缓冲区溢出

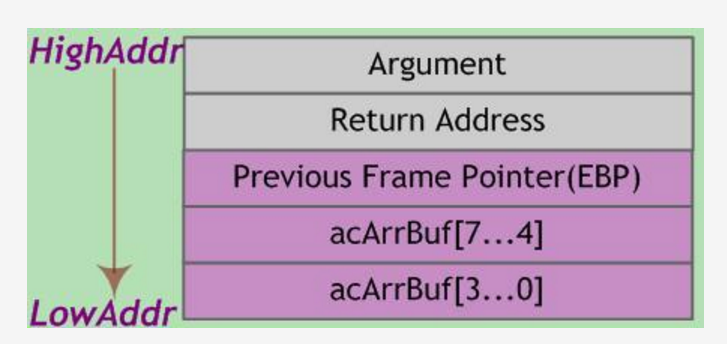
1.缓冲区溢出原理

缓冲区是一块连续的计算机内存区域，可保存相同数据类型的多个实例。缓冲区可以是栈(局部变量，函数的参数值)、堆(由new创建的对象)和静态数据区(全局或静态)。在C/C++语言中，通常使用字符数组和malloc/new之类内存分配函数实现缓冲区。溢出指数据被添加到分配给该缓冲区的内存块之外。缓冲区溢出是最常见的程序缺陷。

     栈帧结构的引入为高级语言中实现函数或过程调用提供直接的硬件支持，但由于将函数返回地址这样的重要数据保存在程序员可见的堆栈中，因此也给系统安全带来隐患。若将函数返回地址修改为指向一段精心安排的恶意代码，则可达到危害系统安全的目的。此外，堆栈的正确恢复依赖于压栈的EBP值的正确性，但EBP域邻近局部变量，若编程中有意无意地通过局部变量的地址偏移窜改EBP值(对于一个栈来说，寄存器EBP和ESP分别指向指向系统栈最上面一个[栈帧](http://baike.baidu.com/view/8128123.htm" \t "http://blog.csdn.net/ylyuanlu/article/details/_blank)的底部(高地址)和栈帧顶部（低地址）)，则程序的行为将变得非常危险。

     由于C/C++语言没有数组越界检查机制，当向局部数组缓冲区里写入的数据超过为其分配的大小时，就会发生缓冲区溢出。攻击者可利用缓冲区溢出来窜改进程运行时栈，从而改变程序正常流向，轻则导致程序崩溃，重则系统特权被窃取。

     例如，对于下图的栈结构：

IMG_256

     若攻击者用一个有意义的地址(否则会出现段错误)覆盖返回地址的内容，函数返回时就会去执行该地址处事先安排好的攻击代码。最常见的手段是通过制造缓冲区溢出使程序运行一个用户shell，再通过shell执行其它命令。若该程序有root或suid执行权限，则攻击者就获得一个有root权限的shell，进而可对系统进行任意操作。

     除通过使堆栈缓冲区溢出而更改返回地址外，还可改写局部变量(尤其函数指针)以利用缓冲区溢出缺陷。

2.缓冲区溢出实例

 (1)修改邻接变量

函数的局部变量在栈中一个接一个排列，如果这些局部变量有数组之类的缓冲区，并且存在数组越界的缺陷，那么越界的数组元素就有可能破坏栈中相邻变量的值

#define PASSWORD "1234567“

int verify\_password (char \*password)

{

int authenticated;

char buffer[8];

authenticated=strcmp(password,PASSWORD);

strcpy(buffer,password);

return authenticated;

}

main()

{

int valid\_flag=0;

char password[1024];

while(1)

{

printf("please input password: ");

scanf("%s",password);

valid\_flag = verify\_password(password);

if(valid\_flag)

{

printf("incorrect password!\n\n");

}

else

{

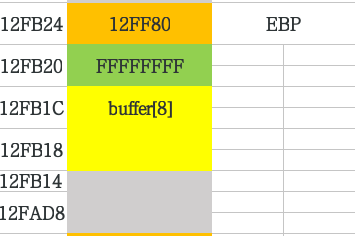
printf("Congratulation! \n");

break;

}

}

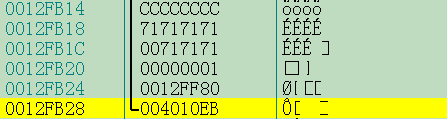
}



authenticated:12FB20

buffer[8]:12FB18-12FB1C

a.输入七个qqqqqqq(0x71)



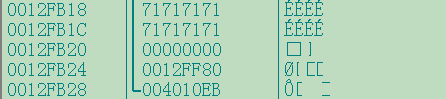
return address:004010EB

ebp:12FF80

authenticated:1

buffer[8]:71717171717171

b.输入八个qqqqqqqqq(0x71)



return address:004010EB

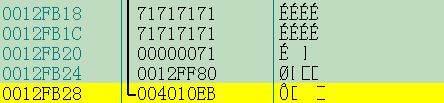
ebp:12FF80

authenticated:0

buffer[8]:7171717171717171

\*字符串末尾有“\0”

c.输入9个qqqqqqqqqq(0x71)



return address:004010EB

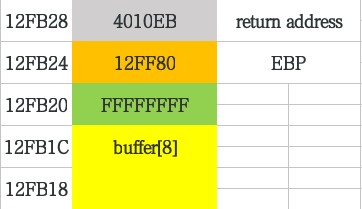
ebp:12FF80

authenticated:71

buffer[8]:7171717171717171

(2)修改函数返回地址

更强大的攻击通过缓冲区溢出改写目标往往不是一个变量，而是瞄准栈帧最下方的EBP与函数返回地址（return address）



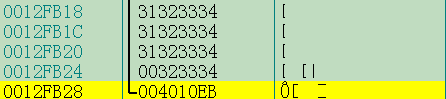
return address:004010EB

ebp:12FF80

authenticated:0

buffer[8]:0

a.输入432143214321432(0x31,0x32,0x33,0x34)覆盖EBP



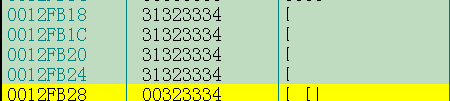
return address:004010EB

ebp:323334

authenticated:31323334

buffer[8]:3132333431323334

b.输入4321432143214321432,覆盖EBP,return addresss



return address:323334

ebp:31323334

authenticated:31323334

buffer[8]:3132333431323334



c.采用读取模式

#define PASSWORD "1234567“

int verify\_password (char \*password)

