Shrinking C++ Executables

by R4ndom on Sep.10, 2012, under Intermediate, Reverse Engineering, Tutorials

Over the course of creating these tutorials, I have been confronted with attempting to make the compiled binaries small. Usually, after entering a three line program in C++, Visual Studio will assume I would like every DLL, API and function ever created by Microsoft to be included in my binary, and I end up having something close to a 6 meg file. (Don't even get me started on the fact that you can open a new Word document, type one letter, and the file no longer fits on a 32Gig USB key!)

Because you don't want the binary filled with a bunch of useless crap to detract from the learning process, the binary should ONLY contain the instructions you want used, and nothing else. You would think this would be easy- perhaps a button somewhere that says "De-crapify" or something, but this is Microsoft, so you actually have to do quite a bit of experimenting in Visual Studio to get the binary size even close to what it should actually be.

Over the weekend I did some experimenting, attempting to get the binary as small as possible and trying to figure out what all this crap is that gets inserted into our binary, and this tutorial covers what I learned. A lot of this info was performed by Zer0Flag, so many thanks (and kudos) go out to him for his hard work. If you would rather have the PDF of this tutorial, you can download it on the tutorials page. Otherwise, read on…

First off, here is the source code. It simply opens a message box with a string in it, then closes the app:

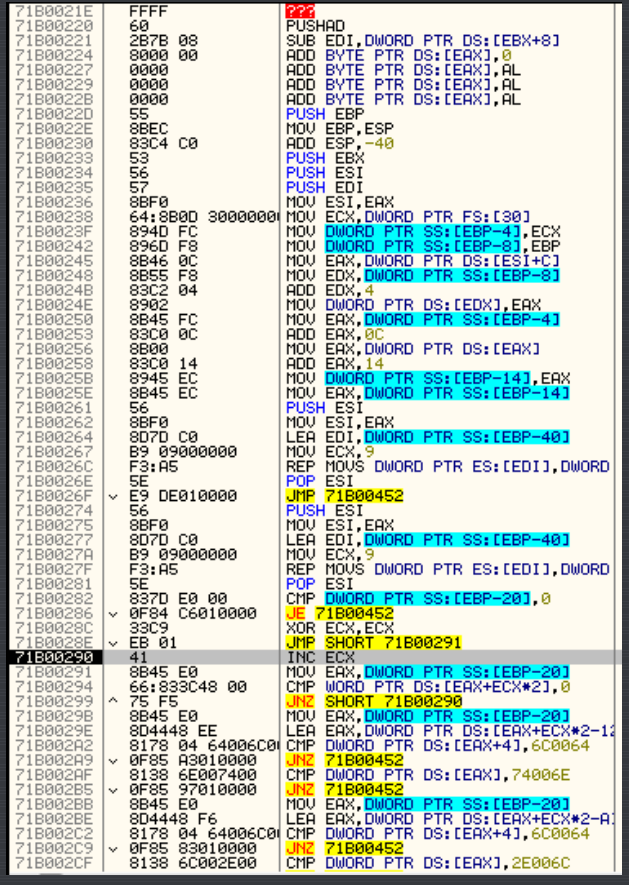
```
#include<Windows.h>

INT WinMain(HINSTANCE hInstance,
            HINSTANCE hPrevInstance,
            LPSTR lpCmdLine,
            int nCmdShow)
{
    MessageBoxA( NULL,
                 "Shrink- the incredible shrinking program!",
                 "Shrink!", MB_OK );
    return true;
}
```

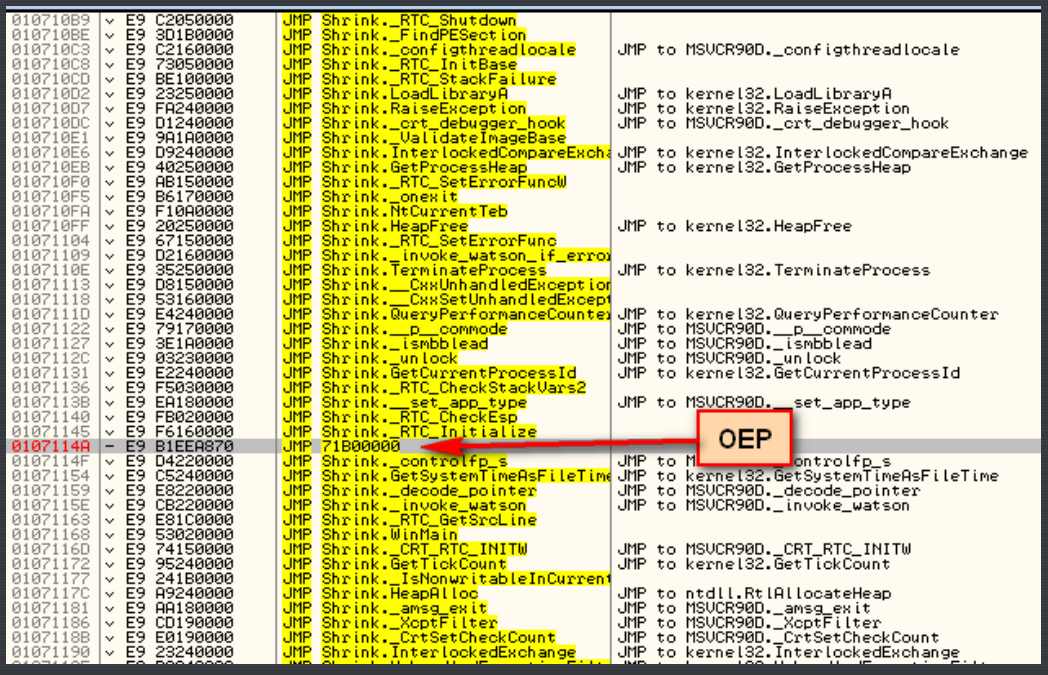
Because this is basically two lines of code, you would guess that it should be like, oh, maybe 50 bytes? You guessed wrong:

