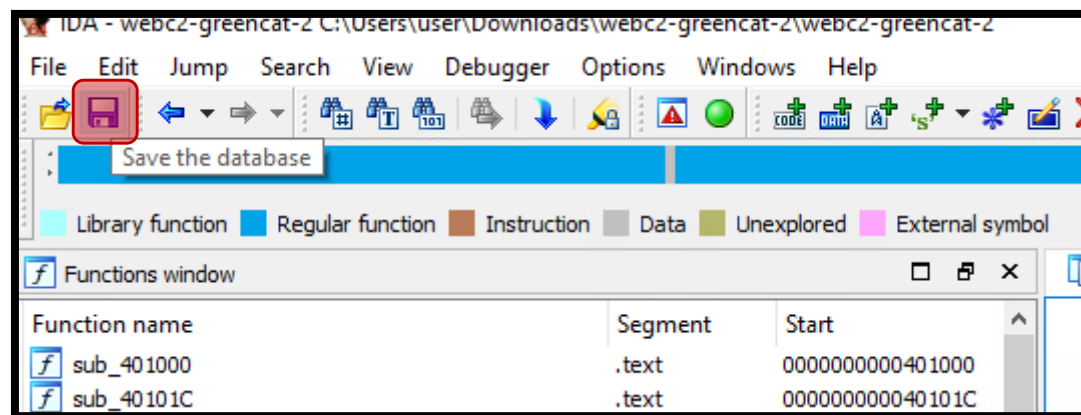
A dialog box titled "Use standard symbolic constant" is open, showing a list of constants. The constant `1056` is selected, which corresponds to the value `0x420` in the assembly code.

A separate window shows the constant `ERROR_SERVICE_ALREADY_RUNNING` with its value `1056 (0x420)`. Below this, a snippet of text reads: "An instance of the service is already running."

Another window shows the "Symbol name" list, where `ERROR_SERVICE_ALREADY_RUNNING` is highlighted. The list includes various constants and their values, such as `CV_IA64_IntR32`, `DISPID_EVMETH_ON`, `DISPID_IHTMLDOCUMENT2_OPEN`, `DISPID_IHTMLEVENTOBJ5_ISSESSION`, `DISPID_ONOBJECTCONTENTSROLLED`, `DISPID_WMPMETADATA_TEXT_DESCRIPTION`, `ERROR_SERVICE_ALREADY_RUNNING`, `MSG_FTS_JUMP_HASH`, `SDP_ST_INT128`, `TBM_SETBUDDY`, `TB_SETBITMAPSIZE`, and `TTM_SETTITLE`.

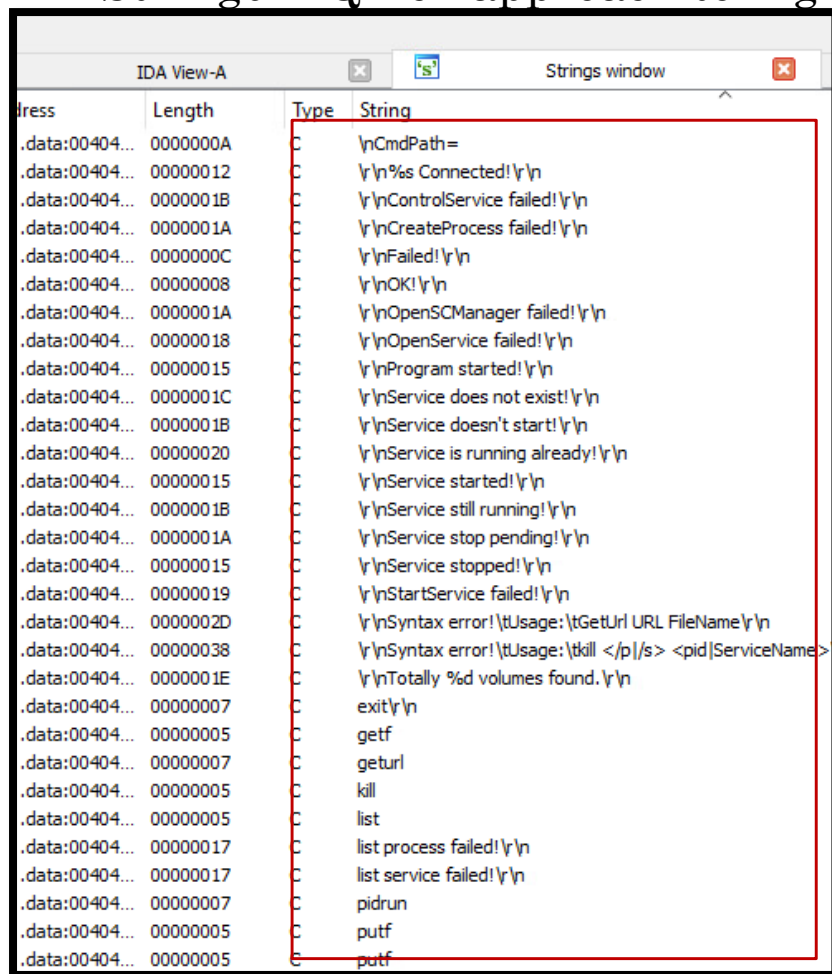
Pro tip #5: Save often!

- Losing hours of reverse engineering can be hazardous to your health!



Pro tip #6: Where to begin?

- Strings -- Quick approach to high-level overview (Shift-F12)

