

CSCI 476: Computer Security

Hashing (Part 3)

Reese Pearsall
Spring 2023

Announcements

Lab 8 due Wednesday April 19th

Research Project due April 23rd

Lab 9 due Sunday April 30th

Final Lab will be posted sometime after the research project

→ Due during finals week

→ You can earn extra credit for the final lab by attending lecture on 4/24, 4/26, 4/28, 5/1

CRC32

MD5

SHA-256

Doctor's prescription note

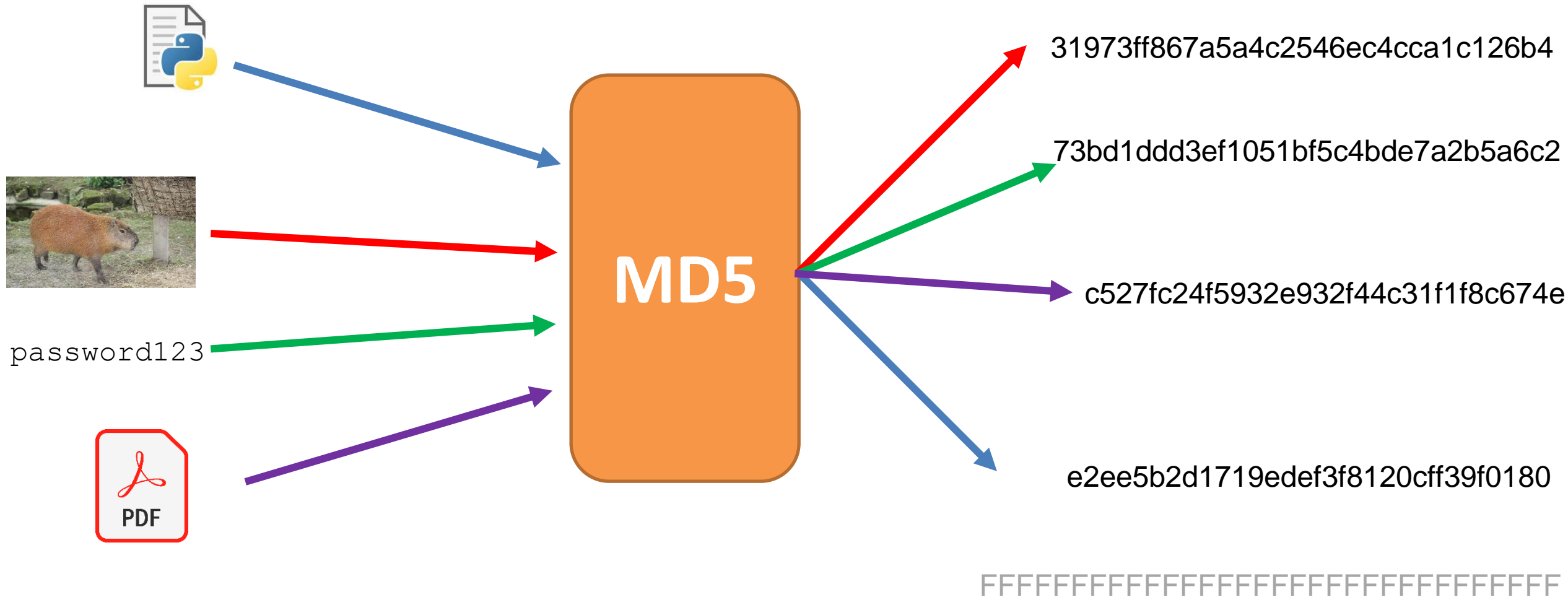
You & Yours are
please at your
Use your signature in
the letter long at
in image
the place at



Applications of Hashing

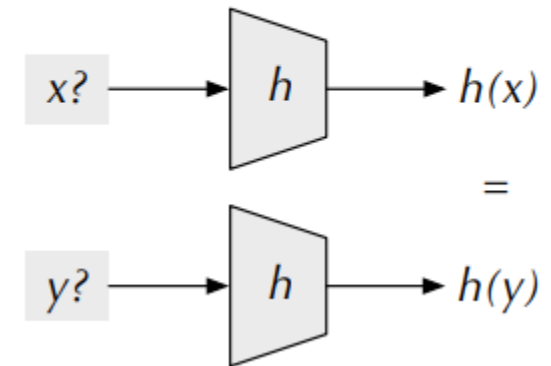
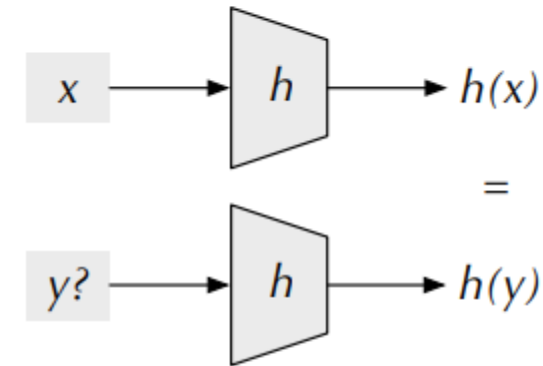
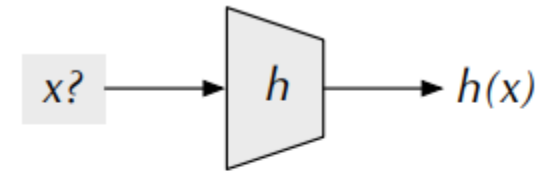
Output space of MD5 (128 bits)