Shrink.exe	Application	30 KB	9/9/2012 9:30 AM
Shrink.ilk	Incremental Linke...	307 KB	9/9/2012 9:30 AM
Shrink.pdb	Program Debug D...	459 KB	9/9/2012 9:30 AM

Yes, try 30,000 bytes. For two lines of code? Oh, come on! Let's see what's going on in Olly:



This is where Olly first breaks, at address 71B00220. After a little digging, though, I found this is not the real EP. Looking in the PE header, the real entry point is at107114A:



This jumps to our initialization code. Interestingly, after we perform this code, the jump at 107114A will be dynamically changed to point to CRTMain later on. But for now, this jumps to the code in the picture above, starting at address 71B00220.

This initialization code looks for command line arguments and loads in DLLs for the application. At the end, we return to our original jump that is now changed to point to WinMainCRTStartup:

```
01071127  E9 3E1A0000   JMP Shrink._ismbblead
0107112C  E9 03230000   JMP Shrink._unlock
01071131  E9 E2240000   JMP Shrink.GetCurrentProcessId
01071136  E9 F5030000   JMP Shrink._RTC_CheckStackVars2
0107113B  E9 EA180000   JMP Shrink.__set_app_type
01071140  E9 FB020000   JMP Shrink._RTC_CheckEsp
01071145  E9 F6160000   JMP Shrink._RTC_Initialize
0107114A  E9 91060000   JMP Shrink.WinMainCRTStartup
0107114F  E9 D4220000   JMP Shrink._controlfp_s
01071154  E9 C5240000   JMP Shrink.GetSystemTimeAsFileTime
01071159  E9 E8220000   JMP Shrink._decode_pointer
0107115E  E9 CB220000   JMP Shrink._invoke_watson
01071163  E9 E81C0000   JMP Shrink._RTC_GetSrcLine
01071168  E9 53020000   JMP Shrink.WinMain
0107116D  E9 74150000   JMP Shrink._CRT_RTC_INITW
01071172  E9 95240000   JMP Shrink.GetTickCount
01071177  E9 241B0000   JMP Shrink._IsNonwritableInCurrent
```

CRTStartup is used for loading the C RunTime libraries. The CRT provides the fundamental C++ runtime support, including:

- setup the C++ exception model
- making sure the constructor of global variables get called before entering main function
- parse command line arguments, and call the main function
- initialize the heap
- setup the atexit chain

After the runtime is initialized, CRTStartup calls the __security_init_cookie function:

```
00CD1046  E9 CB170000   JMP Shrink.__setusermatherr     JMP to MSVCR90D.__setusermatherr
00CD104B  E9 6E250000   JMP Shrink.Sleep                JMP to kernel32.Sleep
00CD1050  E9 E1250000   JMP Shrink.GetModuleFileNameW   JMP to kernel32.GetModuleFileNameW
00CD1055  E9 E6190000   JMP Shrink.__security_init_cookie
00CD105A  E9 A1250000   JMP Shrink.SetUnhandledExceptionFilter  JMP to kernel32.SetUnhandledExceptionFilter
00CD105F  E9 FA1A0000   JMP Shrink._cexit               JMP to MSVCR90D._cexit
00CD1064  E9 8D1D0000   JMP Shrink._CrtDbgReportW       JMP to MSVCR90D._CrtDbgReportW
00CD1069  E9 CE250000   JMP Shrink.VirtualQuery         JMP to kernel32.VirtualQuery
00CD106E  E9 8D190000   JMP Shrink.atexit
```

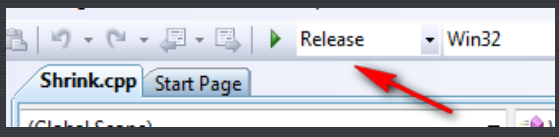
This function detects some buffer overruns that overwrite a function's return address, exception handler address, or certain types of parameters. Causing a buffer overrun is a technique used by hackers to exploit code that does not enforce buffer size restrictions.

After this function checks the code for potential buffer overruns, we finally get to our actual code:

