软件安全一恶意代码机理与防护 C3 PE文件格式

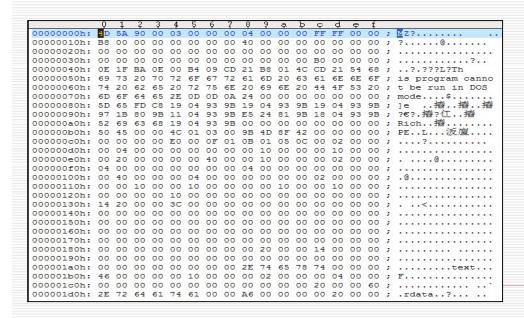
彭国军 教授 武汉大学国家网络安全学院 guojpeng@whu.edu.cn

本讲的内容提纲

- 3.1 PE文件及其表现形式
- 3.2 PE文件格式与恶意软件的关系
- 3.3 PE文件格式总体结构
- 3.4 代码节与数据节
- 3.5 引入函数节: PE文件的引入函数机制
- 3.6 引出函数节: DLL文件的函数引出机制
- 3.7 资源节: 文件资源索引、定位与修改
- 3.8 重定位节: 镜像地址改变后的地址自动修正

3.3 PE文件格式总体结构

□ test.exe





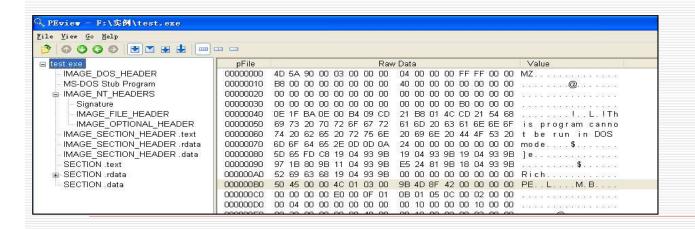
看雪学院工具集:

http://tools.pediy.com



PE文件格式查看工具1-PEView

□ PEView: 可按照PE文件格式对目标文件的 各字段进行详细解析。



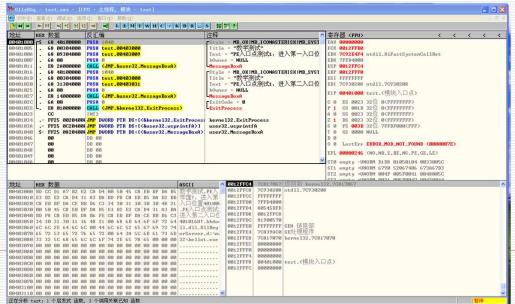
PE文件格式查看工具2-Stud_PE

□ Stud_PE: 可按照PE文件格式对目标文件的各字段进行详细解析。



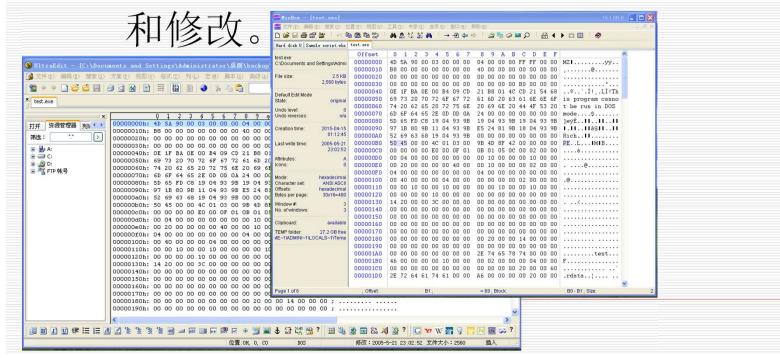
PE程序调试工具-Ollydbg

□ Ollydbg: 可跟踪目标程序的执行过程,属于用户态调试工具。



16进制文件编辑工具-UltraEdit

□ UltraEdit: 可对目标文件进行16进制查看



PE文件格式总体结构

■ test.exe - IMAGE_DOS_HEADER - MS-DOS Stub Program - IMAGE_NT_HEADERS - Signature - IMAGE_FILE_HEADER - IMAGE_OPTIONAL_HEADER - IMAGE_SECTION_HEADER .rdata - IMAGE_SECTION_HEADER .rdata - IMAGE_SECTION_HEADER .data - SECTION .text - SECTION .rdata - SECTION .rdata

