Return-to-libc Attack Lab

1 Lab Overview

The learning objective of this lab is for students to gain first-hand experience with a variant of the buffer-overflow attack; this attack can bypass a protection scheme currently implemented in major Linux operating systems. A common way to exploit a buffer-overflow vulnerability is to overflow the buffer with malicious shellcode, and then cause the vulnerable program to jump to the shellcode that is stored in the stack. To prevent these types of attacks, some operating systems provide an ability to make stacks non-executable; therefore, jumping to the shellcode will cause the program to fail.

Unfortunately, the above protection scheme is not fool-proof; there exists a variant of buffer-overflow attack called the `return-to-libc` attack, which does not need an executable stack; it does not even use shell code. Instead, it causes the vulnerable program to jump to some existing code, such as the `system()` function in the `libc` library, which is already loaded into the memory.

In this lab, students are given a program with a buffer-overflow vulnerability; their task is to develop a `return-to-libc` attack to exploit the vulnerability and finally to gain the root privilege. In addition to the attacks, students will be guided to walk through several protection schemes that have been implemented in Ubuntu to counter against the buffer-overflow attacks. Students need to evaluate whether the schemes work or not and explain why.

2 Lab Environment

This lab runs in the Labtainer framework, available at http://my.nps.edu/web/c3o/labtainers. That site includes links to a pre-built virtual machine that has Labtainers installed, however Labtainers can be run on any Linux host that supports Docker containers.

From your labtainer-student directory start the lab using:

```
labtainer retlibc
```

Links to this lab manual and to an empty lab report will be displayed. If you create your lab report on a separate system, be sure to copy it back to the specified location on your Linux system.

3 Lab Tasks

You are required to exploit the "retlib" program. You are provided the source code to the program, but you are not to alter the program source code.

3.1 Initial Setup

Address Space Randomization.   `Ubuntu` and several other Linux-based systems uses address space randomization to randomize the starting address of heap and stack. This makes guessing the exact addresses difficult; guessing addresses is one of the critical steps of buffer-overflow attacks. In this lab, we disable this feature using the following command:

```
sudo sysctl -w kernel.randomize_va_space=0
```

The StackGuard Protection Scheme.   The GCC compiler implements a security mechanism called "Stack Guard" to prevent buffer overflows. In the presence of this protection, buffer overflows will fail. You can disable this protection if you compile the program using the *-fno-stack-protector* switch. For example, to compile a program example.c with Stack Guard disabled, you may use the following command:

```
$ gcc -m32 -fno-stack-protector example.c
```

The retlib program was compiled without StackGuard.

Non-Executable Stack.   `Ubuntu` used to allow executable stacks, but this has now changed: the binary images of programs (and shared libraries) must declare whether they require executable stacks or not, i.e., they need to mark a field in the program header. The kernel or dynamic linker uses this marking to decide whether to make the stack of this running program executable or non-executable. This marking is done automatically by the recent versions of `gcc`, and by default, the stack is set to be non-executable. To change that, use the following option when compiling programs:

```
For executable stack:
$ gcc -m32 -z execstack  -o test test.c

For non-executable stack:
$ gcc -m32 -z noexecstack  -o test test.c
```

Because the objective of this lab is to show that the non-executable stack protection does not work, the retlib program was compiled with the `"-z noexecstack"` option.

Note for Instructors: For this lab, a lab session is desirable, especially if students are not familiar with the tools and the enviornments. If an instructor plans to hold a lab session (by himself/herself or by a TA), it is suggested the following to be covered in the lab session [1]:

1. The use of Labtainers.

2. Basic use of `gdb` debug commands and stack stucture.

Review the retlib.c program. It has a buffer overflow vulnerability. It first reads an input of size 80 bytes from a file called "badfile" into a buffer that is smaller than 80 bytes, causing the overflow. The function fread() does not check boundaries, so a buffer overflow will occur. Since this program is a set-root-uid program, if a normal user can exploit this buffer overflow vulnerability, the normal user might be able to get a root shell. It should be noted that the program gets its input from a file called "badfile". This file is under the users' control. Our objective is to create the contents for "badfile", such that when the vulnerable program copies the contents into its buffer, a root shell can be spawned.

---

[1]We assume that the instructor has already covered the concepts of the attacks in the lecture, so we do not include them in the lab session.

3.2 Task 1: Exploiting the Vulnerability

Create the **badfile**. Your home directory includes an "exploit.c" program that you can modify in order to create the badfile.

You need to figure out the values for the addresses in that program, as well as where to store those addresses. If you incorrectly calculate the locations, your attack might not work.

After you modify the exploit.c program, compile using `compile_exploit.sh` and run it; this will generate the contents for "badfile". Run the vulnerable program `retlib`. If your exploit is implemented correctly, when the function `bof` returns, it will return to the `system()` libc function, and execute `system("/bin/sh")`. If the vulnerable program is running with the root privilege, you can get the root shell at this point.