000



Hash Functions Properties

- **Preimage Resistance ("One-Way")**
Given $h(x) = z$, hard to find x
(or any input that hashes to z for that matter)
- **Second Preimage Resistance**
Given x and $h(x)$, hard to find y s.t. $h(x) = h(y)$
- **Collision Resistance (or, ideally, "Collision Free")**
Difficult to find x and y s.t. $hash(x) = hash(y)$



Hash Collisions

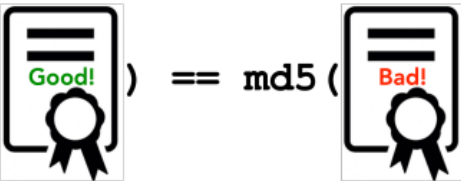
Goal: Create two **different files** with the **same md5 hash**

Our **ultimate goal** would be to create two executables (one benign, one malicious) with the same hash
(This is difficult to do, but we will show that it can theoretically happen)

Motivation

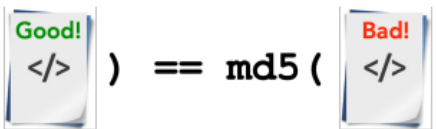
Forging public-key certificates

- Assume two certificate requests for www.example.com and www.attacker.com have same hash due to a collision
- CA signing of either request would be equivalent
- Attacker can get certificate signed for www.example.com without owning it!

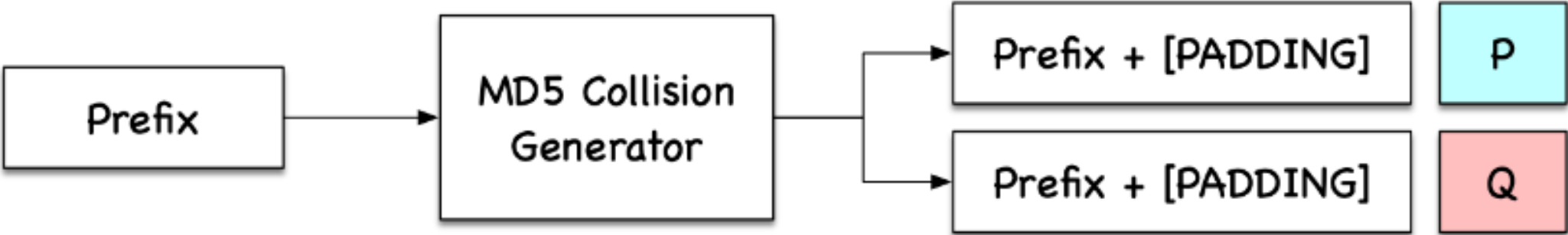
$$\text{md5} \left(\text{Good!} \right) == \text{md5} \left(\text{Bad!} \right)$$


Integrity of Programs

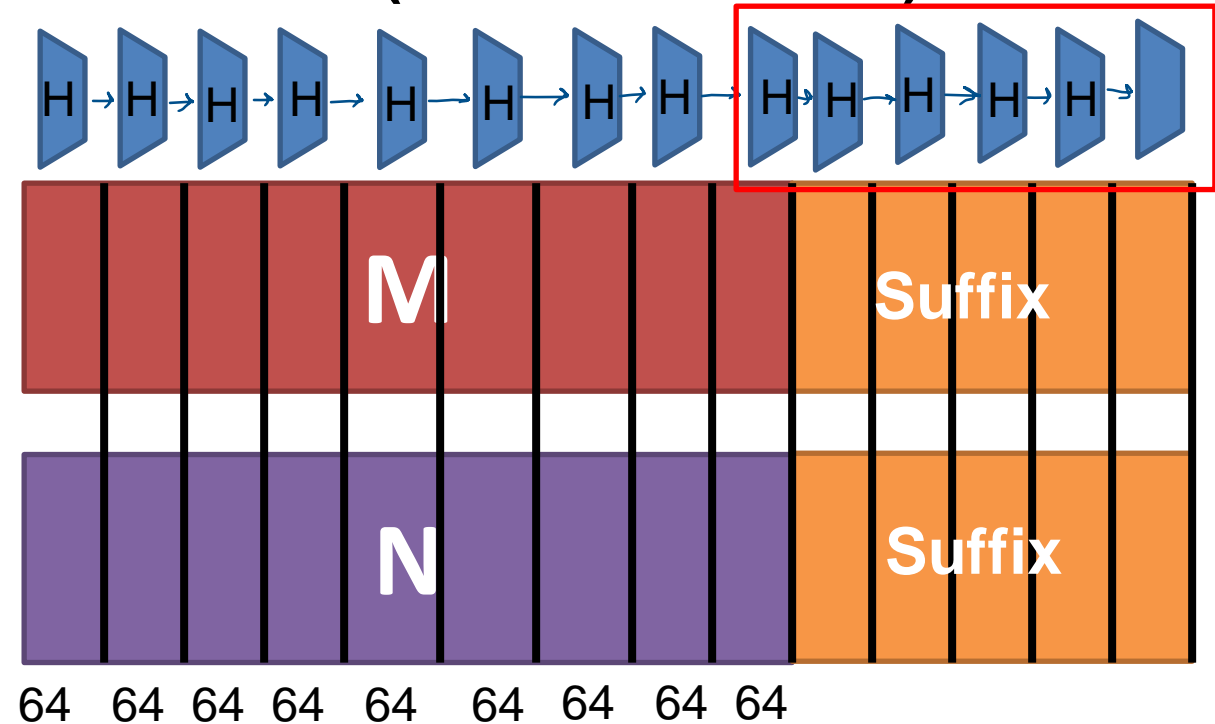
- Ask CA to sign a legitimate program's hash
- Attacker creates a malicious program with same hash
- The certificate for legitimate program is also valid for malicious version