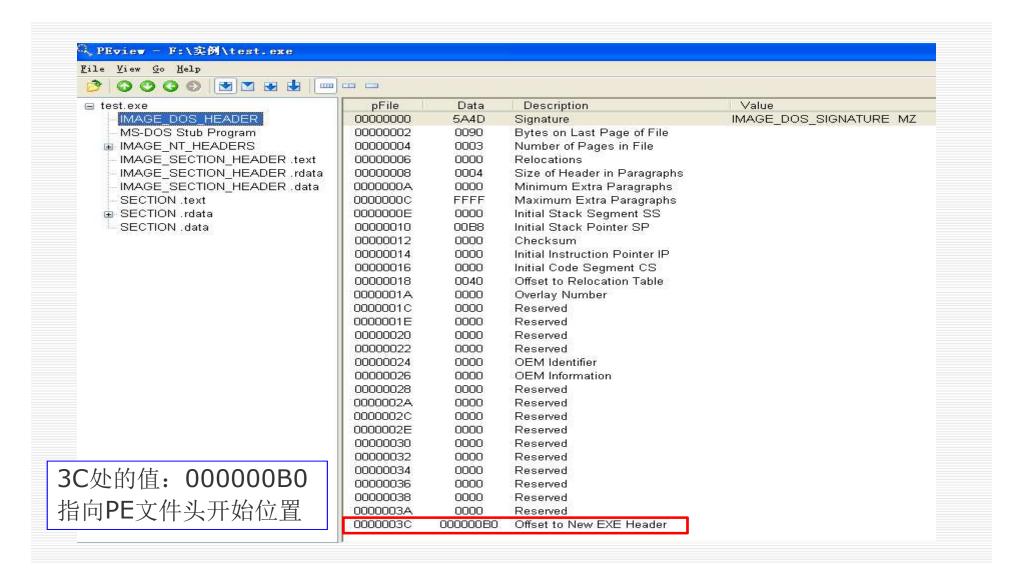
1.DOS MZ header	
2.DOS stub	
3.PE header	
4.Section table	
5-1 Section 1	
5-2 Section 2	
Section	
5.3 Section n	

(1) MS-DOS MZ文件头(0x40)+

(2) DOS Stub

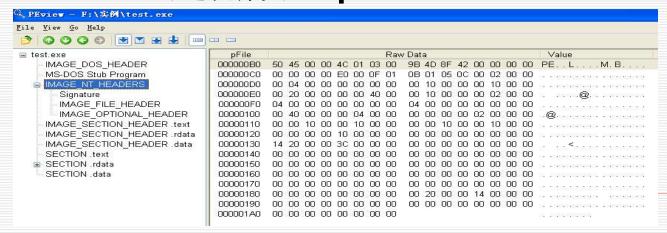
□作用:

- 定位PE文件头开始位置,也可用于PE文件合法性检测
- DOS下运行该程序时,将提示用户: "This Program cannot be run in DOS mode"!



(3) PE header

- □ PE header 由三部分组成【开始于000000B0】
 - 字串 "PE\0\0" (**Signature**)
 - 映像文件头(FileHeader)
 - 可选映像头(OptionalHeader)



PE Header结构

```
000000c0h: 00 00 00 00 E0 00 0F 01 0B 01 05 0C 00 02 00 00 ; ....?......
000000d0h: 00 04 00 00 00 00 00 00 00 00 00 00 10 00 00 : ......
000000e0h: 00 20 00 00 00 00 40 00 00 10 00 00 00 02 00 00 : . . . . . @. . . .
00000100h: 00 40 00 00 04 00 00 00 00 00 02 00 00 00 : .@........
00000180h: 00 00 00 00 00 00 00 00 20 00 00 14 00 00 00 : ........
000001a0h: 00 00 00 00 00 00 00 00
```

1)字串 "PE\0\0"

- hello-2.5.exe

 IMAGE_DOS_HEADER

 MS-DOS Stub Program

 IMAGE_NT_HEADERS

 IMAGE_TILE_HEADER

 IMAGE_OPTIONAL_HEADER

 IMAGE_SECTION_HEADER.rdata

 IMAGE_SECTION_HEADER.data

 SECTION.text

 SECTION.rdata

 SECTION.rdata
- □ Signature —dword类型, 值为50h, 45h, 00h, 00h(PE\0\0)。
 - 本域为PE标记,可以此识别给定文件是否为有效PE 文件。

PE\0\0 0000000b0h: 50 45 00 00 4C 01 03 00 9B 4D 8F 42 00 00 00 00 ; PE.L... 液度.... 0000000c0h: 00 00 00 00 E0 00 0F 01 0B 01 05 0C 00 02 00 00 ; ...?......