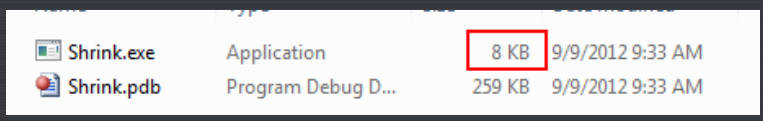
```
004013C0  .  55             PUSH EBP
004013C1  .  8BEC           MOV EBP,ESP
004013C3  .  81EC C0000000  SUB ESP,0C0
004013C9  .  53             PUSH EBX
004013CA  .  56             PUSH ESI
004013CB  .  57             PUSH EDI
004013CC  .  8DBD 40FFFFFF  LEA EDI,[LOCAL.48]
004013D2  .  B9 30000000    MOV ECX,30
004013D7  .  B8 CCCCCCCC    MOV EAX,CCCCCCCC
004013DC  .  F3:AB          REP STOS DWORD PTR ES:[EDI]
004013DE  .  8BF4           MOV ESI,ESP
004013E0  .  6A 00          PUSH 0                          ; Style = MB_OK|MB_APPLMODAL
004013E2  .  68 70574000    PUSH Shrink.00405770            ; Title = "Shrink!"
004013E7  .  68 3C574000    PUSH Shrink.0040573C            ; Text = "Shrink- the incredible shrinking program!"
004013EC  .  6A 00          PUSH 0                          ; hOwner = NULL
004013EE  .  FF15 50834000  CALL DWORD PTR DS:[<&USER32.MessageBoxA>]  ; MessageBoxA
004013F4  .  3BF4           CMP ESI,ESP
004013F6  .  E8 45FDFFFF    CALL Shrink.00401140
004013FB  .  B8 01000000    MOV EAX,1
00401400  .  5F             POP EDI                         ; Shrink.00401A88
00401401  .  5E             POP ESI                         ; Shrink.00401A88
00401402  .  5B             POP EBX                         ; Shrink.00401A88
00401403  .  81C4 C0000000  ADD ESP,0C0
00401409  .  3BEC           CMP EBP,ESP
0040140B  .  E8 30FDFFFF    CALL Shrink.00401140
00401410  .  8BE5           MOV ESP,EBP
00401412  .  5D             POP EBP                         ; Shrink.00401A88
00401413  .  C2 1000        RETN 10
00401416     CC             INT3
```

Changing the Build

The first thing we should notice is that Visual Studio defaults to debug mode, so we should definitely change to Release:



Now when we check the size, we see already a big difference:



Wow, that debug information was almost 75% of the binary's size! Loading this in CFF Explorer, we see that we lost the .textbss and .idata sections, and the other sections have been reduced drastically.

Debug version:

Name	Virtual Size	Virtual Address	Raw Size	Raw Address	Reloc Address
Byte[8]	Dword	Dword	Dword	Dword	Dword
.textbss	00010000	00001000	00000000	00000000	00000000
.text	000035BC	00011000	00003600	00000400	00000000
.rdata	00001CB2	00015000	00001E00	00003A00	00000000
.data	0000059C	00017000	00000200	00005800	00000000
.idata	000008F8	00018000	00000A00	00005A00	00000000
.rsrc	00000C09	00019000	00000E00	00006400	00000000
.reloc	00000458	0001A000	00000600	00007200	00000000

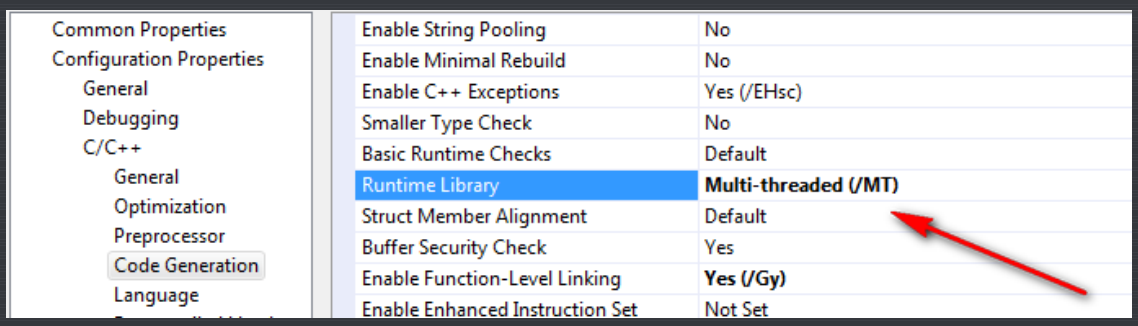
Release version:

Name	Virtual Size	Virtual Address	Raw Size	Raw Address	Reloc Address
Byte[8]	Dword	Dword	Dword	Dword	Dword
.text	0000087E	00001000	00000A00	00000400	00000000
.rdata	0000065E	00002000	00000800	00000E00	00000000
.data	00000388	00003000	00000200	00001600	00000000
.rsrc	000002B0	00004000	00000400	00001800	00000000
.reloc	00000192	00005000	00000200	00001C00	00000000

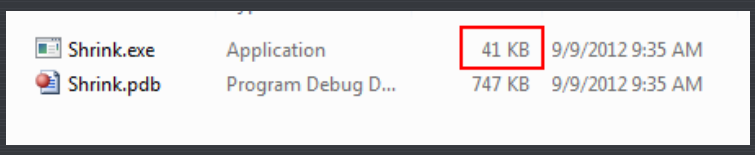
Removing the C Runtime

Of course, 8,000 bytes is already pretty good, but who wants to stop at “pretty good”?

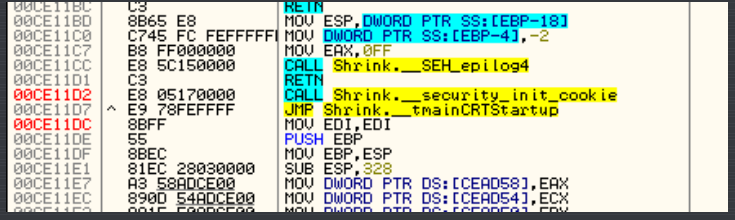
Next, we have to take a step backward in order to take a couple steps forward. Right-clicking the main project’s name in the Project Explorer and selecting Properties, we have the main properties window. Open the C/C++ tree and select the “Code Generation” item. We want to change the “Runtime Library” to “Multi-threaded (/MT)”. This will make the binary load the C++ runtime files when the executable is loaded. The reason we want to do this is so we can manually delete it later.