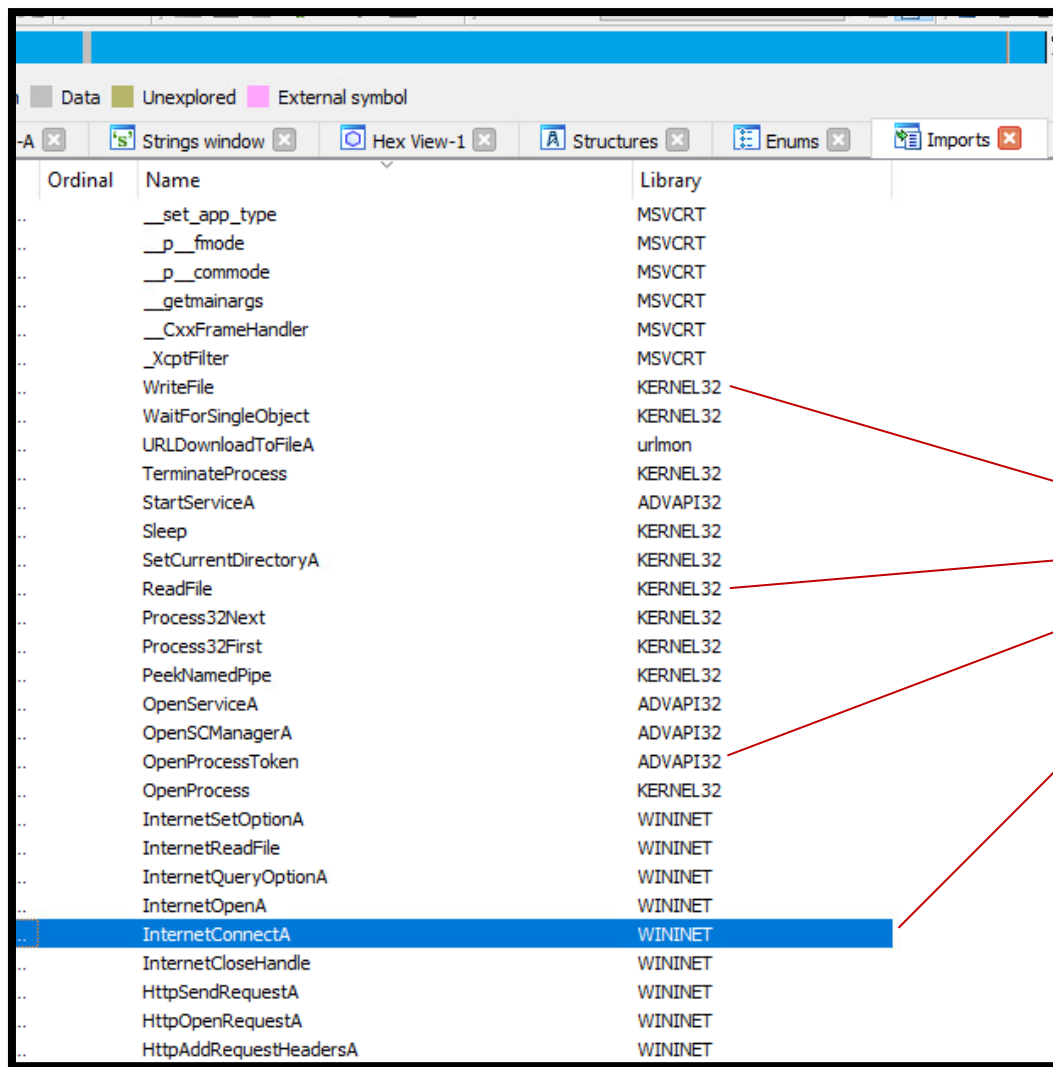


Address	Length	Type	String
.data:00404...	0000000A	C	\nCmdPath=
.data:00404...	00000012	C	\r\n%s Connected!\r\n
.data:00404...	0000001B	C	\r\nControlService failed!\r\n
.data:00404...	0000001A	C	\r\nCreateProcess failed!\r\n
.data:00404...	0000000C	C	\r\nFailed!\r\n
.data:00404...	00000008	C	\r\nOK!\r\n
.data:00404...	0000001A	C	\r\nOpenSCManager failed!\r\n
.data:00404...	00000018	C	\r\nOpenService failed!\r\n
.data:00404...	00000015	C	\r\nProgram started!\r\n
.data:00404...	0000001C	C	\r\nService does not exist!\r\n
.data:00404...	0000001B	C	\r\nService doesn't start!\r\n
.data:00404...	00000020	C	\r\nService is running already!\r\n
.data:00404...	00000015	C	\r\nService started!\r\n
.data:00404...	0000001B	C	\r\nService still running!\r\n
.data:00404...	0000001A	C	\r\nService stop pending!\r\n
.data:00404...	00000015	C	\r\nService stopped!\r\n
.data:00404...	00000019	C	\r\nStartService failed!\r\n
.data:00404...	0000002D	C	\r\nSyntax error!\tUsage:\tGetUrl URL FileName\r\n
.data:00404...	00000038	C	\r\nSyntax error!\tUsage:\tkill </p>/s> <pid>ServiceName>
.data:00404...	0000001E	C	\r\nTotally %d volumes found.\r\n
.data:00404...	00000007	C	exit\r\n
.data:00404...	00000005	C	getf
.data:00404...	00000007	C	geturl
.data:00404...	00000005	C	kill
.data:00404...	00000005	C	list
.data:00404...	00000017	C	list process failed!\r\n
.data:00404...	00000017	C	list service failed!\r\n
.data:00404...	00000007	C	pidrun
.data:00404...	00000005	C	putf
.data:00404...	00000005	C	putf



What are those????????!!!!