$$\text{md5} \left(\text{Good!} \right) == \text{md5} \left(\text{Bad!} \right)$$


Hash Collisions (MD5collgen)



Hash Collisions (Suffix Extension)



If we append the same suffix, then this computation will also be the exact same for M and N

```
[11/17/22] seed@VM: ~/.../07_hash$ echo "suffix" > suffix.txt
[11/17/22] seed@VM: ~/.../07_hash$ cat out1.bin suffix.txt > out1suffix.bin
[11/17/22] seed@VM: ~/.../07_hash$ cat out2.bin suffix.txt > out2suffix.bin
```

$$H(m) == H(n)$$

$$H(m || s) == H(n || s) \quad s = \text{shared suffix}$$

```
[11/17/22] seed@VM: ~/.../07_hash$ md5sum out1suffix.bin
a63075af11518048cff11bf3d11a5462 out1suffix.bin
[11/17/22] seed@VM: ~/.../07_hash$ md5sum out2suffix.bin
a63075af11518048cff11bf3d11a5462 _out2suffix.bin
```

Hash Collisions (Generating Two executable files with the same MD5 hash)

```
[11/17/22]seed@VM:~/.../07_hash$ cat print_array.c
```

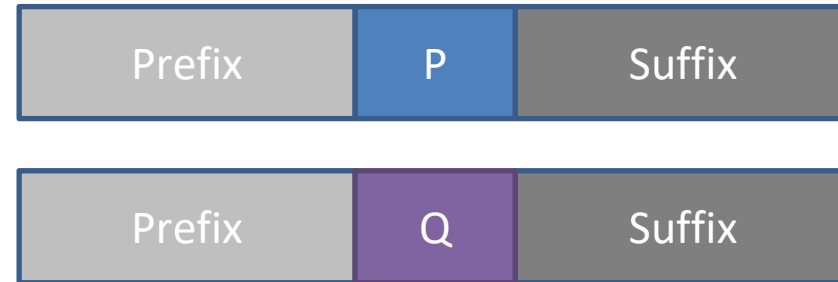
```
#include <stdio.h>
```

[illegible]

```
int main()
{
    int i;
    for (i=0; i<200; i++){
        printf("%x", xyz[i]);
    }
    printf("\n");
}
```

This is a program that will print out the contents of an array

We will create two variants of this program, but the program will have the same hash




```
[11/17/22] seed@VM: ~/ /07 hash$ cat print_array.c
```

Prefix

P

Suffix

```
int main()
{
    int i;
    for (i=0; i<200; i++){
        printf("%x", xyz[i]);
    }
    printf("\n");
}
```