Changing this adds a significant amount back in, but will allow us to delete it (and more) later:



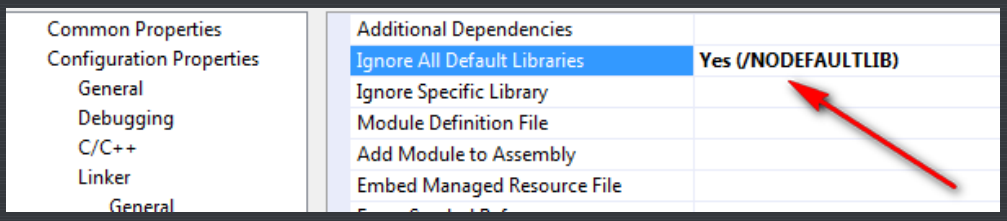
One thing you will notice in OllyDBG is that our jump table has all but disappeared:



This is because our DLLs have been inserted into our binary, so they will be called directly.

Ignore Default Libraries

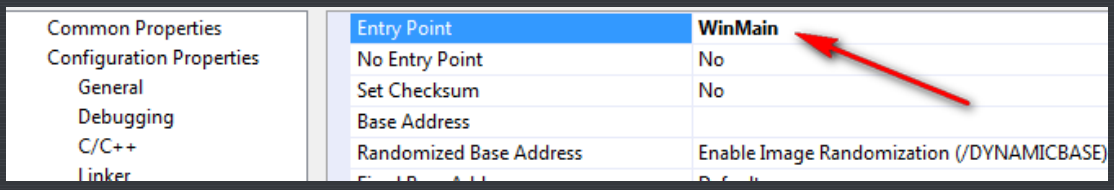
Clicking on the “Input” label under the “Linker” tree, we can force Visual Studio to ignore all the default libraries usually automatically loaded.



Changing this to “Yes” and trying to build the program gives us an error though:

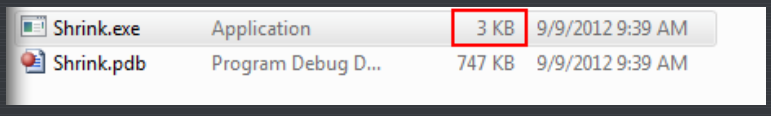
```
1>Linking...
1>Shrink.obj : error LNK2001: unresolved external symbol @__security_check_cookie@4
1>LINK : error LNK2001: unresolved external symbol _WinMainCRTStartup
1>C:\Users\Random\Documents\Visual Studio 2008\Projects\Shrink\Release\Shrink.exe : fatal error LNK1120: 2 unresolved externals
1>Build log was saved at "file://c:\Users\Random\Documents\Visual Studio 2008\Projects\Shrink\Shrink\Release\BuildLog.htm"
1>Shrink - 3 error(s), 0 warning(s)
========== Build: 0 succeeded, 1 failed, 0 up-to-date, 0 skipped ==========
```

To fix this we must change the entry point of our program. The reason for this is that Visual Studio incorporates several function calls before our program actually starts, namely the CRTStartup and security_cookie calls. That means the entry point is set to these functions instead of the true beginning of our app. Since we just told Visual Studio to ignore these functions, if we don’t change the entry point it is still pointing to these functions, that are now being ignored. Clicking on the “Advanced” label under “Linker” we can change this to our actual entry point, WinMain:



**** You may also need to change the “Buffer Security Check” option to “No (/GS-)” under C/C++ in the Code Generation tab to make it build properly. ****

Now when we build it we get no errors and also a file size of 3,000 bytes:



Now we’re talking! Loading this in Olly, we start to see some improvements:

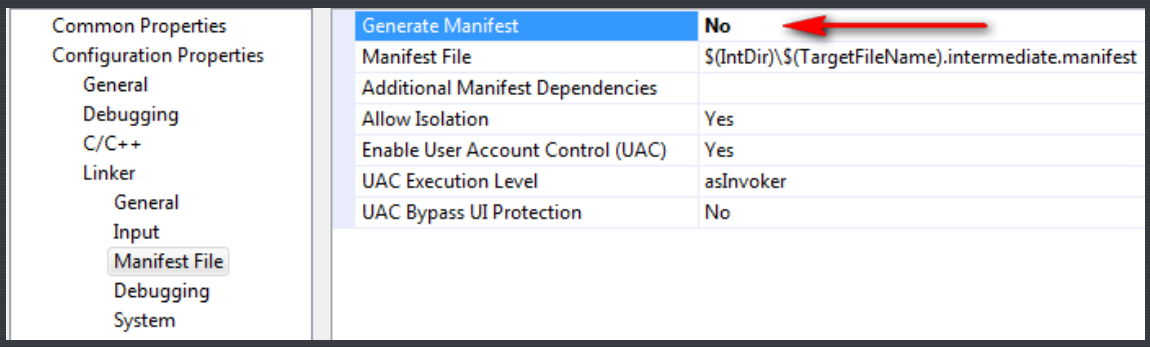
```
011A1000  - E9 FBEF9570    JMP 71B00000
011A1005    1A01           SBB AL,BYTE PTR DS:[ECX]
011A1007    68 34201A01    PUSH Shrink.011A2034                ASCII "Shrink- the incredible shrinking program!"
011A100C    6A 00          PUSH 0
011A100E    FF15 00201A01  CALL DWORD PTR DS:[<&USER32.MessageBoxA>]  USER32.MessageBoxA
011A1014    B8 01000000    MOV EAX,1
011A1019    C2 1000        RETN 10
011A101C    0000           ADD BYTE PTR DS:[EAX],AL
011A101E    0000           ADD BYTE PTR DS:[EAX],AL
011A1020    0000           ADD BYTE PTR DS:[EAX],AL
011A1022    0000           ADD BYTE PTR DS:[EAX],AL
011A1024    0000           ADD BYTE PTR DS:[EAX],AL
011A1026    0000           ADD BYTE PTR DS:[EAX],AL
011A1028    0000           ADD BYTE PTR DS:[EAX],AL
011A102A    0000           ADD BYTE PTR DS:[EAX],AL
011A102C    0000           ADD BYTE PTR DS:[EAX],AL
011A102E    0000           ADD BYTE PTR DS:[EAX],AL
011A1030    0000           ADD BYTE PTR DS:[EAX],AL
011A1032    0000           ADD BYTE PTR DS:[EAX],AL
011A1034    0000           ADD BYTE PTR DS:[EAX],AL
```