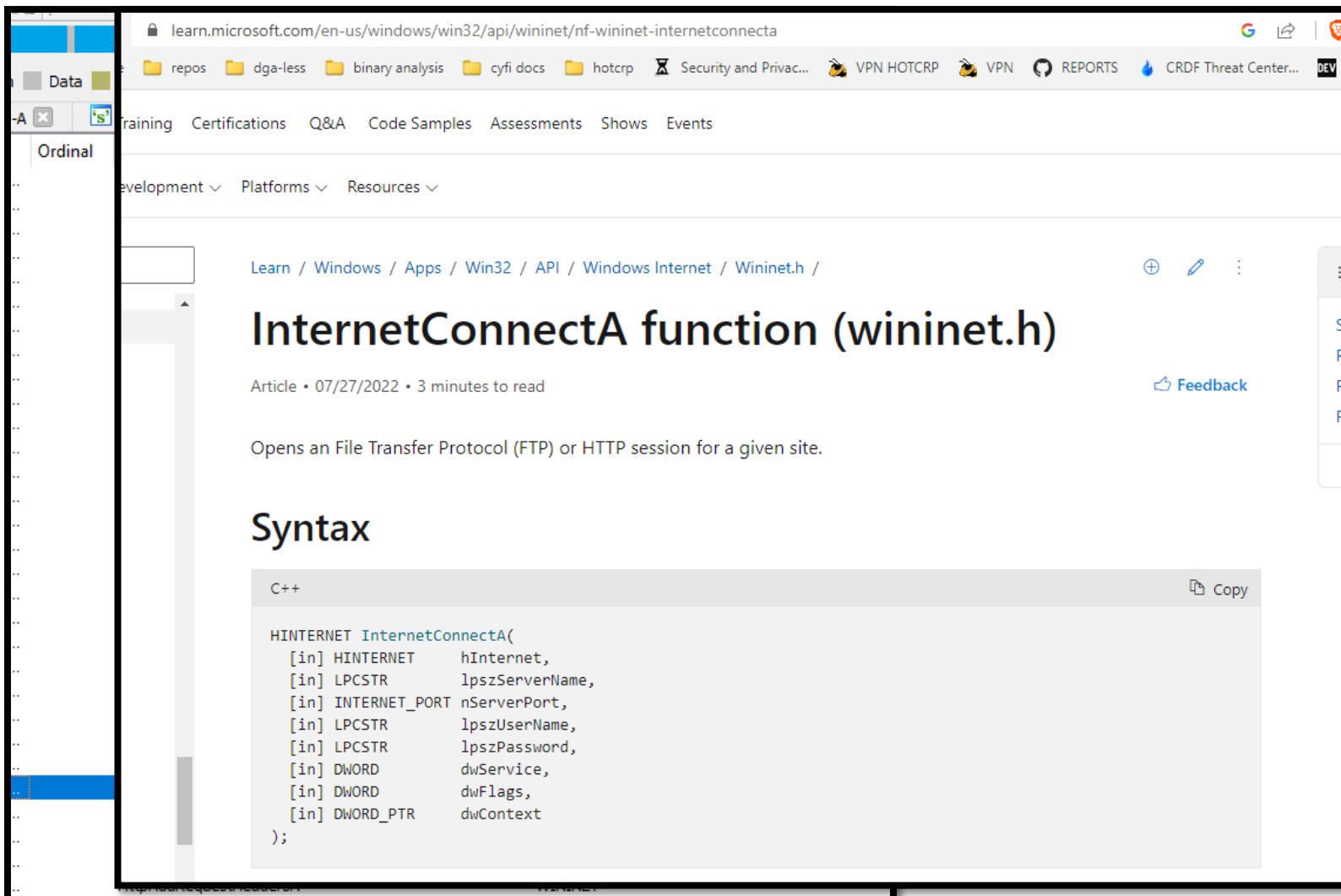
Pro tip #6: Where to begin? IAT



Ordinal	Name	Library
...	__set_app_type	MSVCRT
...	__p__fmode	MSVCRT
...	__p__commode	MSVCRT
...	__getmainargs	MSVCRT
...	__CxxFrameHandler	MSVCRT
...	_XcptFilter	MSVCRT
...	WriteFile	KERNEL32
...	WaitForSingleObject	KERNEL32
...	URLDownloadToFileA	urlmon
...	TerminateProcess	KERNEL32
...	StartServiceA	ADVAPI32
...	Sleep	KERNEL32
...	SetCurrentDirectoryA	KERNEL32
...	ReadFile	KERNEL32
...	Process32Next	KERNEL32
...	Process32First	KERNEL32
...	PeekNamedPipe	KERNEL32
...	OpenServiceA	ADVAPI32
...	OpenSCManagerA	ADVAPI32
...	OpenProcessToken	ADVAPI32
...	OpenProcess	KERNEL32
...	InternetSetOptionA	WININET
...	InternetReadFile	WININET
...	InternetQueryOptionA	WININET
...	InternetOpenA	WININET
...	InternetConnectA	WININET
...	InternetCloseHandle	WININET
...	HttpSendRequestA	WININET
...	HttpOpenRequestA	WININET
...	HttpAddRequestHeadersA	WININET

What do these do?

Pro tip #6: Where to begin? IAT

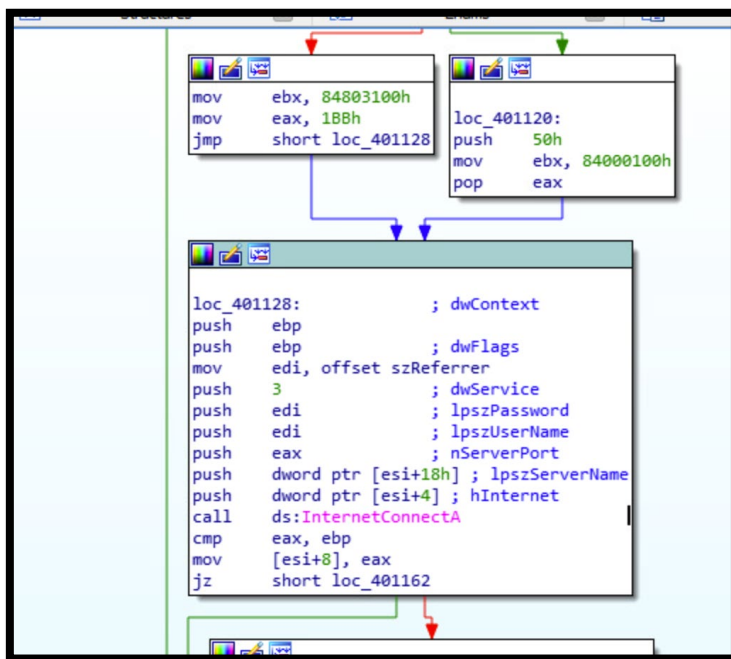
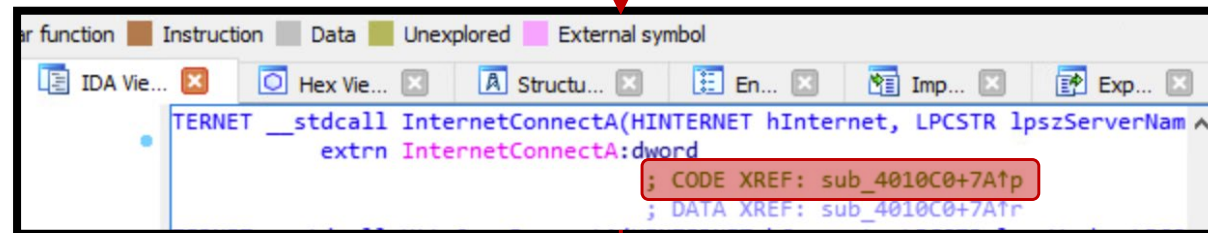
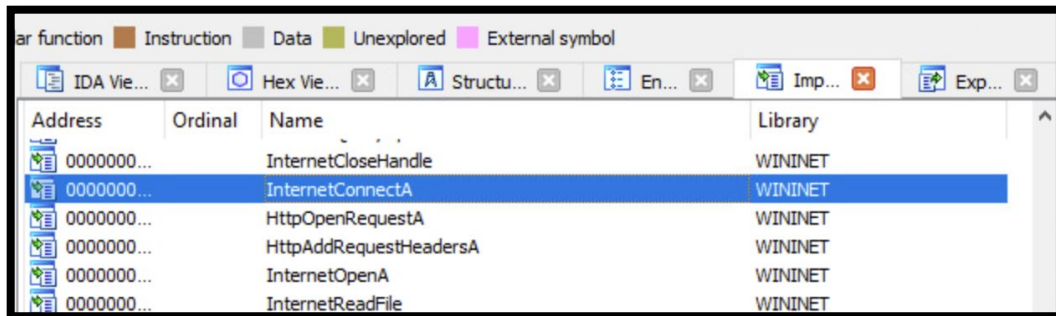


