

Lecture 0x07: DLLs, Linking Continued

Crackmes: Fall 2021

Crackme 0

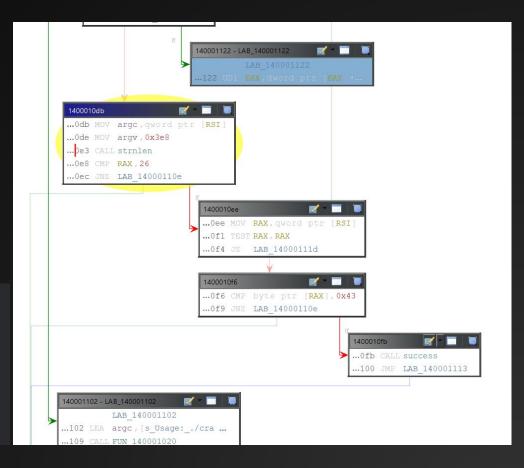
- Intended difficulty: trivial (so long as you know how to find the main entry point!)
- Run strings on the binary.
- Find main
- See that it is directly comparing your input to a static string
- GG

Crackme 1:

Intended Difficulty: Easy

The binary checks if the length of the input is 26, and if it starts with a C

```
uVar1 = strnleh (*ppcVar2,1000);
if (uVar1 == 26) {
   if (*ppcVar2 == (char *)0x0) goto LAB_14000111d;
   if (**ppcVar2 == 'C') {
      success();
      return 0;
   }
}
else {
```



Crackme 2 (Learn from my mistakes.)

Intended Difficulty:
Moderate

Question to answer: WTF is happening inside of that do-while loop?

```
🥱 📭 📓 📹 -
Decompile: main - (crackme2.exe)
LAB 1400011ad:
     if (((ulonglong)user input & 7) != 0) goto LAB 1400011b7;
     if (uVar2 != 7) goto LAB 140001185;
           ) || (CARRY8((ulonglong)&local 38,uVar2))) goto LAB 1400011ad;
         if ((longlong) &local 38 + uVar2 == 0) goto LAB 1400011ad;
         goto LAB 140001185;
```

Crackme 2 (Again, learn from my mistakes!)

- A stack variable is declared, with values inside of the ASCII range (HINT)
- We iterate through values in argv[1]
- Copy that character into a buffer
- Copy the char from the stack var into a buffer
- Checks to see if the the input char + 5 is the same as the stack variable
- If not, Fail. If yes, continue
- Once we are done checking all the characters in the stack string with no errors, Success

Crackme 2 (...Oops)

Moral of the story: Don't leave your work until the night before it is due.

You will make careless mistakes, and if you're like me and don't have tenure it could be a bad look!

Also...maybe don't use aggressive Debug statements.

```
OR git commit messages: https://www.youtube.com/watch?v=Kj
YBh7rq0-Y
```

Review: PE File Format

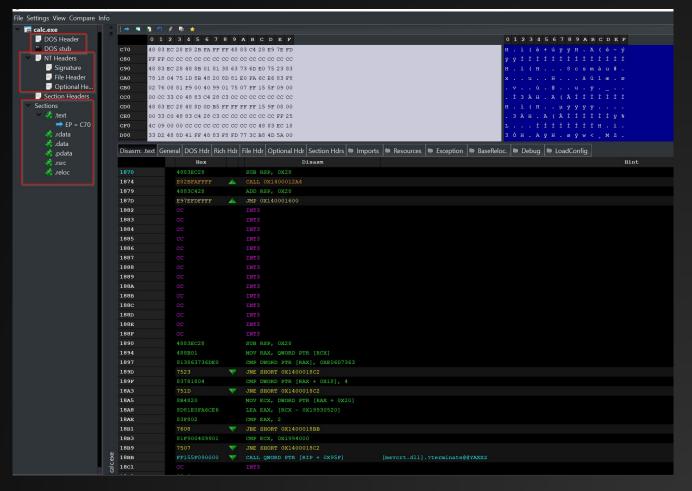
PE File Format Basic Definitions and Concepts

- Portable Executable (PE) is an executable file format used by Windows NT
- It contains information about code to execute, and how it should be executed
- In this discussion, we will use an open source tool PE-Bear to look at the structure of a PE file.