It should be noted that the `exit()` function is not very necessary for this attack; however, without this function, when `system()` returns, the program might crash, causing suspicions.

```
$ ./compile_exploit.sh
$./exploit          // create the badfile
$./retlib           // launch the attack by running the vulnerable program
# <---- You've got a root shell!
```

Questions.   In your report, please answer the following questions:

- Please describe how you decide the values for X, Y and Z. Either show us your reasoning, or if you use trial-and-error approach, show your trials.

- After your attack is successful, change the file name of `retlib` to a different name, making sure that the length of the file names are different. For example, you can change it to `newretlib`. Repeat the attack (without changing the content of `badfile`). Is your attack successful or not? If it does not succeed, explain why.

3.3 Task 2: Address Randomization

In this task, let us turn on the Ubuntu's address randomization protection. We run the same attack developed in Task 1. Can you get a shell? If not, what is the problem? How does the address randomization make your return-to-libc attack difficult? You should describe your observation and explanation in your lab report. You can use the following instructions to turn on the address randomization:

```
sudo sysctl -w kernel.randomize_va_space=2
```

3.4 Task 3: Stack Guard Protection

In this task, let us turn on the Ubuntu's Stack Guard protection. Please remember to turn off the address randomization protection. We run the same attack developed in Task 1. Can you get a shell? If not, what is the problem? How does the Stack Guard protection make your return-to-libc attack difficult? You should describe your observation and explanation in your lab report. You can use the following instructions to compile your program with the Stack Guard protection turned on.

```
sudo su
# gcc -m32 -z noexecstack  -o retlib retlib.c
# chmod 4755 retlib
# exit
```

4 Guidelines: Understanding the function call mechanism

4.1 Find out the addresses of libc functions

To find out the address of any libc function, you can use the following gdb commands (a.out is an arbitrary program):

```
$ gdb a.out

(gdb) b main
(gdb) r
(gdb) p system
 $1 = {<text variable, no debug info>} 0x9b4550 <system>
(gdb) p exit
 $2 = {<text variable, no debug info>} 0x9a9b70 <exit>
```

From the above gdb commands, we can find out that the address for the system() function is 0x9b4550, and the address for the exit() function is 0x9a9b70. The actual addresses in your system might be different from these numbers.

4.2 Putting the shell string in the memory

One of the challenge in this lab is to put the string "/bin/sh" into the memory, and get its address. This can be achieved using environment variables. When a C program is executed, it inherits all the environment variables from the shell that executes it. The environment variable **SHELL** points directly to /bin/bash and is needed by other programs, so we introduce a new shell variable **MYSHELL** and make it point to zsh

```
$ export MYSHELL=/bin/sh
```

We will use the address of this variable as an argument to system() call. The location of this variable in the memory can be found out easily using the following program:

```
void main(){
  char* shell =  getenv("MYSHELL");
  if (shell)
     printf("%x\n", (unsigned int)shell);
}
```

If the address randomization is turned off, you will find out that the same address is printed out. However, when you run the vulnerabile program retlib, the address of the environment variable might not be exactly the same as the one that you get by running the above program; such an address can even change when you change the name of your program (the number of characters in the file name makes difference). The good news is, the address of the shell will be quite close to what you print out using the above program. Therefore, you might need to try a few times to succeed.