2) 映像文件头(0x14)

hello-2.5.exe

IMAGE_DOS_HEADER

MS-DOS Stub Program

IMAGE_NT_HEADERS

Signature

IMAGE_OPTIONAL_HEADER

IMAGE_SECTION_HEADER.rdata

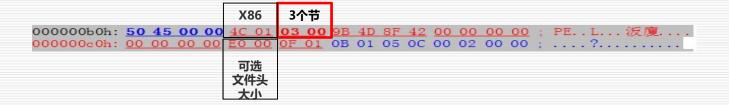
IMAGE_SECTION_HEADER.data

SECTION .text

SECTION .text

SECTION .rdata

□ 该结构域包含了关于PE文件物理分布的信息, 比如节数目、后续可选文件头大小、机器类 型等。



映像文件头的结构

顺序	名字	大小 (字节)	描述
1	Machine *	2	机器类型,x86为14CH
2	NumberOfSection **	2	文件中节的个数
3	TimeDataStamp	4	生成该文件的时间
4	PointerToSymbleTabl e	4	COFF符号表的偏移
5	NumberOfSymbols	4	符号数目
6	SizeOfOptionalHead er*	2	可选头的大小
7	Characteristics *	2	关于文件信息的标记,比如文件是 exe 还是 dll

3) 可选文件头

hello-2.5.exe

IMAGE_DOS_HEADER

MS-DOS Stub Program

IMAGE_NT_HEADERS

Signature

IMAGE_OPTIONAL_HEADER

IMAGE_SECTION_HEADER.rdata

IMAGE_SECTION_HEADER.data

SECTION.text

SECTION.text

SECTION.text

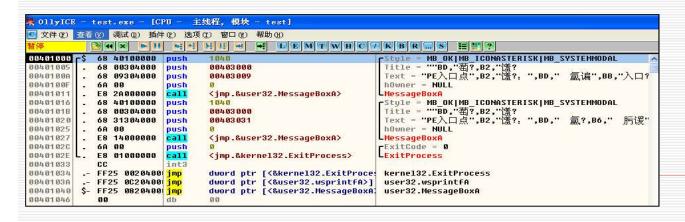
SECTION.data

- □ 定义了PE文件的很多关键信息
 - □ 内存镜像加载地址(ImageBase)
 - □ 程序入口点(代码从哪里开始执行?)
 - □ 节在文件和内存中的对齐粒度
 - □ 本程序在内存中的镜像大小、文件头大小等

```
000000b0h: <u>50 45 00 00 4C 01</u> <u>03 00 9B 4D 8F 42 00 00 00 00</u> ; PE..L...  液廈....
000000c0h: 00 00 00 00 E0 00 0F 01 0B 01 05 0C 00 02 00 00 ; ....?......
000000d0h: 00 04 00 00 00 00 00 00 00 10 00 00 00 10 00 00 : . . . . .
000000e0h: 00 20 00 00 00 00 40 00 00 10 00 00 02 00 00 ; . ....@.
00000100h: 00 40 00 00 00 04 00 00 00 00 00 02 00 00 00 : .@.....
00000120h: 00 00 00 00 10 00 00 00 00 00 00 00 00 00 00 ; ..
00000130h: 14 20 00 00 3C 00 00 00 00 00 00 00 00 00 00 : . . .
00000180h: 00 00 00 00 00 00 00 00 20 00 14 00 00 00 ; ......
000001a0h: 00 00 00 00 00 00 00 2E 74 65 78 74 00 00 00
```

几个概念-1

- ☐ ImageBase:
 - PE文件在内存中的优先装载地址。
- □ RVA地址:
 - Relative Virtual Address,相对虚拟地址,它是相对内存中 ImageBase的偏移位置。



几个概念-2

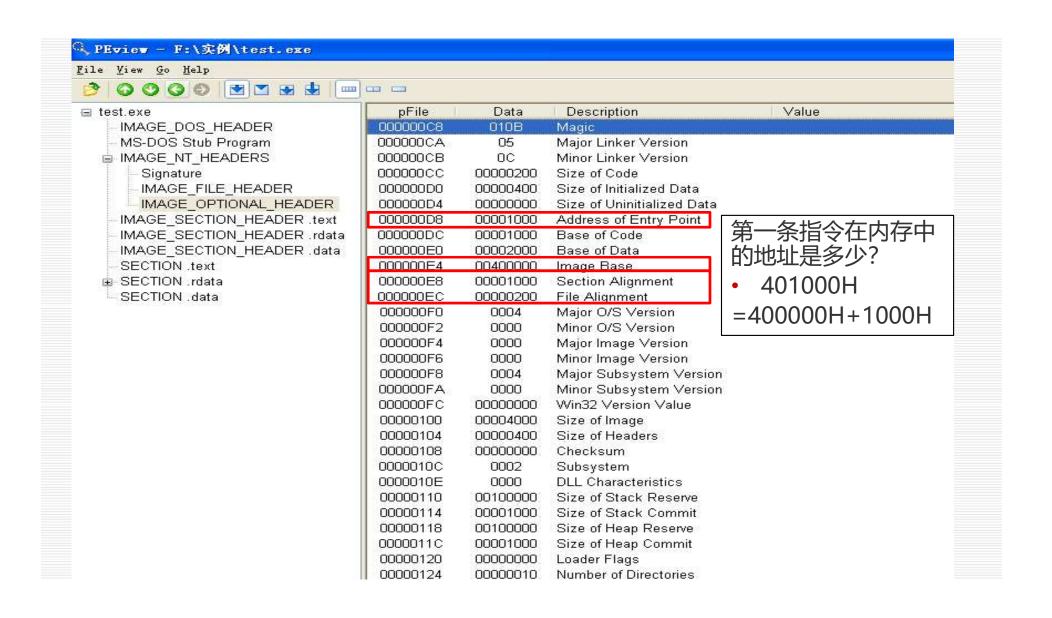
- □ 对齐粒度
 - 比喻:桶的容量为100升,现有367升水,请问需要使用多少个桶?
- □ 问题: 代码节的代码实际长度为0x46字节
 - 文件中节对齐粒度为0x200,
 - 内存中节对齐粒度为0x1000字节,