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

Prefix

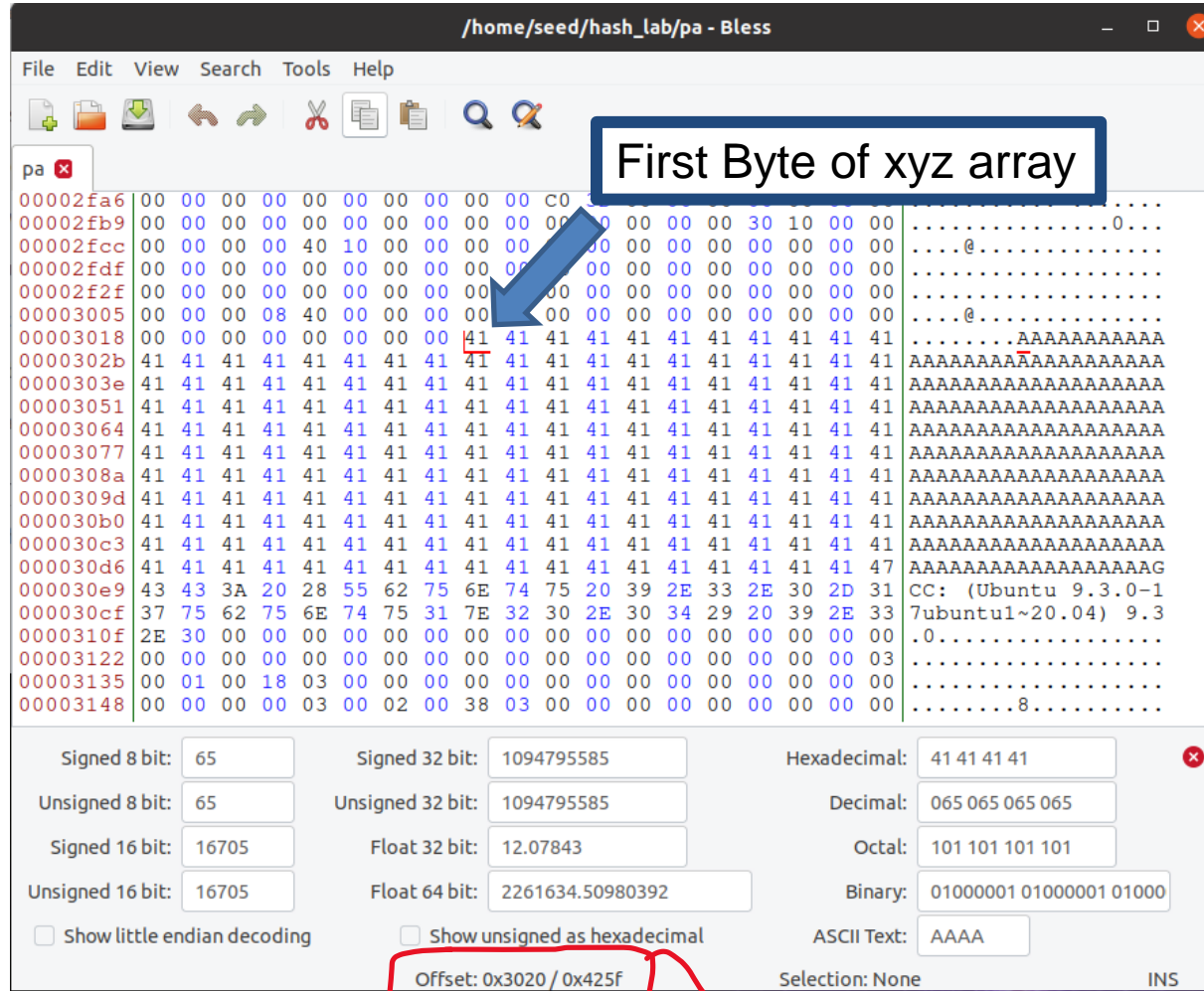
1,1G

Suffix

```
int main()
{
    int i;
    for (i=0; i<200; i++){
        printf("%x", xyz[i]);
    }
    printf("\n");
}
```

Hash Collisions (Generating Two executable files with the same MD5 hash but behave very differently)