PE File Format

- PE file format is used for both userland and kernel mode executables
 - Userland: file.exe, file.dll, file.obj
 - Kernel mode: driver.sys, ntoskrnl.exe
- ullet PE is based on the Common Object File Format (COFF).
- PE format is not architecture specific (hence "portable")
 - Note this means the format can be used across multiple different architectures. The target architecture is still specified inside of the PE though
- Data is grouped together in blocks called sections, identified by headers

Calc.exe



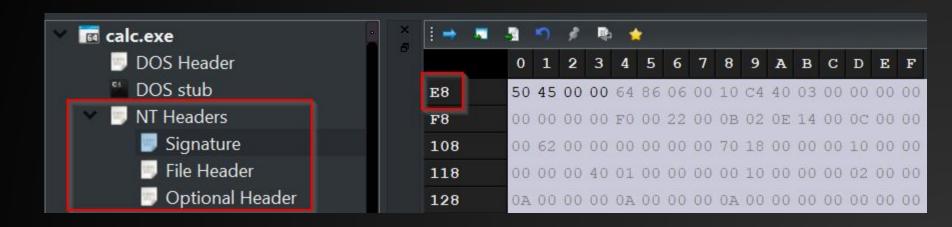
DOS Header

- The DOS header contains the magic bytes MZ that identify it as a PE
- The final entry (offset 0x3c referenced as ->e_lfanew) is the offset the of NT Headers

Offset	Name	Value
	Magic number	5A4D
	Bytes on last page of file	90
	Pages in file	3
	Relocations	0
	Size of header in paragraphs	4 MZ magic
	Minimum extra paragraphs needed	0
	Maximum extra paragraphs needed	FFFF
	Initial (relative) SS value	0
	Initial SP value	B8
	Checksum	0
	Initial IP value	0
	Initial (relative) CS value	0
	File address of relocation table	
	Overlay number	0 Address
	Reserved words[4]	0, 0, 0, 0
	OEM identifier (for OEM information)	0 Headers
	OEM information; OEM identifier specific	0
	Reserved words[10]	0, 0, 0, 40, 0, 0, 0, 0
	File address of new exe header	

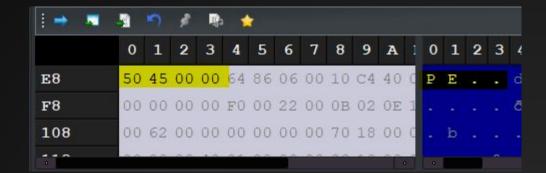
NT Headers

- Signatures
- File Header
- Optional Header



Signature

- Usually 4 bytes containing
- "PE\0\0"
- For our purposes, it is only used to verify the file format.



File Headers

Following the Signature, we have the File Headers. This gives us

- The number of sections (NumberOfSections)
- Whether or not we have a DLL/EXE (Characteristics)
- The Compilation timestamp
- A pointer to a symbol table if one exists

Disasm	General DOS Hdr Ric	h Hdr File Hdr	Optional Hdr Section Hdrs Imports
Offset	Name	Value	Meaning
EC	Machine	8664	AMD64 (K8)
	Sections Count	6	6
	Time Date Stamp	340c410	Friday, 24.09.1971 16:02:24 UTC
	Ptr to Symbol Table	0	0
	Num. of Symbols	0	0
	Size of OptionalHead	er f0	240
→ FE	Characteristics	22	

Optional Headers

I don't know why it is listed as optional. I don't think a PE can run without this section (but I could be wrong?)

The optional headers contains most of the data required to load PE

Specifically, values found here are used to build the *Import*Address Table, and perform Base Relocations

Offset	Name	Value	Value
	Magic	lagic 20B	
	Linker Ver. (Major)		
	Linker Ver. (Minor)	14	
	Size of Code	C00	
	Size of Initialized Data	6200	
	Size of Uninitialized Data		
	Entry Point		
	Base of Code		
	Image Base		
	Section Alignment	1000	
	File Alignment	200	
	OS Ver. (Major)	A	
	OS Ver. (Minor)		
	Image Ver. (Major)	A	
	Image Ver. (Minor)		
	Subsystem Ver. (Major)	A	
	Subsystem Ver. Minor)		
	Win32 Version Value		
	Size of Image	B000	
	Size of Headers	400	
	Checksum	14163	
	Subsystem	2	Windows GUI
× 146	DLL Characteristics	C160	
		40	DLL can move
		100	Image is NX compatible
		4000	Guard
		8000	TerminalServer aware
	Size of Stack Reserve	80000	
	Size of Stack Commit	2000	
	Size of Heap Reserve	100000	
	Size of Heap Commit	1000	