The screenshot shows a web browser displaying a Microsoft Learn article. The address bar shows the URL: `learn.microsoft.com/en-us/windows/win32/api/wininet/nf-wininet-internetconnecta`. The page title is "InternetConnectA function (wininet.h)". Below the title, it says "Article • 07/27/2022 • 3 minutes to read". The description states: "Opens an File Transfer Protocol (FTP) or HTTP session for a given site." The "Syntax" section contains a C++ code block with the following function signature:

```
C++  
HINTERNET InternetConnectA(  
    [in] HINTERNET    hInternet,  
    [in] LPCSTR        lpszServerName,  
    [in] INTERNET_PORT nServerPort,  
    [in] LPCSTR        lpszUserName,  
    [in] LPCSTR        lpszPassword,  
    [in] DWORD          dwService,  
    [in] DWORD          dwFlags,  
    [in] DWORD_PTR      dwContext  
);
```

se do?

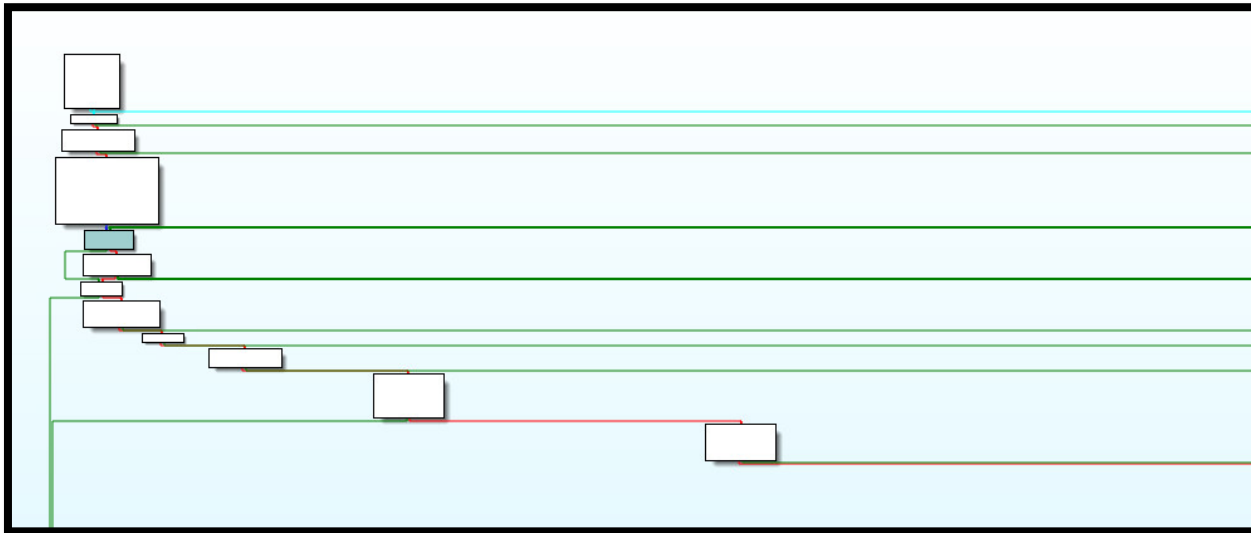
Pro tip #6: IAT Motivated Search



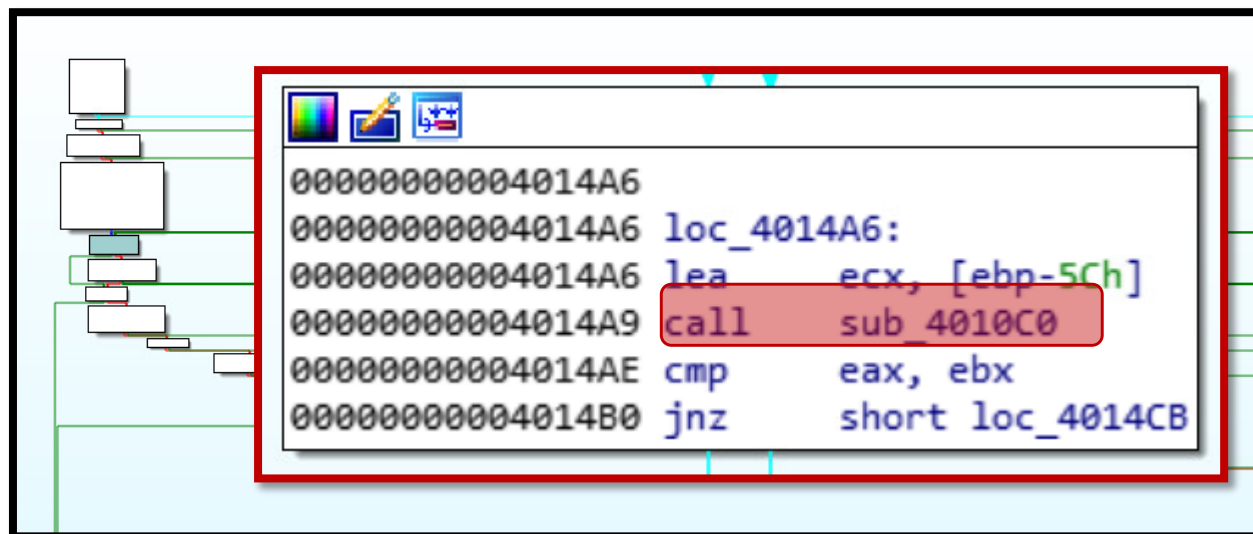
Recursive Descent RE?

- How should we go about REing malware?
- We can learn from recursive descent disassembly
- Follow a path to the end
 - Through switch/case statements
 - Through loops
 - Across multiple functions
- Create labels
 - Name function
 - Name variables
 - Comment where you're fairly sure