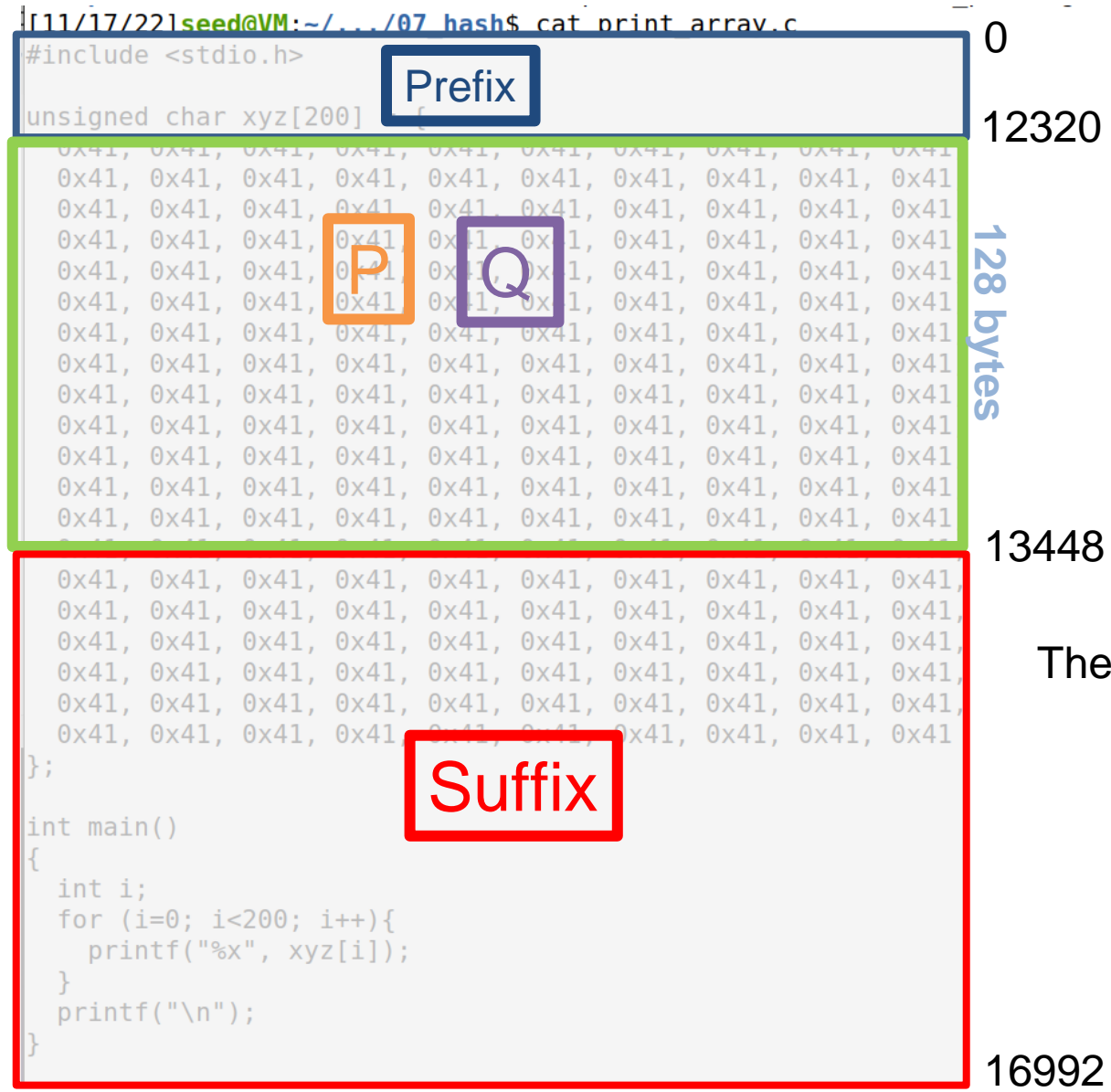
```
[11/17/22] seed@VM:~/hash_lab$ gcc print_array.c -o pa
[11/17/22] seed@VM:~/hash_lab$ bless pa
```



We can find where xyz begins in our program easily, because we filled it with A's

Start of XYZ = 0x3020 (Hexadecimal)
12320 (decimal)

Task 4 on the lab



Our prefix will be bytes 0-12320 of the program!

We want our **P** and **Q** to be 128 bytes

- Why 128?
- Multiple of 64
 - Wont overflow an array of size 200

Therefore, our suffix will begin at byte # $12320 + 128 = 13448$

Task 4 on the lab

```
[11/17/22]seed@VM:~/.../07_hash$ cat print_array.c
```

Prefix

P

Q

Suffix

0
12320

128 bytes

13448

16992 (size of executable)

①

Get contents of prefix and suffix

```
[11/17/22]seed@VM:~/hash_lab$ head -c 12320 pa > prefix  
[11/17/22]seed@VM:~/hash_lab$ tail -c +12448 pa > suffix
```

②

Use collision tool to get (prefix + P) and (prefix + Q)

```
[11/17/22]seed@VM:~/hash_lab$ md5collgen -p prefix -o prefix_and_P prefix_and_Q  
MD5 collision generator v1.5  
by Marc Stevens (http://www.win.tue.nl/hashclash/)  
  
Using output filenames: 'prefix_and_P' and 'prefix_and_Q'  
Using prefixfile: 'prefix'  
Using initial value: fa3f7a62525b9c90471862a4a04139a5  
  
Generating first block: ..  
Generating second block: S01..  
Running time: 1.78726 s
```

③

Add suffix to programs

```
[11/17/22]seed@VM:~/hash_lab$ cat prefix_and_P suffix > program1.out  
[11/17/22]seed@VM:~/hash_lab$ cat prefix_and_Q suffix > program2.out
```

④

Verify that executables are different, but have the same hash

```
[11/17/22]seed@VM:~/hash_lab$ diff program1.out program2.out  
Binary files program1.out and program2.out differ  
[11/17/22]seed@VM:~/hash_lab$ md5sum program1.out  
f489a326ed9c692f31eabccab06062ce program1.out  
[11/17/22]seed@VM:~/hash_lab$ md5sum program2.out  
f489a326ed9c692f31eabccab06062ce program2.out
```



Task 4 on the lab

```
[11/17/22] seed@VM:~/.../07 hash$ cat print_array.c
```

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

Prefix

```
unsigned char xyz[200]
```

A 10x10 grid of hexadecimal values 0x41. A 3x3 orange box labeled 'P' is centered at row 4, column 4. A 3x3 purple box labeled 'Q' is centered at row 5, column 5.

