Memory Safety

Question 1 ***Software Vulnerabilities*** ()

For the following code, assume an attacker can control the value of `basket` passed into `search_basket`. The value of `n` is constrained to correctly reflect the number of cats in `basket`.

The code includes several security vulnerabilities. **Circle *three* such vulnerabilities** in the code and **briefly explain** each of the three on the next page.

```
1  struct cat {
2      char name[1024];
3      int age;
4  };
5
6  /* Searches through a basket of cats of size at most 32.
7     Returns the number of kittens or -1 if there are no kittens */
8  int search_basket(struct cat basket[], size_t n) {
9      struct cat kittens[32];
10     char kitten_names[1024], cmd[1024];
11     int i, total_kittens = 0, names_size = 0;
12
13     if (n > 32) return -1;
14
15     for (i = 0; i <= n; i++) {
16         if (basket[i].age < 12) {
17             size_t len = strlen(basket[i].name);
18             snprintf(kitten_names + names_size, len, "%s ", basket[i].name);
19             kittens[total_kittens] = basket[i];
20             names_size += len;
21             total_kittens += 1;
22         }
23     }
24
25     if (total_kittens > 5) {
26             const char *fmt = "adopt-kittens --num_kittens %d --names %s";
27             snprintf(cmd, sizeof cmd, fmt, total_kittens, kitten_names);
28             system(cmd);
29     }
30     return total_kittens;
31 }
```

Reminders:

- **snprintf(buf, len, fmt, ...)** works like **printf**, but instead writes to **buf**, and won't write more than **len - 1** characters. It terminates the characters written with a '\0'.
- **system** runs the shell command given by its first argument.

1. Explanation:

> **Solution:** Line **15** has a fencepost error: the conditional test should be $i < n$ rather than $i <= n$. The test at line **13** assures that **n** doesn't exceed 32, but if it's equal to 32, and if all of the cats in **basket** are kittens, then the assignment at line **19** will write past the end of **kittens**, representing a buffer overflow vulnerability.

2. Explanation:

> **Solution:** At line **18**, there's an error in that the length passed to **snprintf** is *supposed* to be available space in the buffer, but instead it's the length of the string being copied (along with a blank) into the buffer. Therefore by supplying large names for cats in **basket**, the attacker can write past the end of **kitten_names**, again representing a buffer overflow vulnerability.

3. Explanation:

> **Solution:** At line **28**, a shell command is run based on the contents of **cmd**, which in turn includes values from **kitten_names**, which in turn is derived from input provided by the attacker. That input could include shell command characters such as pipes ('|') or command separators (';'), facilitating *command injection*.

> **Solution:** Some more minor issues concern the **name** strings in **basket** possibly not being correctly terminated with '\0' characters, which could lead to reading of memory outside of **basket** at line **17** or line **18**.
>
> Note that there are no issues with format string vulnerabilities at any of lines **18**, **26**, or **27**. For each of those, the format itself does not include any elements under the control of the attacker.

Describe how an attacker could exploit these vulnerabilities to obtain a shell:

> **Solution:** Each vulnerability could lead to code execution. The easiest method would be to use the **cat.name** field for command injection ie. make the last kitten have name ";/bin/sh". An attacker could also use the fencepost or the bound-checking error to overwrite the rip and execute arbitrary code.

Question 2 ***C Memory Defenses*** ()

Mark the following statements as True or False and justify your solution. Please feel free to discuss with students around you.

1. Stack canaries completely prevent a buffer overflow from overwriting the return instruction pointer.

   Solution:

   False, stack canaries can be defeated if they are revealed by information leakage, or if there is not sufficient entropy, in which case an attacker can guess the value. Also, format string vulnerabilities can simply skip past the canary.

2. A format-string vulnerability can allow an attacker to overwrite values below the stack pointer

   Solution:

   True, format string vulnerabilities can write to arbitrary addresses by using a '%n' in junction with a pointer.

3. An attacker exploits a buffer overflow to redirect program execution to their input. This attack no longer works if the data execution prevention/executable space protection/NX bit is set.

   Solution:

   True, the definition of the NX bit is that it prevents code from being writable and executable at the same time. An attacker who can write code into memory cannot execute it.

4. If you have a non-executable stack and heap, buffer overflows are no longer exploitable.

   Solution:

   False. Many attacks rely on writing malicious code to memory and then executing them. If we make writable parts of memory non-executable, these attacks cannot succeed. However there are other types of attacks which still work in these cases, such as Return Oriented Programming.