Let's try 0x401406

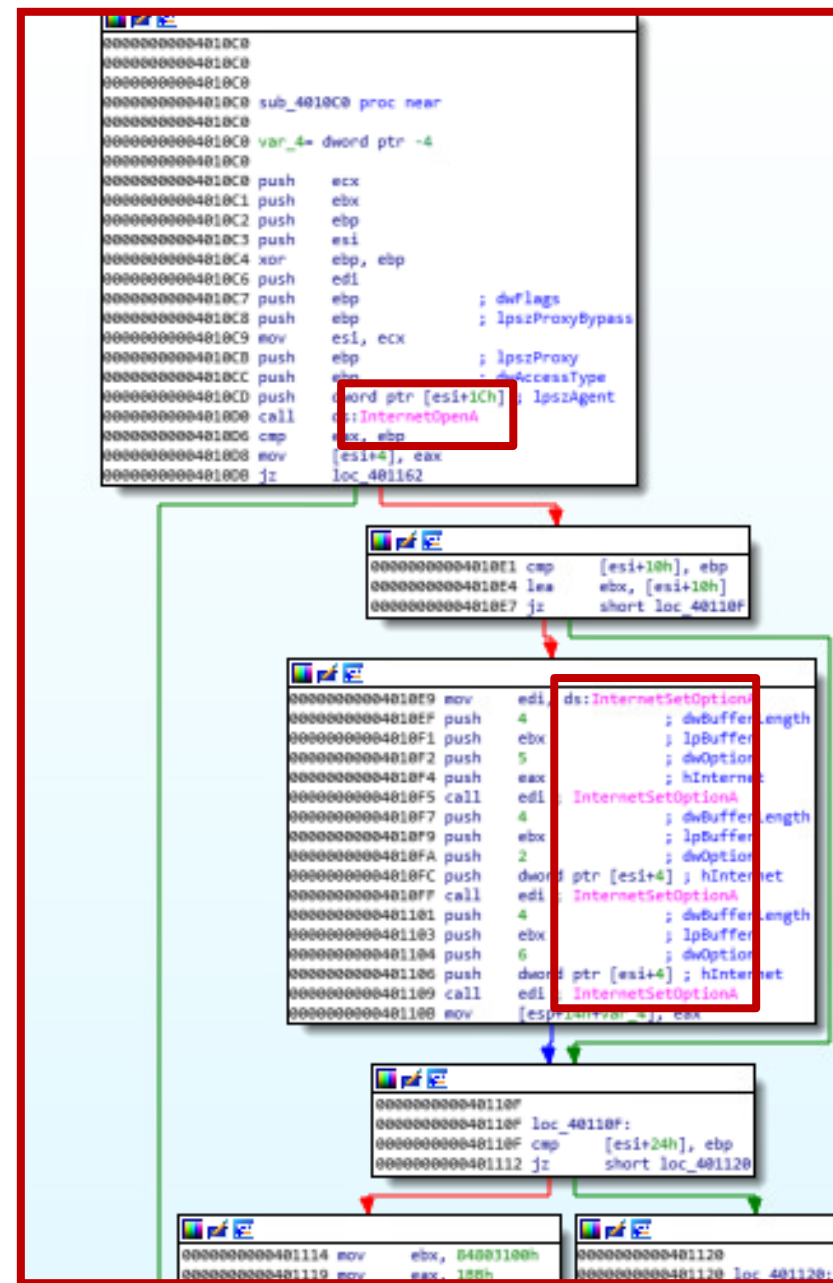
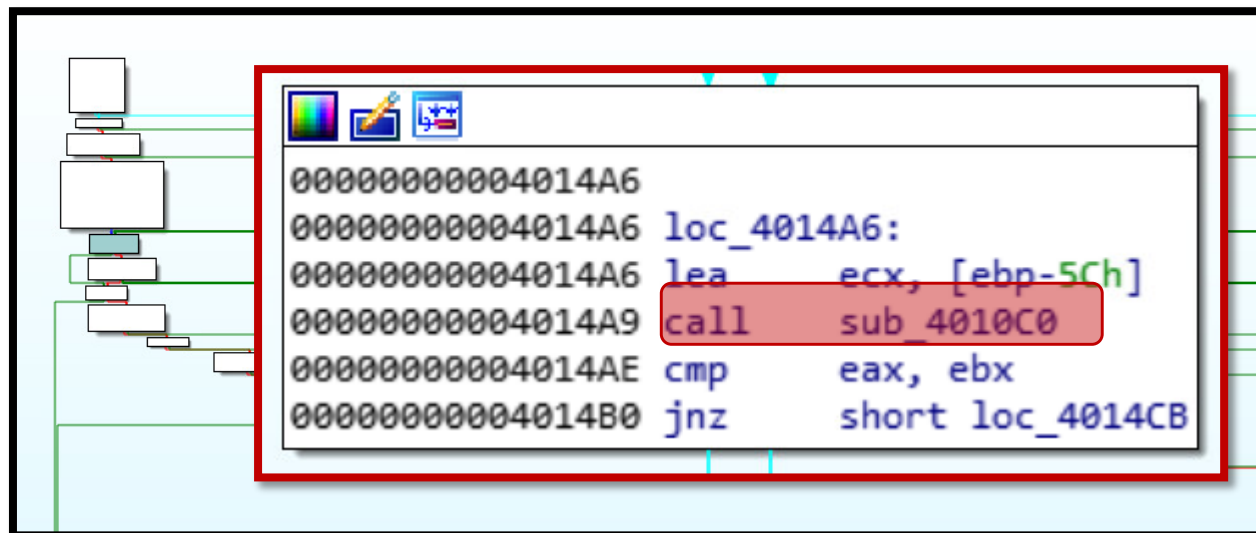


Let's try 0x401406



```
00000000004014A6  
00000000004014A6 loc_4014A6:  
00000000004014A6 lea     ecx, [ebp-5Ch]  
00000000004014A9 call    sub_4010C0  
00000000004014AE cmp     eax, ebx  
00000000004014B0 jnz     short loc_4014CB
```

Let's try 0x401406



Let's try 0x401406

internetopena

About 18,400 results (0.37 seconds)

<https://learn.microsoft.com/.../Wininet.h>

InternetOpenA function (wininet.h) - Win32 apps

Jul 27, 2022 — **InternetOpen** is the first WinINet function called by an application. It tells the

```
000000004010C0 sub_4010C0 proc near
000000004010C0
000000004010C0 var_4= dword ptr -4
000000004010C0
000000004010C0 push ecx
000000004010C1 push ebx
000000004010C2 push ebp
000000004010C3 push esi
000000004010C4 xor ebp, ebp
000000004010C6 push edi
000000004010C7 push ebp ; dwFlags
000000004010C8 push ebp ; lpzProxyBypass
000000004010C9 mov esi, ecx
000000004010CB push ebp ; lpzProxy
000000004010CC push ebp ; dwAccessType
000000004010CD push dword ptr [esi+1Ch] ; lpzAgent
```

```
00000000401101 push 4 ; dwBufferLength
00000000401103 push ebx ; lpBuffer
00000000401104 push 6 ; dwOption
00000000401106 push dword ptr [esi+4] ; hInternet
00000000401109 call edi ; InternetSetOptionA
0000000040110B mov [esp+14h+var_4], eax
```

```
0000000040110F loc_40110F:
0000000040110F cmp [esi+24h], ebp
00000000401112 jz short loc_401120
```

```
00000000401114 mov ebx, 04003100h
00000000401119 mov eax, 100h
```

```
00000000401120 loc_401120:
```

Let's try 0x401406

```
000000004010C0
000000004010C0
000000004010C0
000000004010C0 sub_4010C0 proc near
000000004010C0
000000004010C0 var_4= dword ptr -4
000000004010C0
000000004010C0 push ecx
000000004010C1 push ebx
000000004010C2 push ebp
000000004010C3 push esi
000000004010C4 xor ebp, ebp
000000004010C6 push edi
```

[Learn](#) / [Windows](#) / [Apps](#) / [Win32](#) / [API](#) / [Windows Internet](#) / [Wininet.h](#) /

InternetOpenA function (wininet.h)

Article • 07/27/2022 • 3 minutes to read

[Feedback](#)

Initializes an application's use of the WinINet functions.