4.3 Understand the Stack

To know how to conduct the return-to-libc attack, it is essential to understand how the stack works. We use a small C program to understand the effects of a function invocation on the stack.

```
/* foobar.c */
#include<stdio.h>
void foo(int x)
{
  printf("Hello world: %d\n", x);
}

int main()
{
  foo(1);
  return 0;
}
```
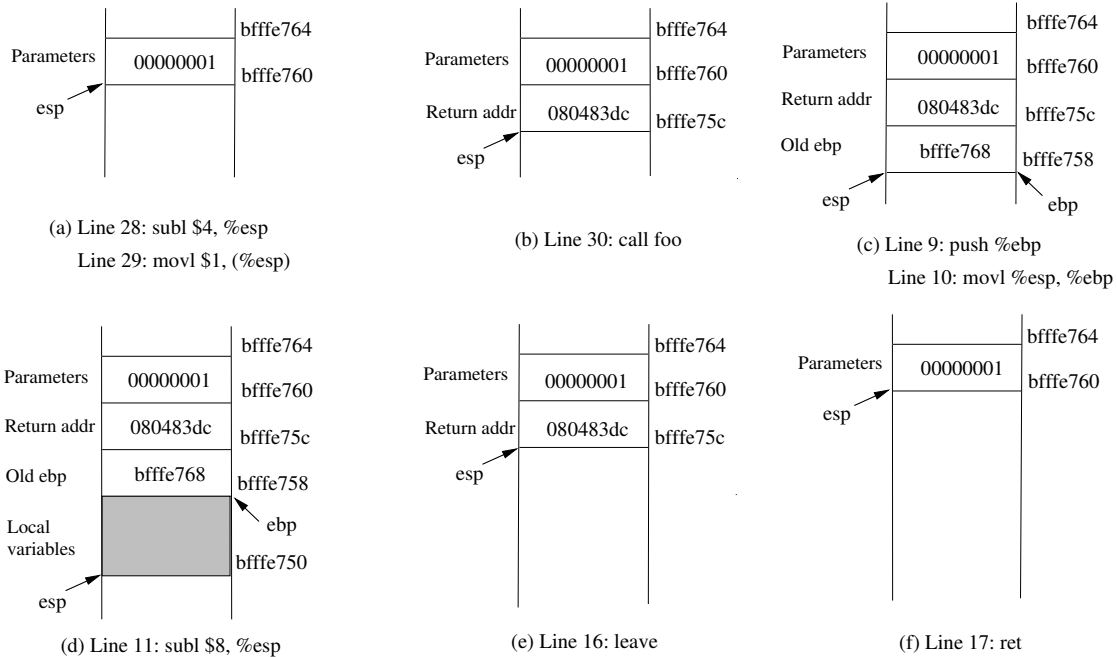
We can use "`gcc -m32 -S foobar.c`" to compile this program to the assembly code. The resulting file `foobar.s` will look like the following:

```
    ......
  8 foo:
  9         pushl   %ebp
 10         movl    %esp, %ebp
 11         subl    $8, %esp
 12         movl    8(%ebp), %eax
 13         movl    %eax, 4(%esp)
 14         movl    $.LC0, (%esp)   : string "Hello world: %d\n"
 15         call    printf
 16         leave
 17         ret
    ......
 21 main:
 22         leal    4(%esp), %ecx
 23         andl    $-16, %esp
 24         pushl   -4(%ecx)
 25         pushl   %ebp
 26         movl    %esp, %ebp
 27         pushl   %ecx
 28         subl    $4, %esp
 29         movl    $1, (%esp)
 30         call    foo
 31         movl    $0, %eax
 32         addl    $4, %esp
 33         popl    %ecx
 34         popl    %ebp
 35         leal    -4(%ecx), %esp
 36         ret
```

4.4 Calling and Entering `foo()`

Let us concentrate on the stack while calling `foo()`. We can ignore the stack before that. Please note that line numbers instead of instruction addresses are used in this explanation.

(a) Line 28: subl $4, %esp
Line 29: movl $1, (%esp)

(b) Line 30: call foo

(c) Line 9: push %ebp
Line 10: movl %esp, %ebp

(d) Line 11: subl $8, %esp

(e) Line 16: leave

(f) Line 17: ret

Figure 1: Entering and Leaving `foo()`

- **Line 28-29:**: These two statements push the value 1, i.e. the argument to the `foo()`, into the stack. This operation increments `%esp` by four. The stack after these two statements is depicted in Figure **??**(a).

- **Line 30: `call foo`**: The statement pushes the address of the next instruction that immediately follows the `call` statement into the stack (i.e the return address), and then jumps to the code of `foo()`. The current stack is depicted in Figure **??**(b).

- **Line 9-10**: The first line of the function `foo()` pushes `%ebp` into the stack, to save the previous frame pointer. The second line lets `%ebp` point to the current frame. The current stack is depicted in Figure **??**(c).

- **Line 11: `subl $8, %esp`**: The stack pointer is modified to allocate space (8 bytes) for local variables and the two arguments passed to `printf`. Since there is no local variable in function `foo`, the 8 bytes are for arguments only. See Figure **??**(d).

4.5 Leaving `foo()`

Now the control has passed to the function `foo()`. Let us see what happens to the stack when the function returns.

- **Line 16: `leave`**: This instruction implicitly performs two instructions (it was a macro in earlier x86 releases, but was made into an instruction later):

```
mov   %ebp, %esp
pop   %ebp
```

The first statement release the stack space allocated for the function; the second statement recover the previous frame pointer. The current stack is depicted in Figure **??**(e).

- **Line 17: `ret`**: This instruction simply pops the return address out of the stack, and then jump to the return address. The current stack is depicted in Figure **??**(f).

- **Line 32: `addl $4, %esp`**: Further resotre the stack by releasing more memories allocated for foo. As you can clearly see that the stack is now in exactly the same state as it was before entering the function foo (i.e., before line 28).

References

[1] c0ntext   Bypassing   non-executable-stack   during   exploitation   using   return-to-libc http://www.infosecwriters.com/text_resources/pdf/return-to-libc.pdf

[2] Phrack by Nergal Advanced return-to-libc exploit(s) *Phrack 49*, Volume 0xb, Issue 0x3a. Available at http://www.phrack.org/archives/58/p58-0x04