{

int authenticated;

char buffer[8];

authenticated=strcmp(password,PASSWORD);

strcpy(buffer,password);

return authenticated;

}

main()

{

int valid\_flag=0;

char password[1024];

while(1)

{

FILE \*fp;

if(!(fp=fopen(“password.txt”,”rw”)))exit(0);

fscanf()fp,”%s”,password);

valid\_flag = verify\_password(password);

if(valid\_flag)

{

printf("incorrect password!\n\n");

}

else

{

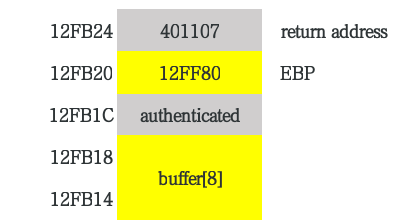
printf("Congratulation! \n");

break;

}

}

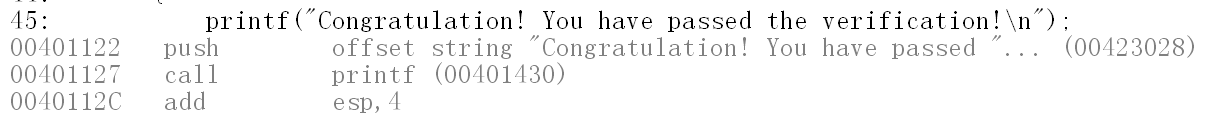
}



return address:00401107

ebp:12FF80

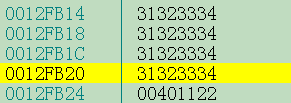
1）成功源码在内存的地址



2）password.txt（16进制）

4

3）栈



4）寄存器



(3)代码植入

利用栈溢出让经程执行输入数据中的植入代码

int verify\_password (char \*password)

{

int authenticated;

char buffer[44];

authenticated=strcmp(password,PASSWORD);

strcpy(buffer,password);//over flowed here!

return authenticated;

}

main()

{

int valid\_flag=0;

char password[1024];

FILE \* fp;

LoadLibrary("user32.dll");//prepare for messagebox

if(!(fp=fopen("password.txt","rw+")))

{

exit(0);

}

fscanf(fp,"%s",password);

valid\_flag = verify\_password(password);

if(valid\_flag)

{

printf("incorrect password!\n");

}

else

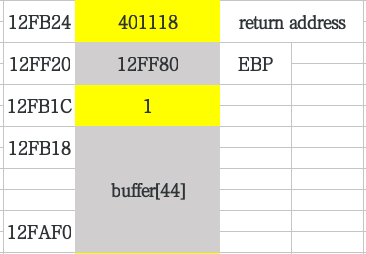
{

printf("Congratulation! You have passed the verification!\n");

}

fclose(fp);

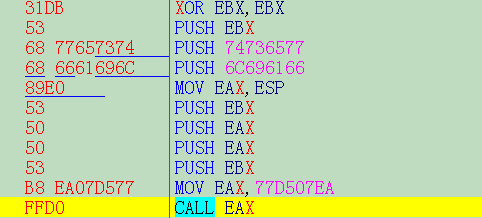
}



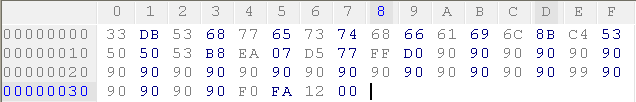
return address:401118

EBP:12FF80

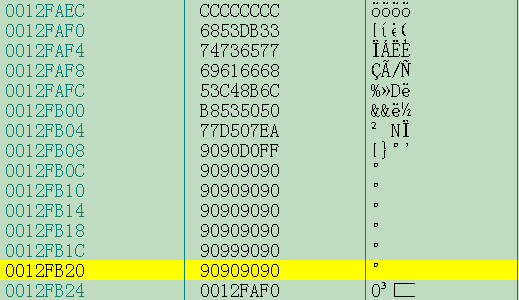
1)用到函数int messagebox(0,”str”,”str”,0);77d507ea函数地址



2)password.txt



3)寄存器



3.缓冲区溢出防范

 防范缓冲区溢出问题的准则是：确保做边界检查(通常不必担心影响程序效率)。不要为接收数据预留相对过小的缓冲区，大的数组应通过malloc/new分配堆空间来解决；在将数据读入或复制到目标缓冲区前，检查数据长度是否超过缓冲区空间。同样，检查以确保不会将过大的数据传递给别的程序，尤其是第三方COTS(Commercial-off-the-shelf)商用软件库——不要设想关于其他人软件行为的任何事情。

     若有可能，改用具备防止缓冲区溢出内置机制的高级语言(Java、C#等)。但许多语言依赖于C库，或具有关闭该保护特性的机制(为速度而牺牲安全性)。其次，可以借助某些底层系统机制或检测工具(如对C数组进行边界检查的编译器)。