Optional Headers (pt 1)

- Magic: Architecture of image
- Entry Point: Relative virtual address (RVA) from the Base Address
- Image Base: (prefered) Base address: Where in memory the PE "prefers" to be loaded. If the location is unavailable, the Image needs to be relocated



Optional Headers (pt 2)

- SizeOfImage: the virtual size of the image
- SizeOfHeaders: the size of the headers
- DLLCharacteristics: flags including knowledge of hardening features such as ASLR/ CFG...etc. Not super important for us other than assuming knowledge of ASLR.

Offset	Name	Value	Value
100	Magic	20B	NT64
	Linker Ver. (Major)	E	
	Linker Ver. (Minor)	14	
	Size of Code	C00	
	Size of Initialized Data	6200	
	Size of Uninitialized Data	0	
	Entry Point		
	Base of Code		
	Image Base	140000000 lmag	
	Section Alignment	1000	
	File Alignment	200 head	
	OS Ver. (Major)	A	
	OS Ver. (Minor)	0	
	Image Ver. (Major)	A	GUI?)
	Image Ver. (Minor)	0	
	Subsystem Ver. (Major)	A	
	Subsystem Ver. Minor)	0	
	Win32 Version Value	0	
	Size of Image	B000	
	Size of Headers	400	
	Checksum	14163	
	Subsystem	2	Windows GUI
146	DLL Characteristics	C160	
		40	DLL can move
		100	Image is NX compatible
		4000	Guard
		8000	TerminalServer aware
	Size of Stack Reserve	80000	
	Size of Stack Commit	2000	
	Size of Heap Reserve	100000	
	Size of Heap Commit	1000	

When a DLL is loaded by a process:

- It is not guaranteed to be placed in the same absolute location in memory
- Nor is it guaranteed to be placed in the same relative location

To handle variability in load location, the programer can can declare all imports in the *Import Address Table* (IAT)

This shifts the work to the PE loader to resolve all the imports away from the programmer!

When a dependency is declared in the IAT, the PE loader will attempt to resolve the dependency. If it fails, the program crashes.

Recall this is a subtype of dynamic linking called *Implicit Linking*

The IAT is effectively a collection of key value pairs, where the key is an identifier of an imported function, and the value is the address of a small function that jumps to the location of the real function.

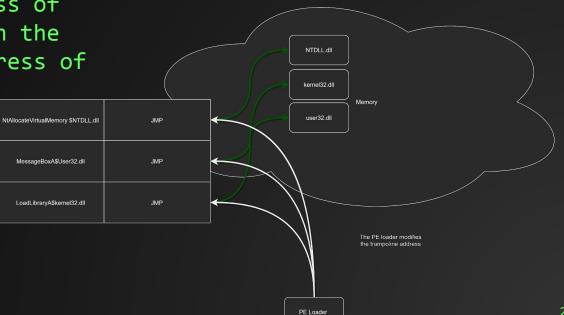
This is reflected in x64dbg when we set a breakpoint at a call to a remote library: the breakpoint is set at the Virtual Addresses of the remote function, but in our code we reference it via the IAT



When resolving imports, the PE loader will load all required DLLs, identify the address of the loaded DLL, and patch the IAT with the correct address of the requested function

