

Vulnérabilités Applicatives - TP1 :

return into libc, débordements dans le tas et dans BSS

Objectifs : ce premier TP est destiné à mettre en évidence certaines classes de vulnérabilités applicatives vues en cours : les débordements dans la pile de type "return into libc", les attaques de type ROP, les débordements dans le tas et dans BSS. Nous les illustrerons au travers de divers exemples tirés du cours.

1 Return into libc

Nous allons dans cette partie étudier comment un débordement de tampon dans la pile peut être exploité même si la pile est non exécutable

Lorsque la pile est non exécutable, il devient impossible de copier dans la pile le code à exécuter. Il faut donc aller le chercher ailleurs ! La technique du `return into libc` permet de faire en sorte d'exécuter du code inclus dans la `libc` qui en principe est liée avec tous les programmes exécutables. Dans l'exemple que nous allons étudier ici, nous allons essayer d'exécuter un shell, ce qui est en général le but recherché par les attaquants, mais nous essaierons aussi d'autres codes exécutables.

1.1 Le programme de test

Soit le programme `tp1.c` :

```
#include <stdio.h>
#include <string.h>

void copie(char * s)
```

```
{
    char ch[8]="BBBBBBB";

    strcpy(ch,s);
}

int main(int argc, char * argv[])
{
    copie(argv[1]);
    return(0);
}
```

Pour compiler ce programme, nous utiliserons la commande suivante :

```
bash$ gcc -Wall -g -fno-stack-protector -mpreferred-stack-boundary=2
      tp1.c -o tp1
```

Et nous désactiverons ALSR (randomisation de l'espace d'adressage) :

```
bash# echo 0 > /proc/sys/kernel/randomize_va_space
```

Nous désactifions donc l'utilisation des canaries, de façon à pouvoir écrire dans la pile mais nous ne désactivons pas la protection qui interdit l'exécution de code dans la pile.

1.2 A la recherche de l'adresse de `system`

La fonction que nous allons exécuter lorsque nous aurons réalisé le débordement de tampon est la fonction `system`. Cette fonction, incluse dans la `libc` a la bonne idée de pouvoir exécuter tout binaire qu'on lui donne en paramètres. Ainsi `system("/bin/bash")` provoque le lancement d'un processus qui exécute le binaire `/bin/bash`.

Le premier problème que l'on doit résoudre est trouver l'adresse de cette fonction. Pour cela, on peut utiliser un debugger. La commande à utiliser sous `gdb` est simplement la commande : `p system`.

1. Compilez le programme source proposé avec les options proposées également.
2. A l'aide de `gdb`, déterminez l'adresse de la fonction `system` telle qu'elle est mappée dans votre programme exécutable.

1.3 A la recherche de la chaîne de caractères `/bin/bash` dans la pile

Comme nous l'avons vu en cours, le buffer que nous allons écraser devra contenir l'adresse de la fonction `system` mais aussi l'adresse de la chaîne de caractères `/bin/bash` (ou tout autre shell du même genre) quelques octets plus loin.

Pour déterminer l'adresse d'une telle chaîne dans l'espace d'adressage du processus, il y a deux solutions. Soit, la chercher dans les variables d'environnement (la variable `SHELL` contient une telle chaîne de caractères), soit la chercher dans la `libc` également.

Nous allons tout d'abord commencer par la première solution. Pour cela, il suffit de modifier le code source du programme C et d'ajouter les lignes :

```
char * ch=getenv("SHELL");  
printf("%x\n",ch);
```

au début du programme, pour connaître l'adresse en mémoire de la variable `SHELL`.

1. Modifiez votre programme source comme indiqué ci-dessus.
2. En déduire l'adresse de la chaîne `/bin/bash` dans les variables d'environnement.

1.4 A la recherche de la chaîne de caractères `/bin/sh` dans la `libc`

La `libc` contient la chaîne de caractères `/bin/sh`. Pour déterminer l'offset de cette chaîne dans notre programme binaire, deux simples commandes vont nous suffir. La commande `ldd` nous permet de connaître l'adresse dans notre binaire où est chargée la `libc`. La simple commande `ldd a.out` fournit ce résultat. Il faut ensuite déterminer l'offset de la chaîne `/bin/sh` dans la `libc`. Pour cela, il suffit d'utiliser la commande `strings` avec les bonnes options. Par exemple :

```
bash$ strings -t x /lib/tls/i686/cmov/libc.so.6 | grep /bin/sh
```

Pour obtenir l'adresse qui nous intéresse, il suffira de sommer les deux adresses trouvées.