The setup code has also shrunk:

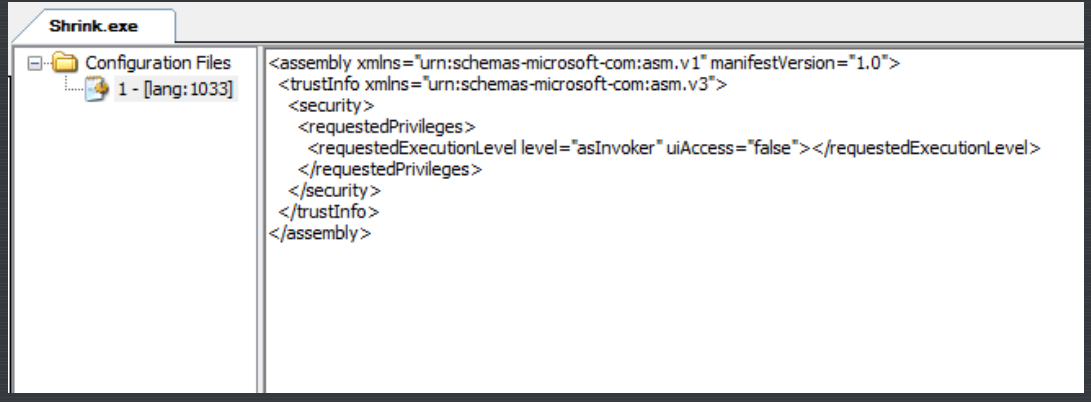
```
71B00000    B8 0000B071    MOV EAX,71B00000
71B00005    B9 2D02B071    MOV ECX,71B0022D
71B0000A    FFD1           CALL ECX
71B0000C    0010           ADD BYTE PTR DS:[EAX],DL
71B0000E    1A01           SBB AL,BYTE PTR DS:[ECX]
71B00010    6A 00          PUSH 0
71B00012    68 2C203E00    PUSH 3E202C
71B00017    40             INC EAX
71B00018    001D 00B07143  ADD BYTE PTR DS:[4371B000],BL
71B0001E    003A           ADD BYTE PTR DS:[EDX],BH
71B00020    005C00 57      ADD BYTE PTR DS:[EAX+EAX+57],BL
71B00024    0069 00        ADD BYTE PTR DS:[ECX],CH
71B00027    6E             OUTS DX,BYTE PTR ES:[EDI]           I/O command
71B00028    006400 6F      ADD BYTE PTR DS:[EAX+EAX+6F],AH
71B0002C    0077 00        ADD BYTE PTR DS:[EDI],DH
71B0002F  v 73 00          JNB SHORT 71B00031
71B00031    5C             POP ESP                             71B0000C
71B00032    0053 00        ADD BYTE PTR DS:[EBX],DL
71B00035  v 79 00          JNS SHORT 71B00037
71B00037  v 73 00          JNB SHORT 71B00039
71B00039    57             PUSH EDI
71B0003A    004F 00        ADD BYTE PTR DS:[EDI],CL
71B0003D    57             PUSH EDI
71B0003E    0036           ADD BYTE PTR DS:[ESI],DH
71B00040    003400         ADD BYTE PTR DS:[EAX+EAX],DH
71B00043    5C             POP ESP                             71B0000C
71B00044    0067 00        ADD BYTE PTR DS:[EDI],AH
71B00047  v 75 00          JNZ SHORT 71B00049
71B00049    61             POPAD
71B0004A    0072 00        ADD BYTE PTR DS:[EDX],DH
71B0004D    64:0033        ADD BYTE PTR FS:[EBX],DH
71B00050    0032           ADD BYTE PTR DS:[EDX],DH
71B00052    002E           ADD BYTE PTR DS:[ESI],CH
71B00054    006400 6C      ADD BYTE PTR DS:[EAX+EAX+6C],AH
71B00058    006C00 00      ADD BYTE PTR DS:[EAX+EAX],CH
71B0005C    0000           ADD BYTE PTR DS:[EAX],AL
71B0005E    0000           ADD BYTE PTR DS:[EAX],AL
71B00060    0000           ADD BYTE PTR DS:[EAX],AL
71B00062    0000           ADD BYTE PTR DS:[EAX],AL
71B00064    0000           ADD BYTE PTR DS:[EAX],AL
71B00066    0000           ADD BYTE PTR DS:[EAX],AL
71B00068    0000           ADD BYTE PTR DS:[EAX],AL
71B0006A    0000           ADD BYTE PTR DS:[EAX],AL
71B0006C    0000           ADD BYTE PTR DS:[EAX],AL
```