MessageBoxA\$User32.dll

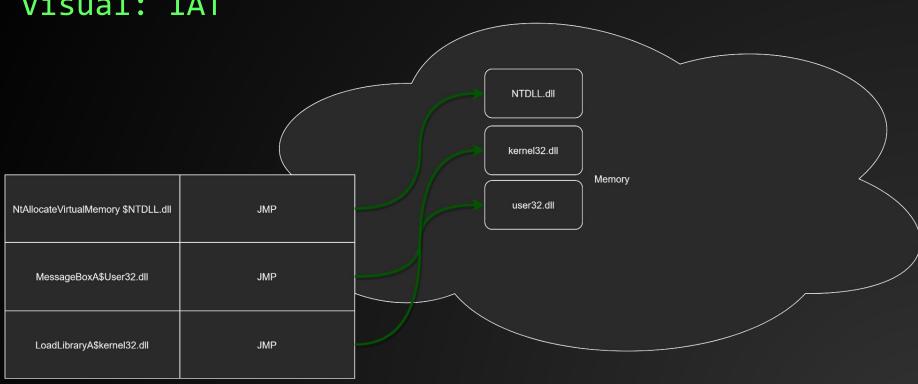
LoadLibraryA\$kernel32.dll



- The application code simply calls the function referenced in the IAT, which is itself a trampoline to the real code.
- I.e., it is a static location user code can can reference, which is simply an unconditional jump to the real virtual address of the required function.
- WHen a PE lives on disk, this jump table is null, and is set at runtime by the PE loader.

000000000040103E 0000000000040103F	90	nop nop	
0000000000401040	FF25 F23F0000	jmp qword ptr ds:[<&MessageBoxA>]	JMP.&MessageBoxA
0000000000401046	90	nop	
0000000000401047	90	nop	
0000000000401048	0F1F8400 00000000	nop dword ptr ds:[rax+rax],eax	
0000000000401050	FF	????	
0000000000401051	FF	???	

Visual: IAT

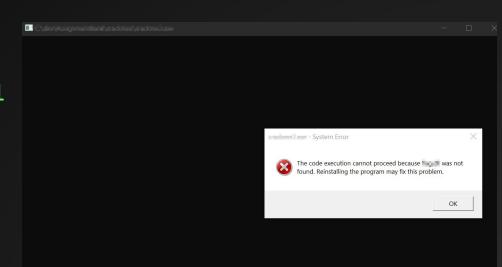


Example of a crash:

Double click an EXE with a missing dependency, and you will likely see a message like this:

...Hint Hint Nudge Nudge...

You might need a bit of programming for 1 of the crackmes!





Dynamically Linked Libraries:

- Refresh: What is a DLL?
- How do we build a DLL?
- How do we execute a DLL?
- How do we export functions in a DLL?
- How do we call functions from a DLL?

DLL

A PE with with DLL characteristic field set.

Usually it has exported functions which can be referenced by code outside of the DLL

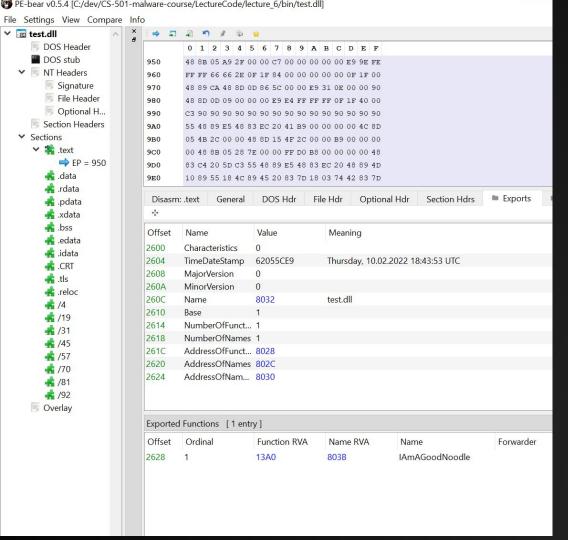
Building DLLs

Exporting Code

```
__declspec(dllexport) int IAmAGoodNoodle()
{
    MessageBoxA(NULL, "I am a good noodle!", "Very good!", MB_OK );
    return 0;
}
```

Note for Cpp, you will need an extern "C" directive

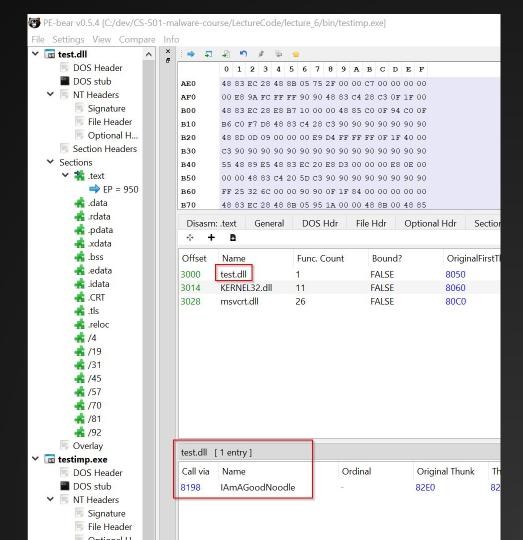
Demo Building a DLL you can link against