1. Utilisez la commande `ldd` pour trouver l'offset de la `libc` dans votre programme binaire.

2. Utilisez la commande `strings` pour trouver l'offset de `/bin/sh` dans la `libc`.
3. En déduire l'adresse de la chaîne `/bin/sh` dans votre programme binaire.

1.5 L'exploitation

Il suffit maintenant de provoquer un débordement du buffer `ch` de la fonction `copie` de façon à écraser l'adresse de retour de cette fonction et à insérer 8 octets plus loin dans la pile l'adresse de la chaîne `/bin/sh`, ainsi que nous l'avons vu en cours. Nous utilisons ici la seconde méthode présentée ci-dessus, à l'aide de l'adresse de `/bin/sh` dans la `libc`.

Le langage `Perl` va nous être d'un grand secours sur ce point. Pour générer une chaîne de caractères composée de différents motifs et adresses, il suffit d'utiliser la commande suivante :

```
perl -e 'print "A" x 12 . "\xaa\xbb\xcc\xdd"'
```

Dans cet exemple, nous fabriquons une chaîne de caractères constituée de 12 caractères 'A' puis de l'adresse `0xddccbbaa` (attention à la notation inversée)

1. A l'aide de la commande ci-dessus, constituez la chaîne de caractères permettant de faire déborder le buffer `ch` de façon à provoquer la déroutement de la fonction `copie` vers la fonction `system` avec la chaîne `/bin/sh` en paramètres.
2. Exécutez votre programme exécutable en lui passant en paramètre cette chaîne.
3. Testez le fonctionnement.

2 Une ROP attaque

Nous allons à présent réutiliser le programme C vulnérable précédent mais nous allons l'exploiter en utilisant une attaque de type ROP. Nous allons donc pour cela devoir fabriquer une ROPchain.

De la même que pour l'exercice précédent, nous allons compiler le programme en désactivant les canaries et en désactivant la randomization de l'espace d'adressage. Le principe des attaques ROP étant de pouvoir fabriquer une ROPchain avec des gadgets trouvés dans les sections de code du binaire, nous allons également augmenter nos chances de trouver ces gadgets en compilant statiquement notre programme :

```
bash$ gcc -Wall -g -fno-stack-protector -mpreferred-stack-boundary=2
      tp1.c -o tp1 -static
bash# echo 0 > /proc/sys/kernel/randomize_va_space
```

Nous allons utiliser la méthode vue en cours nous permettant de chercher des gadgets réalisant `execve("/bin/sh",NULL,NULL)`. L'utilitaire `ROPgadget` va nous aider dans cette tâche. Cet utilitaire n'est installé que sur la machine hôte et non sur la VM. Vous devrez donc fabriquer la ROPchain sur la machine hôte.

1. À l'aide de l'utilitaire `ROPgadget`, trouvez les gadgets utiles à la copie en mémoire d'une chaîne de caractères choisie (telle que `"/bin/sh"`); pensez qu'il faut terminer une chaîne de caractères avec le caractère `NULL`
2. À l'aide de l'utilitaire `ROPgadget`, trouvez les gadgets utiles à l'initialisation de registres à 0 et à leurs incréments (en particulier pour `eax`)
3. Fabriquez ensuite la ROPchain
4. Recopiez la ROPchain sur la machine virtuelle et exécutez le programme vulnérable en lui fournissant un paramètre incluant cette ROPchain; constatez l'exploitation.

3 Débordement dans le tas

3.1 Les meta données du tas

Nous allons dans cette partie étudier la structure des données allouées dans le tas et les parties que l'on peut corrompre ainsi que leur conséquence. Nous n'allons donc pas exécuter d'exploit mais plutôt étudier comment est organisé le tas.