Removing the Manifest

Next we want to ditch the manifest as it's never used (at least not in our case). Under Linker, click Manifest File and change "Generate Manifest" to "No":



Doing this only saves about 200 bytes, but hey, that's something 🙂 Here we can see exactly what the manifest looks like (in CFF Explorer):



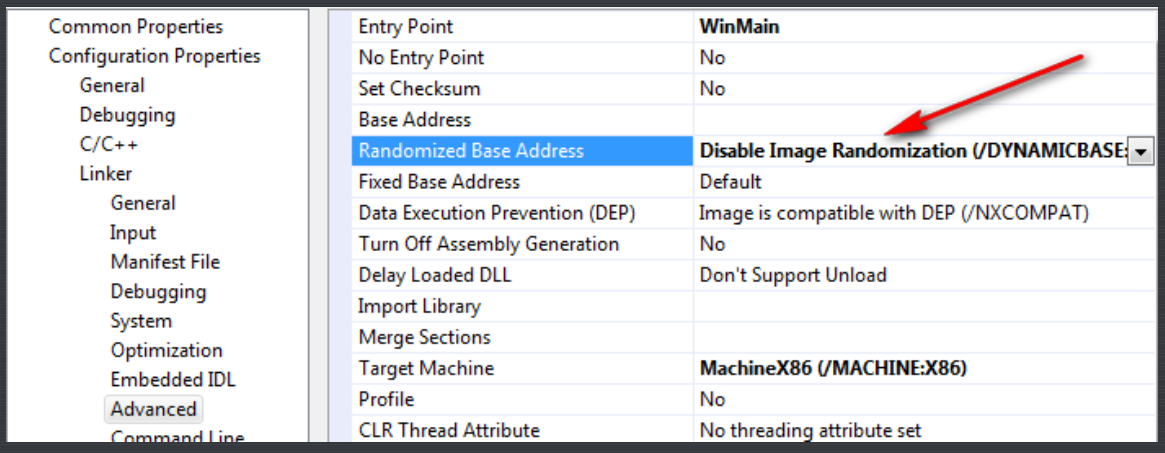
The next thing we may notice is that our binary has four sections:

Name	Virtual Size	Virtual Address	Raw Size	Raw Address	Reloc Address
Byte[8]	Dword	Dword	Dword	Dword	Dword
.text	0000001C	00001000	00000200	00000400	00000000
.rdata	00000114	00002000	00000200	00000600	00000000
.rsrc	000001B4	00003000	00000200	00000800	00000000
.reloc	0000001C	00004000	00000200	00000A00	00000000

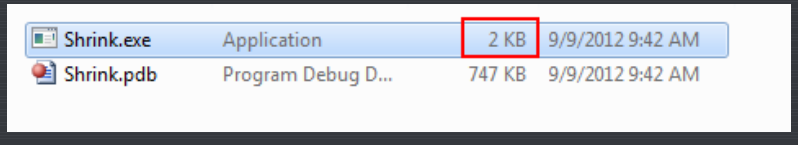
One that we could potentially lose is the .reloc section…

Removing Randomized Base Addresses

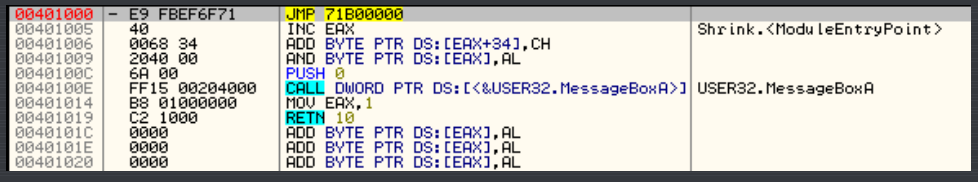
We don't need a relocations section if we never relocate code, so let's turn random relocations off:



Doing that and rebuilding automatically removes our .reloc section, shaving off another 1,000 bytes:

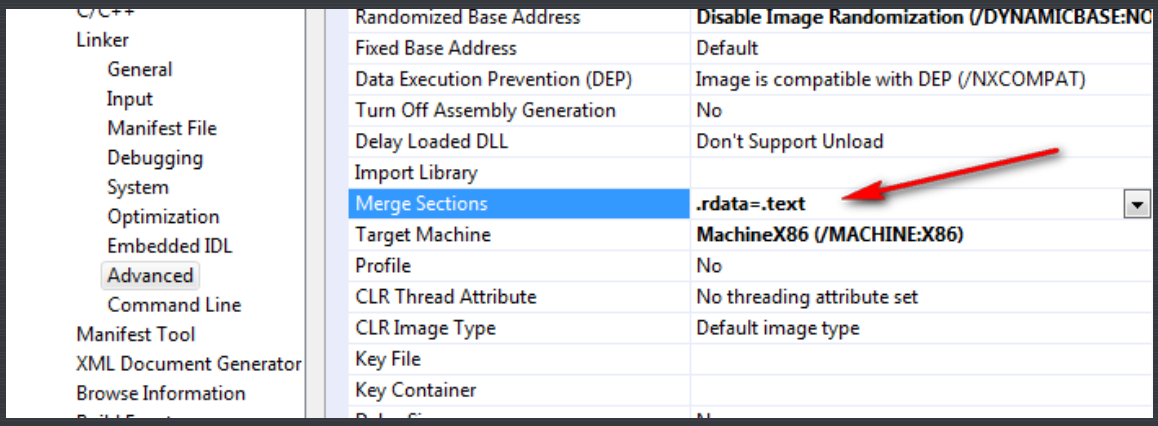


This also has the nice quality of loading our binary in at the usual address of 401000:

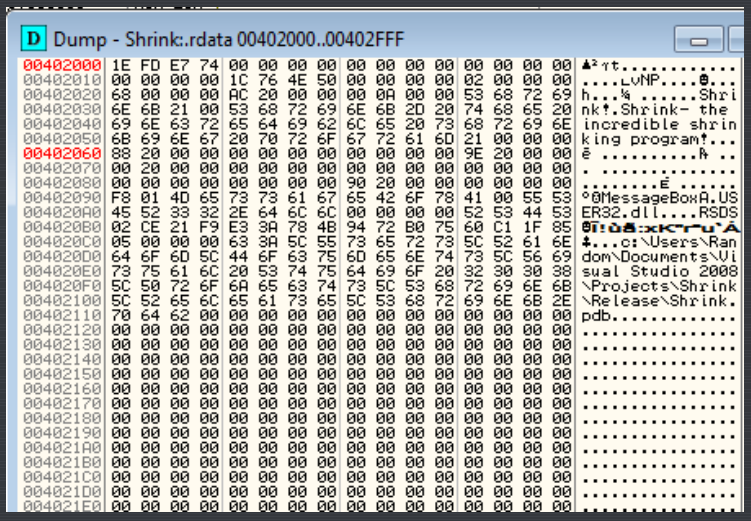


Combining Sections

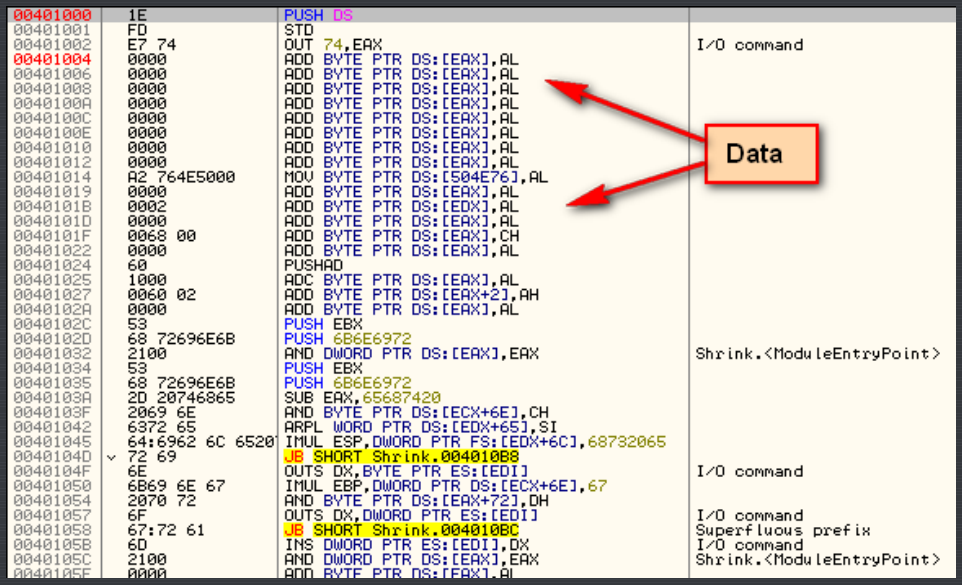
Next, we don't necessarily need both sections, especially since the second section only needs a couple dozen bytes but takes up 1,000. Here, we can set the .rdata section to share the .text section by merging them. Still in the Advanced tab, enter this for "Merge Sections":



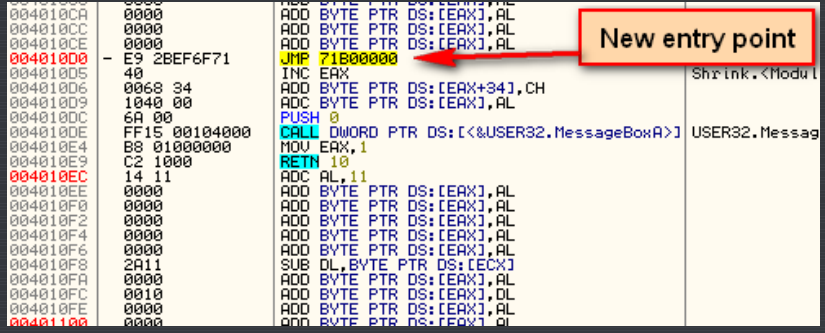
Here is our original .rdata section:



and after combining sections, we can see that this data was inserted into the beginning of our .text section in the binary:

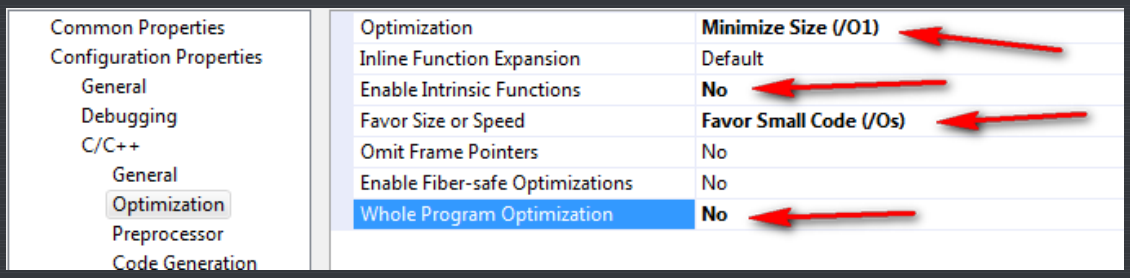


and that our entry point has been changed to 4010D0:

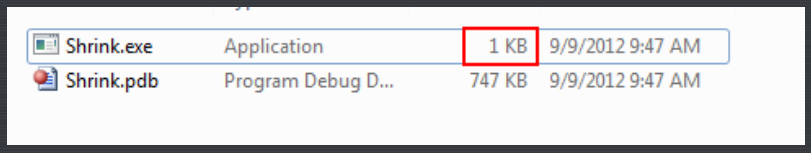


Changing Optimizations

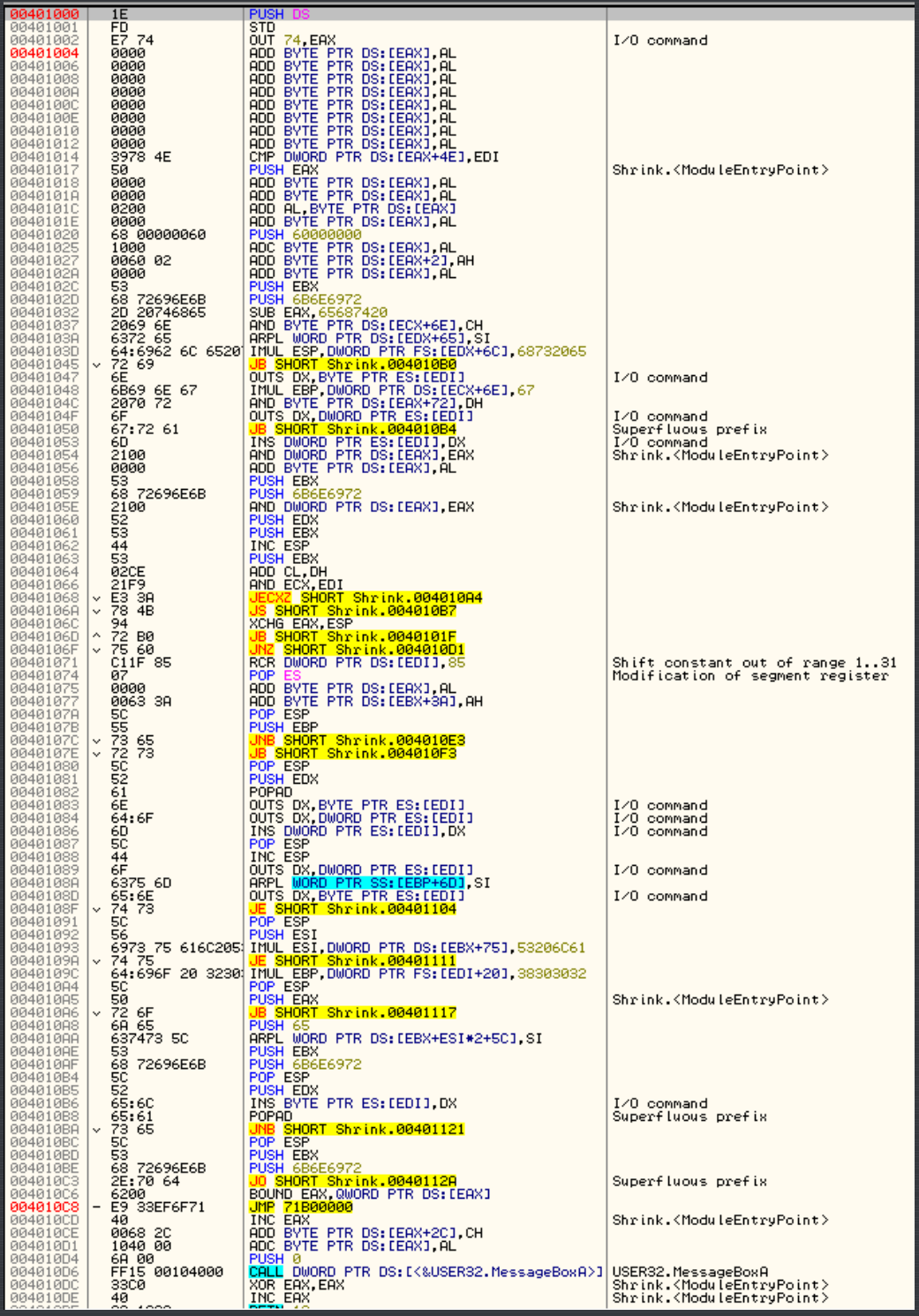
Lastly, we can change the optimizations that Visual Studio uses, telling it to optimize for size over speed. Under C/C++, in the Optimization tab, change these four settings:



One last look at our file size and we see we've done quite a nice job:



And this is the complete disassembly in Olly (the RETN instruction is a little cut off at the bottom):



From 31,000 bytes to less than 1,000 (620 bytes to be exact). I guess the real question we should be asking is “Why didn’t Microsoft just start here and then add things as we need them?” I’m sure they’re crying all the way to the bank.

R4ndom