[illegible]

Suffix

```
};

int main()
{
    int i;
    for (i=0; i<200; i++){
        printf("%x", xyz[i]);
    }
    printf("\n");
}
```

0
12320

128 bytes

13448

16992 (size of executable)

⑤

Make sure you still have a valid program 😊

```
[11/17/22] seed@VM:~/hash_lab$ ./program1.out
```

[illegible]

```
[11/17/22] seed@VM:~/hash_lab$ ./program2.out
```

[illegible]

```
[11/17/22] seed@VM:~/hash_lab$
```

Somewhere in this output, you should find a small difference

Task 4 on the lab

```
[11/17/22] seed@VM:~/.../07 hash$ cat print_array.c
```

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

Prefix

```
unsigned char xyz[200]
```

A 10x10 grid of hexagonal color codes (0x41) with a red 'P' and a blue 'Q' overlaid.

[illegible]

Suffix

```
};

int main()
{
    int i;
    for (i=0; i<200; i++){
        printf("%x", xyz[i]);
    }
    printf("\n");
}
```

0
12320

128 bytes

13448

16992 (size of executable)

⑤

Make sure you still have a valid program 😊

```
[11/17/22] seed@VM:~/hash_lab$ ./program1.out
```

[illegible]

```
[11/17/22] seed@VM:~/hash_lab$ ./program2.out
```

[illegible]

```
[11/17/22] seed@VM:~/hash_lab$
```

*These programs print out different things,
which is very benign*

Our next goal is to write two programs with the same MD5 hash, but one does something malicious, and the other does something benign

```

#include <stdio.h>
#define LENGTH 400

unsigned char X[LENGTH]= {
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
};

unsigned char Y[LENGTH]= {
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
};

int main()
{
    int i = 0;

    for (i = 0; i < LENGTH; i++){
        if (X[i] != Y[i]) break;
    }
    if (i == LENGTH){
        printf("%s\n", "Executing benign code... ");
    }
    else {
        printf("%s\n", "Executing malicious code... ");
    }
    return 0;
}

```

This program has two arrays X and Y

The program compares the contents of these two arrays

```

#include <stdio.h>
#define LENGTH 400

unsigned char X[LENGTH]= {
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
},

unsigned char Y[LENGTH]= {
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
},

int main()
{
    int i = 0;

    for (i = 0; i < LENGTH; i++){
        if (X[i] != Y[i]) break;
    }
    if (i == LENGTH){
        printf("%s\n", "Executing benign code... ");
    }
    else {
        printf("%s\n", "Executing malicious code... ");
    }
    return 0;
}

```

This program has two arrays X and Y

The program compares the contents of these two arrays

If the two arrays are the same, then it will execute the benign code


```

#include <stdio.h>
#define LENGTH 400

unsigned char X[LENGTH]= {
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
},

unsigned char Y[LENGTH]= {
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
},

int main()
{
    int i = 0;
    for (i = 0; i < LENGTH; i++){
        if (X[i] != Y[i]) break;
    }
    if (i == LENGTH){
        printf("%s\n", "Executing benign code... ");
    }
    else {
        printf("%s\n", "Executing malicious code... ");
    }
    return 0;
}

```

This program has two arrays X and Y

The program compares the contents of these two arrays

If the two arrays are the same, then it will execute the benign code

If the two arrays are different, then it will execute the benign code

```

#include <stdio.h>
#define LENGTH 400

unsigned char X[LENGTH]= {
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
},

unsigned char Y[LENGTH]= {
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
},

int main()
{
    int i = 0;

    for (i =0; i< LENGTH; i++){
        if (X[i] != Y[i]) break;
    }
    if (i==LENGTH){
        printf("%s\n", "Executing benign code... ");
    }
    else {
        printf("%s\n", "Executing malicious code... ");
    }
    return 0;
}

```

```

#include <stdio.h>
#define LENGTH 400

unsigned char X[LENGTH]= {
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
},

unsigned char Y[LENGTH]= {
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
},

int main()
{
    int i = 0;

    for (i =0; i< LENGTH; i++){
        if (X[i] != Y[i]) break;
    }
    if (i==LENGTH){
        printf("%s\n", "Executing benign code... ");
    }
    else {
        printf("%s\n", "Executing malicious code... ");
    }
    return 0;
}

```

Goal: Generate two variants of this program. One variant where the two arrays are the same (benign version), and one variant where the two arrays are different (malicious version)

19

Prefix

```
#include <stdio.h>
#define LENGTH 400
```

```
unsigned char X[LENGTH]= {
```

P

suffix1

```
unsigned char Y[LENGTH]= {
```

Q

suffix2

```
int main()
{
    int i = 0;
    for (i =0; i< LENGTH; i++){
        if (X[i] != Y[i]) break;
    }
    if (i==LENGTH){
        printf("%s\n", "Executing benign code... ");
    }
    else {
        printf("%s\n", "Executing malicious code... ");
    }
    return 0;
}
```

Prefix

```
#include <stdio.h>
#define LENGTH 400
```

```
unsigned char X[LENGTH]= {
```