Soit le programme suivant :

```
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#define BUFSIZE 16

int main(int argc, char * argv[])
{
    char *buf1 = (char *)malloc(BUFSIZE);
```

```
char *buf2 = (char *)malloc(BUFSIZE);

memset(buf1, 'A', BUFSIZE-1);
buf1[BUFSIZE-1] = '\0';
memset(buf2, 'B', BUFSIZE-1);
buf2[BUFSIZE-1] = '\0';

strcpy(buf1, argv[1]);

free(buf1);
free(buf2);

return 0;
}
```

1. Modifiez ce programme C de façon à faire apparaître la structure des chunks en mémoire dynamique. En particulier, vous afficherez la “vraie” taille allouée pour `buf1` et `buf2`.
2. Provoquez maintenant un débordement de `buf1` de façon à écraser la zone mémoire où est stockée la “vraie” taille de `buf2` (on ne cherchera donc pas à écraser `buf2` lui-même). Quel est le message d’erreur obtenu à l’exécution ? Que signifie-t-il selon vous ? Commentez l’instruction qui provoque ce message ? Commentez tour à tour les instructions `free` et constatez la différence des messages d’erreurs.
3. Pour finir, provoquez l’écrasement de `buf2`.

3.2 Exploitation de la macro `unlink`

Nous allons dans cette section exploiter la macro `unlink` en faisant exécuter un shellcode à un programme vulnérable. Cet exploit nécessite l’utilisation d’une ancienne `libc` et donc d’une machine virtuelle particulière. Vous devrez donc utiliser le script `start_slack` pour lancer cette VM. Le programme vulnérable est dans le dossier `secu5` du compte `debug`. L’enseignant vous fournira le mot de passe à utiliser.

Un squelette d’exploit vous est fourni dans le dossier et il consiste à utiliser la méthode vu dans les slides du cours. Le shellcode contenant un saut vous est également fourni. Le principe consiste donc à écraser les données d’un chunk alloué légitimement puis à déborder de ce chunk pour écraser les meta données du chunk suivant, alloué lui aussi légitimement.

3.3 Forged chunk

Il est possible de profiter de la structure du tas et des meta données (notamment des pointeurs) pour détourner l'exécution des fonctions malloc et free. Nous allons dans ce dernier exercice coder un programme qui utilise la technique du chunk forgé dans la pile et du détournement de la fonction malloc qui va renvoyer une adresse dans la pile suit à l'écrasement de meta données d'un chunk légitimement alloué puis libéré. Nous devez donc utiliser la méthode vu en cours et écrire un petit programme qui la met en pratique. Réalisez cet exercice sur la première machine virtuelle utilisée dans ce TP.

4 Débordement dans BSS

Nous allons, dans cette partie, étudier comment un débordement dans certaines sections de mémoire telles que la section `bss` peut être exploitée. Nous allons pour cela utiliser l'exemple de programme vulnérable que nous avons vu en cours.

4.1 Le programme vulnérable

Soit le programme suivant :

```
#include <stdio.h>
#include <stdlib.h>
#include <unistd.h>
#include <string.h>

#define ERROR -1
#define BUFSIZE 8

int goodfunc(const char *str);

int main(int argc, char **argv)
{
    static int (*funcptr)(const char *str);
    static char buf[BUFSIZE];

    funcptr = (int (*)(const char *str))goodfunc;
    printf("Avant overflow: funcptr pointe sur %p\n", funcptr);
```

```
memset(buf, 0, sizeof(buf));
strcpy(buf, argv[1]);
printf("Après overflow: funcptr pointe sur %p\n", funcptr);

(void)(*funcptr)(argv[2]);
return 0;
}

int goodfunc(const char *str)
{
    printf("Goodfunc, parametre : %s\n", str);
    return 0;
}
```

4.2 Examen de la section bss

Les 2 variables `static buf` et `funcptr` se trouvent dans la section `bss` ainsi que nous permettent de le vérifier les commandes `nm` et `objdump`.

La commande `nm` permet notamment de lister les différentes sections d'un programme binaire et d'en donner l'adresse. Ainsi `nm a.out | grep bss` nous permet d'obtenir l'adresse du début de la section `bss`.

La commande `objdump` va nous permettre de désassembler la section `bss` de façon à connaître l'adresse exacte des variables de cette section. La commande à utiliser est :

```
objdump -R -D --start-address <adresse_fournie_par_nm> a.out
```

1. A l'aide des deux commandes ci-dessus, déterminez les adresses mémoire de `buf` et de `funcptr`.
2. En déduire la distance séparant ces 2 variables dans `bss`.
3. En déduire la taille nécessaire de `argv[1]` pour provoquer l'écrasement de `funcptr` lors de l'appel de la fonction `strcpy`.