Syntax

C++

[Copy](#)

```
HINTERNET InternetOpenA(
    [in] LPCSTR lpszAgent,
    [in] DWORD dwAccessType,
    [in] LPCSTR lpszProxy,
    [in] LPCSTR lpszProxyBypass,
    [in] DWORD dwFlags
);
```

ebp
[h]
110F

option
buffer length
buffer
option
Internet
[onA
buffer length
buffer
option
Internet
[onA
buffer length
buffer
option
Internet
[onA

ebp
01120

20
20 loc_401120

cdecl

Table 4-9: Conditional Jumps

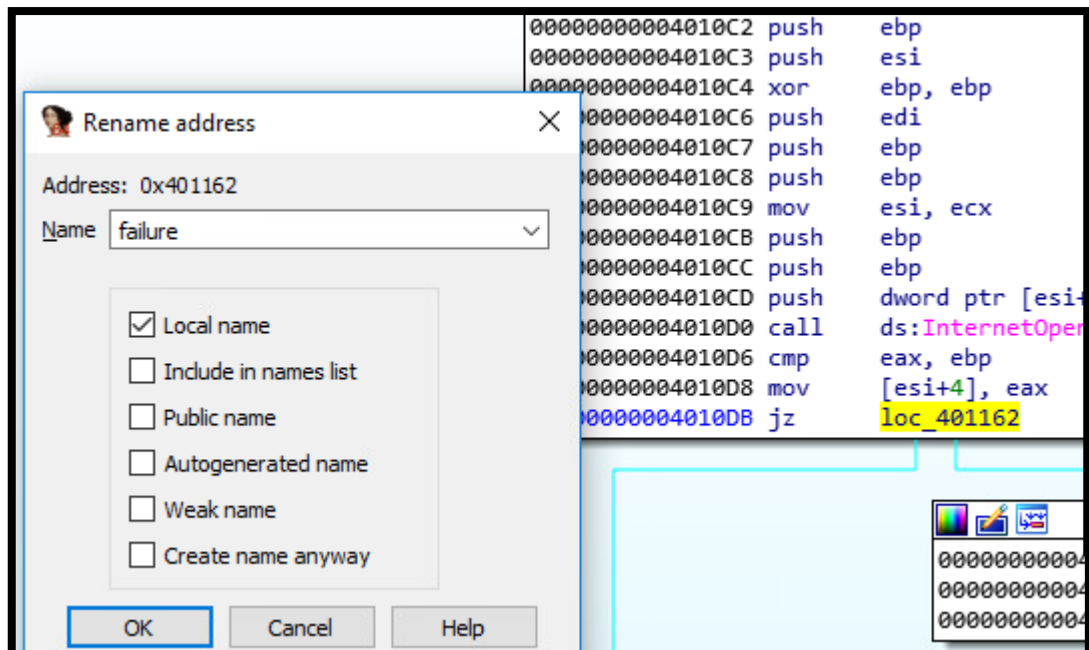
Instruction	Description
jz loc	Jump to specified location if ZF = 1.

```
00000000004010C0
00000000004010C0 push    ecx
00000000004010C1 push    ebx
00000000004010C2 push    ebp
00000000004010C3 push    esi
00000000004010C4 xor     ebp, ebp
00000000004010C6 push    edi
00000000004010C7 push    ebp           ; dwFlags
00000000004010C8 push    ebp           ; lpszProxyBypass
00000000004010C9 mov     esi, ecx
00000000004010CB push    ebp           ; lpszProxy
00000000004010CC push    ebp           ; dwAccessType
00000000004010CD push    dword ptr [esi+1Ch] ; lpszAgent
00000000004010D0 call    ds:InternetOpenA
00000000004010D6 cmp     eax, ebp
00000000004010D8 mov     [esi+4], eax
00000000004010DB jz      loc_401162
```

Table 4-8: cmp Instruction and Flags

cmp dst, src	ZF	CF
dst = src	1	0
dst < src	0	1
dst > src	0	0

Helpful naming...