5. If you use a memory-safe language, some buffer overflow attacks are still possible.

Solution:

False, buffer overflow attacks do not work with memory safe languages.

6. ASLR, stack canaries, and NX bits all combined are insufficient to prevent exploitation of all buffer overflow attacks.

Solution:

True, all of these protections can be overcome.

Short answer!

1. What vulnerability would arise if the canary was above the return address?

Solution:

It doesn't stop an attacker from overwriting the return address. Although if an attacker had absolutely no idea where the return address was, it could potentially detect stack smashing.

2. What vulnerability would arise if the stack canary was between the return address and the saved frame pointer?

Solution:

An attacker can overwrite the saved frame pointer so that the program uses the wrong address as the base pointer after it returns. This can be turned into an exploit.

3. Assume ASLR is enabled. What vulnerability would arise if the instruction **jmp *esp** exists in memory?

Solution:

An attacker can overwrite the return instruction pointer with the address of this command. This will cause the function to execute the instruction one word before the rip. An attacker could place the shellcode after the rip, and have the word before the rip contain a JMP command two words forward.

Question 3 ***TCB (Trusted Computing Base)*** ()

In lecture, we discussed the importance of a TCB and the thought that goes into designing it. Answer these following questions about the TCB:

1. What is a TCB?
2. What can we do to reduce the size of the TCB?
3. What components are included in the (physical analog of) TCB for the following security goals:
   (a) Preventing break-ins to your apartment
   (b) Locking up your bike
   (c) Preventing people from riding BART for free
   (d) Making sure no explosives are present on an airplane

Solution:

1. It is the set of hardware and software on which we depend for correct enforcement of policy. If part of the TCB is incorrect, the system's security properties can no longer be guaranteed to be true. Anything outside the TCB isn't relied upon in any way.
2. Privilege separation and separation of responsibility can help reduce the size of the TCB. You will end up with more components, but not all of them can violate your security goals if they break. The size of the TCB can also be reduced by reducing the application's dependency on third-party components and software.
3. (This list is not necessarily complete)
   (a) the lock, the door, the walls, the windows, the roof, the floor, you, anyone who has a key
   (b) the bike frame, the bike lock, the post you lock it to, the ground
   (c) the ticket machines, the tickets, the turnstiles, the entrances, the employees
   (d) the TSA employees, the security gates, the "one-way" exit gates, the fences surrounding the runway area

Question 4 ***Memory Safetyyy*** ()

The following code runs on a 32-bit x86 system. All memory safety defenses are disabled.

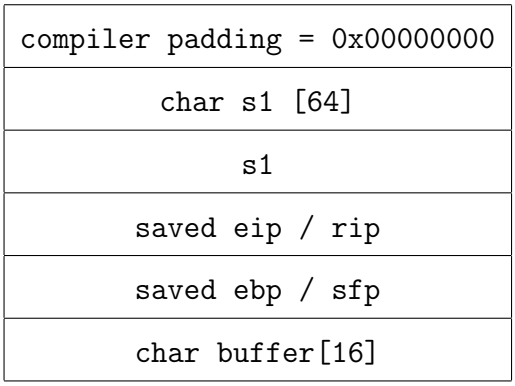
The compiler does not rearrange stack variables.

```
1  void doit(char* s1) {
2      char buffer[16];
3      strcpy(buffer, s1);
4      printf("%s\n%s\n", buffer, s1);
5  }
6
7  int main() {
8      char s1[64];
9      fgets(s1, sizeof s1, stdin);
10     doit(s1);
11 }
```

(a) Which line contains a memory safety vulnerability? What is the name of the vulnerability present on that line?

Solution: Line 3: buffer overflow.

(b) Complete the diagram of the stack, right before line 3. Assume normal (non-malicious) program execution. You do not need to write the values on the stack, only the names. There are no extraneous boxes. As in discussion, the bottom of the page represents the lower addresses.



(c) Now we will exploit the program. There is already shellcode at the address `0xbfffdead`. Complete the Python script below in order to successfully exploit the program.

NOTE: The Python syntax `'A' * n` indicates that the character `'A'` will be repeated `n` times. The syntax `\xRS` indicates a byte with hex value `0xRS`.

```
s1 = 'A' * ____ + '________________________________________________' + \
                  '________________________________________________'
print s1
```

Solution:

```
s1 = 'A' * 20 + '\xad\xde\xff\xbf'
```