P

suffix1

```
unsigned char Y[LENGTH]= {
```

P

suffix2

```
int main()
{
    int i = 0;
    for (i =0; i< LENGTH; i++){
        if (X[i] != Y[i]) break;
    }
    if (i==LENGTH){
        printf("%s\n", "Executing benign code... ");
    }
    else {
        printf("%s\n", "Executing malicious code... ");
    }
    return 0;
}
```

Because we are inserting P/Q at two different points into our program, we will have two suffixes

```
#include <stdio.h>
#define LENGTH 400
```

Prefix

```
unsigned char X[LENGTH]= {
```

```
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
```

P

```
unsigned char Y[LENGTH]= {
```

suffix1

```
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
```

Q

```
int main()
{
    int i = 0;

    for (i = 0; i < LENGTH; i++){
        if (X[i] != Y[i]) break;
    }
    if (i == LENGTH){
        printf("%s\n", "Executing benign code... ");
    }
    else {
        printf("%s\n", "Executing malicious code... ");
    }
    return 0;
}
```

suffix2

First, we must get the starting location of our array X

The screenshot shows a hex editor window titled "/home/seed/csc476-code/09_hashing/be - Bless". The main area displays a memory dump with addresses from 00002e61 to 0000329e. The data is shown in hexadecimal and ASCII. At the bottom, there is a conversion table with the following values:

Signed 8 bit:	65	Signed 32 bit:	1094795585	Hexadecimal:	41 41 41 41
Unsigned 8 bit:	65	Unsigned 32 bit:	1094795585	Decimal:	065 065 065 065
Signed 16 bit:	16705	Float 32 bit:	12.07843	Octal:	101 101 101 101
Unsigned 16 bit:	16705	Float 64 bit:	2261634.50980392	Binary:	01000001 01000001 01000001 01000001

Additional options at the bottom include "Show little endian decoding", "Show unsigned as hexadecimal", "ASCII Text: AAAAA", "Offset: 12320 / 17583", "Selection: None", and "INS".

```
[04/17/23]seed@VM:~/.../09_hashing$ gcc benign_evil.c -o be
[04/17/23]seed@VM:~/.../09_hashing$ bless be
```



```
#include <stdio.h>
#define LENGTH 400
```

Prefix

```
unsigned char X[LENGTH]= {
```

```
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
```

suffix1

```
unsigned char Y[LENGTH]= {
```

```
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
```

suffix2

```
int main()
{
    int i = 0;

    for (i = 0; i < LENGTH; i++){
        if (X[i] != Y[i]) break;
    }
    if (i == LENGTH){
        printf("%s\n", "Executing benign code... ");
    }
    else {
        printf("%s\n", "Executing malicious code... ");
    }
    return 0;
}
```

First, we must get the starting location of our array X

The screenshot shows a hex editor window titled "/home/seed/csc1476-code/09_hashing/be - Bless". The main area displays a memory dump with addresses from 00002e61 to 0000329e. The data column shows a series of '41' bytes, indicating the string "AA". Below the dump, a conversion tool is open, showing various representations of the selected data (offset 12320 / 17583). The tool includes fields for Signed/Unsigned 8, 16, 32, and 64 bits, Signed/Unsigned 32-bit float, Hexadecimal, Decimal, Octal, Binary, ASCII Text, and Selection. The Hexadecimal field shows "41 41 41 41".

```
[04/17/23] seed@VM: ~/.../09_hashing$ gcc benign_evil.c -o be
[04/17/23] seed@VM: ~/.../09_hashing$ bless be
```

Offset = 12320

Padding will mess up our attack, so we must make sure that padding doesn't get added

Offset = 12352 (multiple of 64)


```
#include <stdio.h>
#define LENGTH 400
```

Prefix

```
unsigned char X[LENGTH]= {
```

```
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
```

suffix1

```
unsigned char Y[LENGTH]= {
```

```
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
```

suffix2

```
int main()
{
    int i = 0;

    for (i = 0; i < LENGTH; i++){
        if (X[i] != Y[i]) break;
    }
    if (i == LENGTH){
        printf("%s\n", "Executing benign code... ");
    }
    else {
        printf("%s\n", "Executing malicious code... ");
    }
    return 0;
}
```

First, we must get the starting location of our array X

The screenshot shows a hex editor window titled "/home/seed/csc1476-code/09_hashing/be - Bless". The main area displays a memory dump with addresses on the left, hex values in the middle, and ASCII characters on the right. The dump shows a series of '41' characters (ASCII 'A') starting at address 00002e61. Below the dump, there is a conversion tool with various input fields and checkboxes. The 'Offset' field is highlighted in yellow and contains the value '12320 / 17583'. The 'Selection' dropdown is set to 'None'.

```
[04/17/23] seed@VM: ~/.../09_hashing$ gcc benign_evil.c -o be
[04/17/23] seed@VM: ~/.../09_hashing$ bless be
```

Offset = 12320

Padding will mess up our attack, so we must make sure that padding doesn't get added