DLL Search

The PE loader will check multiple well known paths, finally checking the current directory for the required DLL

If you aren't careful with where you set your dependency, you can open up your code to DLL hijacking

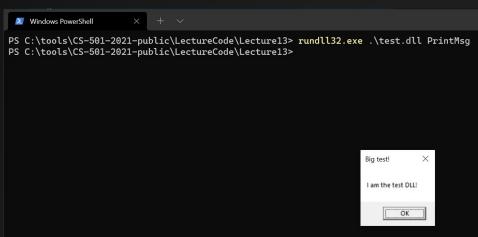


Running DLLs

Rundl32.exe (there is a 32-bit version and a 64-bit version)

Rundl32.exe can execute specific functions, or simply the DLLMain

If you give it the name of an unexported function, it will call ProcessAttach then error out



Testing our DLL: rund132

```
Directory: C:\dev\LS-5UI-malware-course\LectureLode\Lecture_b
                                                                        Very good!
                      LastWriteTime
                                            Length Name
 Mode
                                                                        I am a good noodle!
                                                   bin
                2/10/2022 1:43 PM
                2/10/2022 12:48 PM
                                             89056 a.exe
               2/10/2022 1:43 PM
                                               337 Makefile
                                                                                 OK
               2/10/2022 12:48 PM
                                               618 mbox.c
              2/10/2022 1:39 PM
                                               788 testdll.c
               2/10/2022 1:42 PM
                                                94 testimp.c
                2/10/2022 1:43 PM
                                             57551 testimp.exe
 -a---
PS C:\dev\CS-501-malware-course\LectureCode\lecture_6> cd .\bin\
 PS C:\dev\CS-501-malware-course\LectureCode\lecture_6\bin> rundll32.exe
PS C:\dev\CS-501-malware-course\LectureCode\lecture_6\bin> rundll32.exe .\test.dll
 PS C:\dev\CS-501-malware-course\LectureCode\lecture_6\bin> rundll32.exe .\test.dll IAmAGoodNoodle
 PS C:\dev\CS-501-malware-course\LectureCode\lecture_6\bin>
```

Testing our DLL: Implicit Linking



Dynamic Linking→Explicit Linking

- The programer explicitly calls LoadLibrary followed by GetProcAddress, and has the proper function prototypes inside of their code. Bot of these functions live in Kernel32.dll
- The legitimate reason to do this, is to allow the execution of the program to continue even if a DLL does not exist.
- If an import for an implicitly loaded DLL is missing, the program stops.
- The application using an explicitly loaded DLL can choose how to handle a missing DLL.
- This is a common in malware trying to hide its imports

How does this work?

- LoadLibraryA: gets a handle to a library in memory.
- If the library isn't already loaded into memory, it will load it! If it can't, it returns a NULL pointer
 - The handle it returns is the Virtual Address of the memory mapped library for the current processes Private Address space!
 - In particular, it is the *base address* of the library.
- GetProcAddress: given a handle to a library, and an identifier for a function such as a name, it returns the Virtual Address of the funcion!

Using Function Pointers

- Functions aren't first class objects in C/C++
- However, functions themselves are just memory addresses!
 - with arguments that give the compiler hints on how to call it!
- To invoke a function by a memory address, we first need the address of the function, and information about how to call it
- We can either handle this using raw assembly, or we can declare a function pointer type with the arguments and calling conventions specified
- Then we just cast the raw address to a function pointer type and call it!

Demo: Explicitly Linked Mbox

Reverse Engineering Implications

- If a dll is explicitly linked, you cannot set a breakpoint for the library in the entry point (why?)
- The DLL doesn't show up in the IAT

How can we hide the strings for LoadLibrary/GetProcAddress?

Discussion: Defense Evasion using different types of linking!