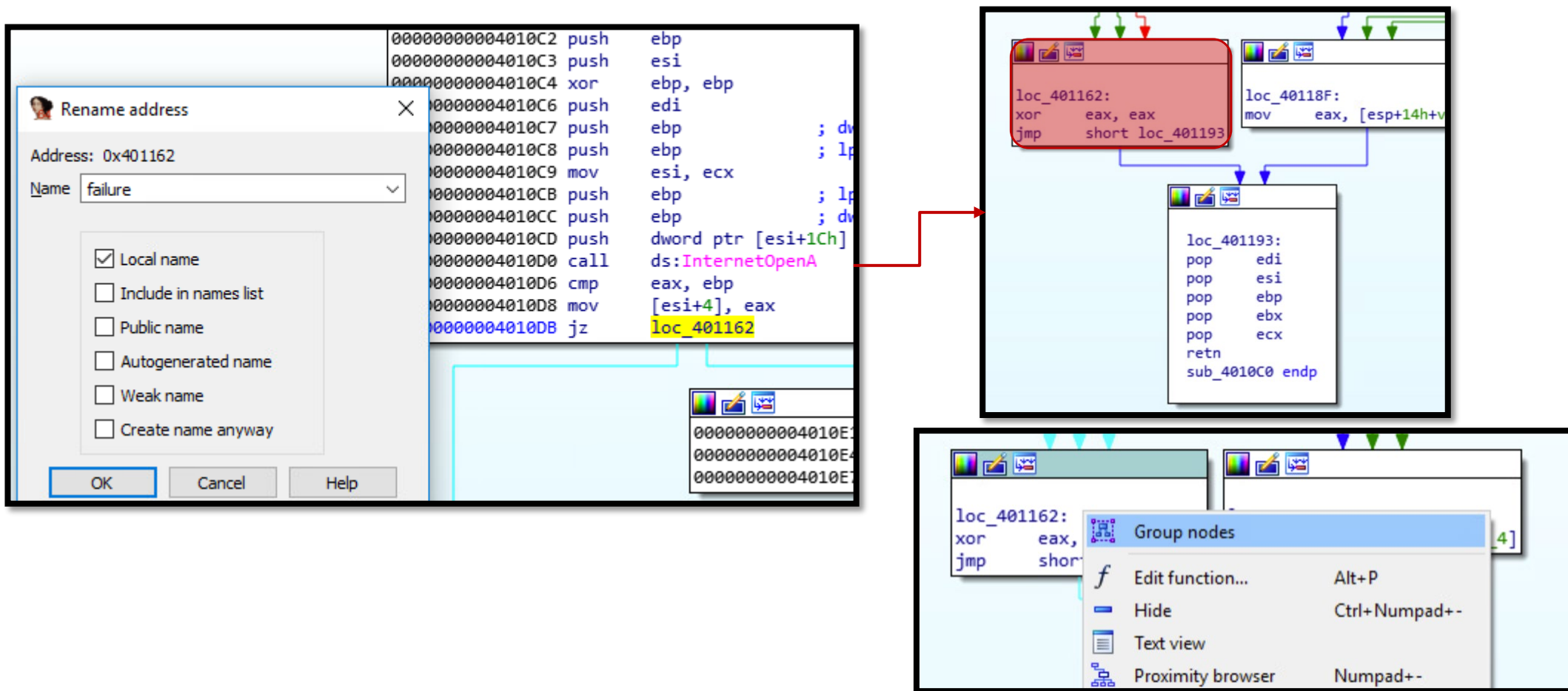


```
00000000004010C2 push ebp
00000000004010C3 push esi
00000000004010C4 xor ebp, ebp
00000000004010C6 push edi
00000000004010C7 push ebp ; dwFlags
00000000004010C8 push ebp ; lpszProxyBypass
00000000004010C9 mov esi, ecx
00000000004010CB push ebp ; lpszProxy
00000000004010CC push ebp ; dwAccessType
00000000004010CD push dword ptr [esi+1Ch] ; lpszAgent
00000000004010D0 call ds:InternetOpenA
00000000004010D6 cmp eax, ebp
00000000004010D8 mov [esi+4], eax
00000000004010DB jz loc_401162
```

```
0000000000401129 push ebp ; dwFlags
000000000040112A mov edi, offset szReferrer
000000000040112F push 3 ; dwService
0000000000401131 push edi ; lpszPassword
0000000000401132 push edi ; lpszUserName
0000000000401133 push eax ; nServerPort
0000000000401134 push dword ptr [esi+18h] ; lpszServerName
0000000000401137 push dword ptr [esi+4] ; hInternet
000000000040113A call ds:InternetConnectA
0000000000401140 cmp eax, ebp
0000000000401142 mov [esi+8], eax
0000000000401145 jz short failure
```

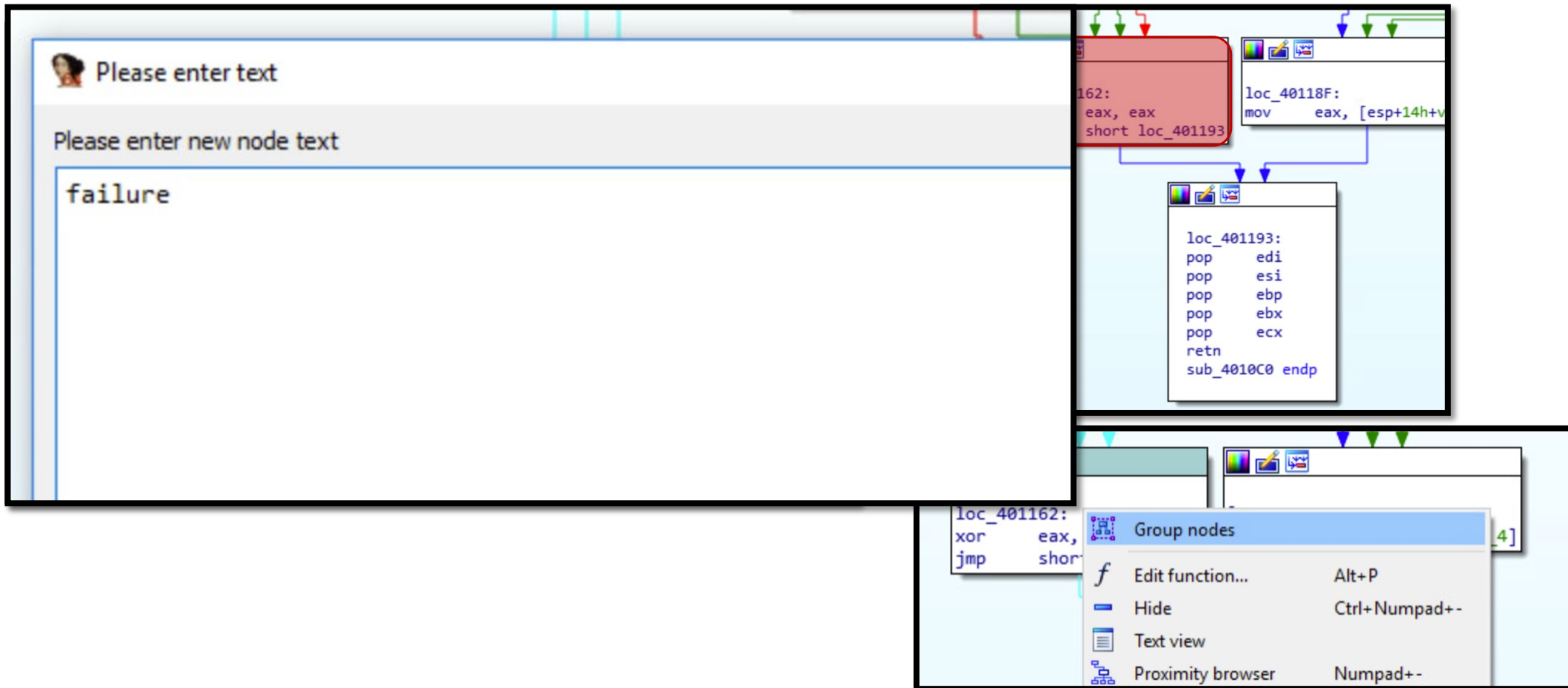
“Easy on the eyes”