4.3 L'exploitation

L'exploitation consiste donc à détourner l'utilisation du pointeur de fonction de façon à le faire pointer sur une fonction choisie par l'attaquant, par exemple, ici encore, la fonction `system`. Il suffira ensuite de fournir en second paramètre (`argv[2]`) la chaîne de caractères `sh` par exemple pour obtenir un shell.

1. Déterminez l'adresse à laquelle la fonction `system` est mappée dans le programme binaire.
2. En déduire les deux arguments à passer au programme pour obtenir un shell (on utilisera `perl` pour fabriquer le premier argument).

5 Débordement dans BSS pour modifier la GOT

Nous allons dans cette dernière partie utiliser l'exemple vu en cours pour profiter d'un débordement dans BSS afin de détourner l'exécution d'une fonction standard de la `libc`.

5.1 Le programme vulnérable

Le programme est le suivant. Il utilise 2 copies de chaînes de caractères passées en paramètres. La seconde copie se fait par l'intermédiaire d'un pointeur que nous allons écraser.

```
#include <string.h>
#include <stdio.h>
#include <stdlib.h>

int main(int argc, char * argv[])
{
    static char * ptr;
    static char buf2[16];
    static char buf1[16];

    printf("buf1 : %p - buf2 : %p - ptr : %p\n",buf1,buf2,&ptr);
    ptr=buf2;
    if (argc < 3) exit(-1);

    strcpy(buf1,argv[1]);
    strcpy(ptr,argv[2]);
    printf("%s\n",buf2);
    return(0);
}
```

5.2 Le principe de l'exploitation

Le but est de profiter du premier appel à `strcpy` pour provoquer un débordement de `buf1` dans BSS et ainsi écraser `buf2` mais aussi `ptr`. Ainsi, nous pourrions maîtriser l'adresse de destination du second `strcpy`. Comme nous maîtrisons aussi la source du second `strcpy` (qui est `argv[2]`), nous pouvons écrire ce que nous voulons là où nous le voulons en quelque sorte.

5.3 Premier strcpy

Il est tout d'abord nécessaire de vérifier que le débordement de `buf1` nous permet bien d'écraser `ptr`. Ensuite, il faut calculer le nombre d'octets séparant `buf1` de `ptr` de façon à pouvoir l'écraser. Il faut ensuite déterminer l'adresse de la fonction `puts` dans la GOT. Enfin, il faut fabriquer `argv[1]` de la façon suivante :

```
NNNNNNNNSSSSSSSSSSSS[adr_got]
```

où `N` est l'instruction NOP, `SSSSSS` représente le shellcode et `adr_got` l'adresse de `puts` dans la GOT. Il faut calculer le tout pour que cette adresse écrase bien la valeur de `ptr`.

1. Compilez le programme précédent.
2. Vérifiez à l'aide d'un debugger ou par simple affichage que `buf1` précède bien `ptr` en mémoire.
3. Il est également important de calculer la longueur de la chaîne `argv[1]` que nous devons utiliser pour écraser `ptr`. Calculez cette distance.
4. Il reste à présent à déterminer l'adresse dans la GOT de la fonction `puts` (il faut en fait utiliser `puts` et non `printf` du fait que nous affichons seulement une chaîne de caractères, dans ce cas, c'est en fait `puts` qui est utilisé). Pour cela, utilisez la commande `objdump -R`.

5.4 Second strcpy et exploitation

Le second `strcpy` nous permet de modifier la GOT (plus précisément, l'indirection correspondant à `puts`) de façon à modifier cette indirection par l'adresse de notre shellcode.

1. Faites en sorte que votre programme affiche l'adresse de `buf1` (c'est cette adresse qui nous servira à écraser la GOT).

2. Exécutez le programme en lui passant en premier paramètre la chaîne calculée dans la section précédente et en second une chaîne contenant l'adresse de `buf1` dans BSS (le shellcode vous sera fourni par l'enseignant).
3. Vérifiez que vous arrivez bien à exécuter un shell.