请问代码节在文件和内存中分别占用多少字节?

■ 为什么PE文件中有很多"00"字节?

可选文件头中的一些关键字段

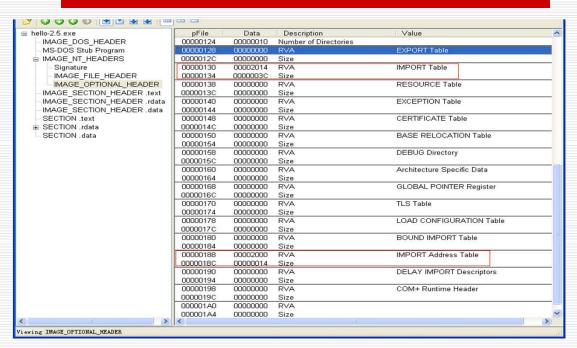
名字	描述
AddressOfEntryPoint *	PE装载器准备运行的PE文件的第一条指 令的RVA。 <i>(病毒感染中通用关键字段)</i>
(位置 D8H,4 字节)	
ImageBase (位置: E4H, 4字节)	PE文件的优先装载地址。比如,如果该值是400000h,PE装载器将尝试把文件装到虚拟地址空间的400000h处。
SectionAlignment	内存中节对齐的粒度。
(位置: E8H, 4字节)	
FileAlignment	文件中节对齐的粒度。
(位置: ECH, 4字节)	

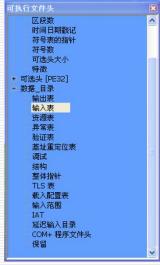


AddressOfEntryPoint的作用

- □<u>是否可以修改</u>AddressOfEntryPoint指
 - 向任意代码?
 - D8H处的4个字节
 - 演示:修改1000H为1016H

Directory—16项 * 8字节

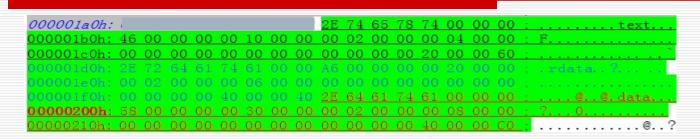


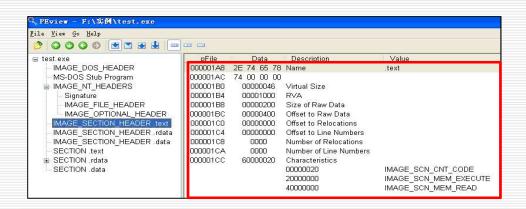


(4) 节表

- □ 节表,是紧挨着 PE header 的一个结构数组。
- □ 每一个节均对应一个节表项:
 - 节名;
 - 节在文件和内存中的开始地址;
 - 长度;
 - 节属性等。

test.exe的节表





Characteristics-节属性

0x00000020? 这个块包含代码。置位

0x00000040? 这个块包含已初始化的数据。

0x0000080? 这个块包含未初始化的数据(如.bss 块)

0x00000200? 这个块包含注释或其它的信息。

0x00000800? 这个块的内容不应放进最终的EXE文件中。

0x02000000? 这个块可以被丢弃,因为一旦它被载入,其进程就不需要它。最通常的可丢弃块是基本重定位块(.reloc)。

0x10000000? 这个块是可共享的。

0x20000000? 这个块是可执行的。

0x40000000? 这个块是可读的。

0x80000000? 这个块是可写的。



(5) 节

- □可执行文件的核心部分。
- □ PE文件一般都有多个"节",比较常见的有:
 - 代码节
 - ■数据节
 - 引入函数节
 - 资源节等(如图标)
 - 引出函数节(DLL文件中常见)
 - 重定位节(DLL文件中常见)