- We can group basic blocks to make visual inspection more appealing



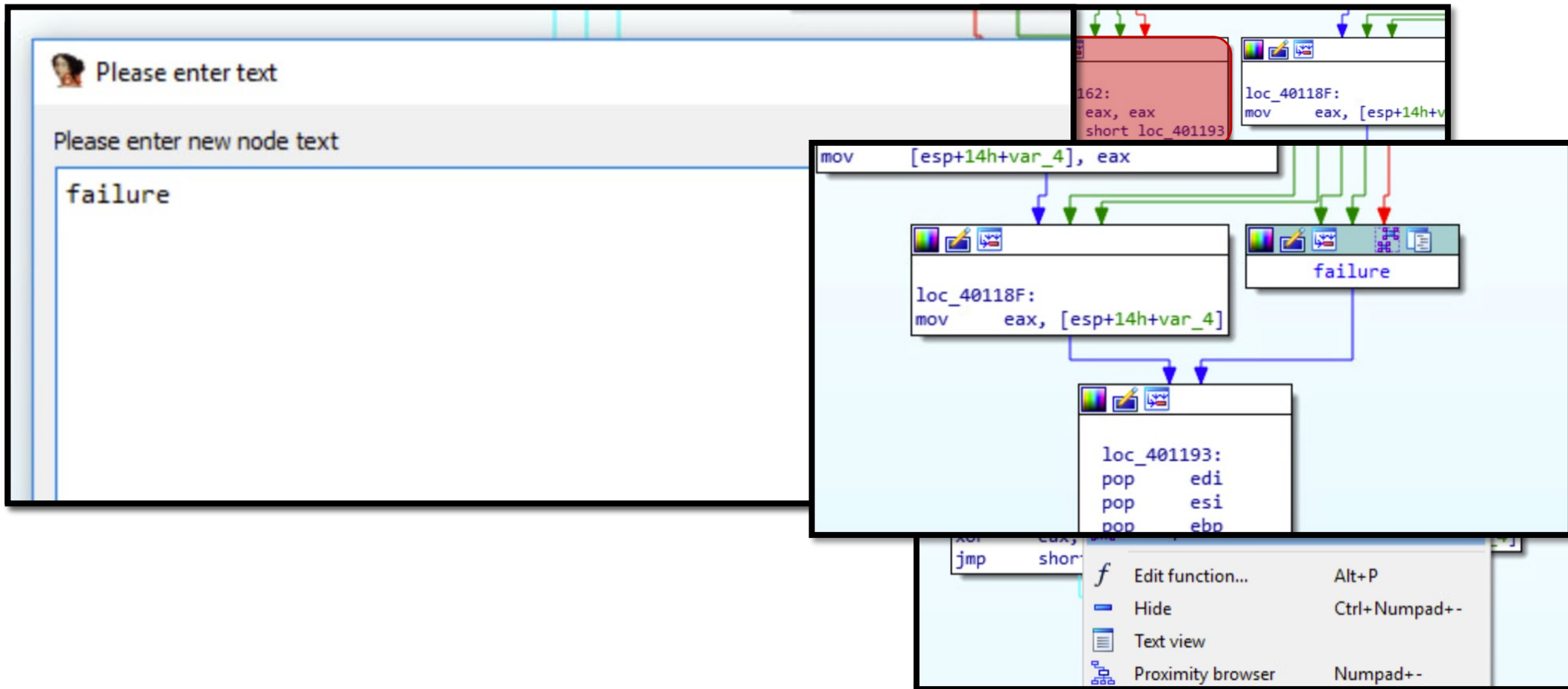
“Easy on the eyes”

- We can group basic blocks to make visual inspection more appealing



“Easy on the eyes”

- We can group basic blocks to make visual inspection more appealing



Unhelpful naming...

```
00000000004010C0 var_4= dword ptr -4
00000000004010C0
00000000004010C0 push    ecx
00000000004010C1 push    ebx
00000000004010C2 push    ebp
00000000004010C3 push    esi
00000000004010C4 xor     ebp, ebp
00000000004010C6 push    edi
00000000004010C7 push    ebp           ; dwFlags
00000000004010C8 push    ebp           ; lpszProxy
00000000004010C9 mov     esi, ecx
00000000004010CB push    ebp           ; lpszProxy
00000000004010CC push    ebp           ; dwAccessTy
00000000004010CD push    dword ptr [esi+1Ch] ; lpszA
00000000004010D0 call    ds:InternetOpenA
00000000004010D6 cmp     eax, ebp
00000000004010D8 mov     [esi+4], eax
00000000004010DB jz      failure_for_internetopen
```

```
0000000000401128
0000000000401128 loc_401128:           ; dwContext
0000000000401128 push    ebp
0000000000401129 push    ebp           ; dwFlags
000000000040112A mov     edi, offset szReferrer
000000000040112F push    3             ; dwService
0000000000401131 push    edi           ; lpszPassword
0000000000401132 push    edi           ; lpszUserName
0000000000401133 push    eax           ; nServerPort
0000000000401134 push    dword ptr [esi+18h] ; lpszServerName
0000000000401137 push    dword ptr [esi+4] ; hInternet
000000000040113A call    ds:InternetConnectA
0000000000401140 cmp     eax, ebp
0000000000401142 mov     [esi+8], eax
0000000000401145 jz      short failure_for_internetopen
```


Exercise #1: Press 'g', enter 4010c0

- As a team, tell me what this function is responsible for. I need a **high-level** overview. (I don't need "the prologue is prepared by.....") 🤔 Who read the book?
- When complete and your team agrees, rename the `sub_4010C0` function to a descriptive term

Hints:

1. Search for API documentation <https://learn.microsoft.com/en-us/windows/win32/api/>
2. Look at returns (cdecl) and how they are used
3. Symbolic constants (API arguments)
4. Ask questions!

Exercise #2: Go to 4013CC

- Super easy!
- Rename it



Exercise #3: Go to 402645

- As a team, tell me what this function is responsible for. I need a **high-level** overview.
- When complete and your team agrees
 1. rename the `sub_402645` function to a descriptive term
 2. Provide the sequence of APIs that lead to your derived capability
`InternetReadFile -> WriteFile -> ShellExecute = Execute Dropped File`

Hints:

1. Search for API documentation <https://learn.microsoft.com/en-us/windows/win32/api/>
2. Look at hard-coded strings for help
3. There are 3 main paths (1 of them is the failure path)

Summary

- Static analysis is “fun”
- Not very useful on sophisticated (obfuscated and packed malware)
- Useful to analyze benign software (when source code isn't available) to find bugs
- Useful to troubleshoot more sophisticated binary analysis tools (i.e., my tool breaks down at instruction X...why?)
- What IOCs can you extract from Greencat?
 - A sequence of APIs can be used to classify maliciousness
 - Sequence of APIs can also be used to identify capabilities (...like CAPA)

For Next Lesson

- Read Chapter 7 from *Practical Malware Analysis*