Offset = 12352 (multiple of 64)

prefix

12352

Prefix

```
#include <stdio.h>
#define LENGTH 400
```

```
unsigned char X[LENGTH]= {
```

```
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
```

P

128 bytes

12481

suffix1

```
unsigned char Y[LENGTH]= {
```

```
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
"AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
```

Q

128 bytes

suffix

```
int main()
{
    int i = 0;

    for (i = 0; i < LENGTH; i++){
        if (X[i] != Y[i]) break;
    }
    if (i == LENGTH){
        printf("%s\n", "Executing benign code... ");
    }
    else {
        printf("%s\n", "Executing malicious code... ");
    }
    return 0;
}
```

suffix2

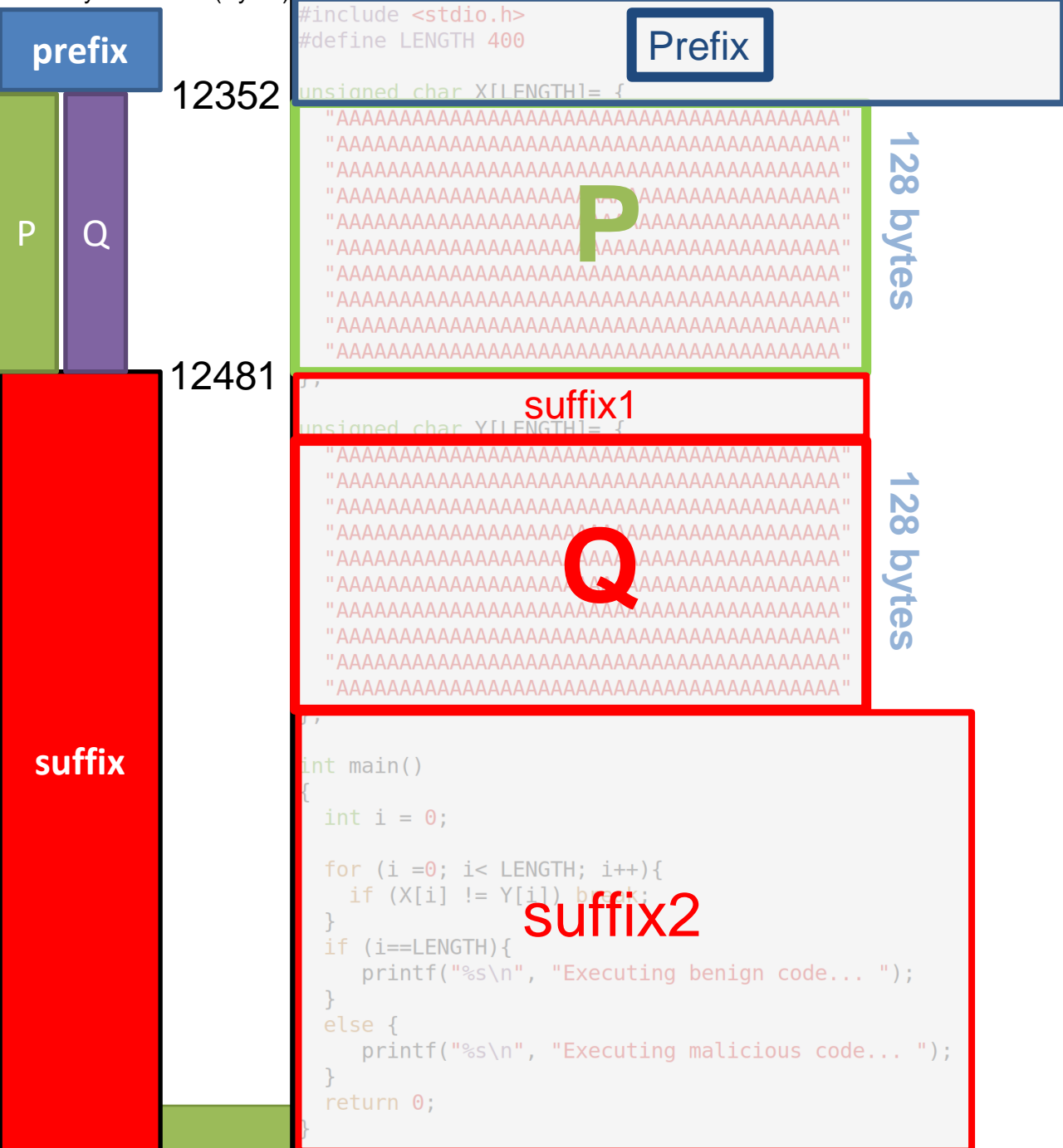
Offset = 12352

P and Q will once again be 128 bytes long

Therefore, P ends at $(12352 + 128)$
= 12480

(we have to add +1 when getting the
suffix to prevent getting an extra byte)

```
[04/17/23] seed@VM:~/.../09_hashing$ head -c 12352 be > prefix
[04/17/23] seed@VM:~/.../09_hashing$ tail -c +12481 be > suffix
```



Now that we have the prefix,
we can generate P and Q

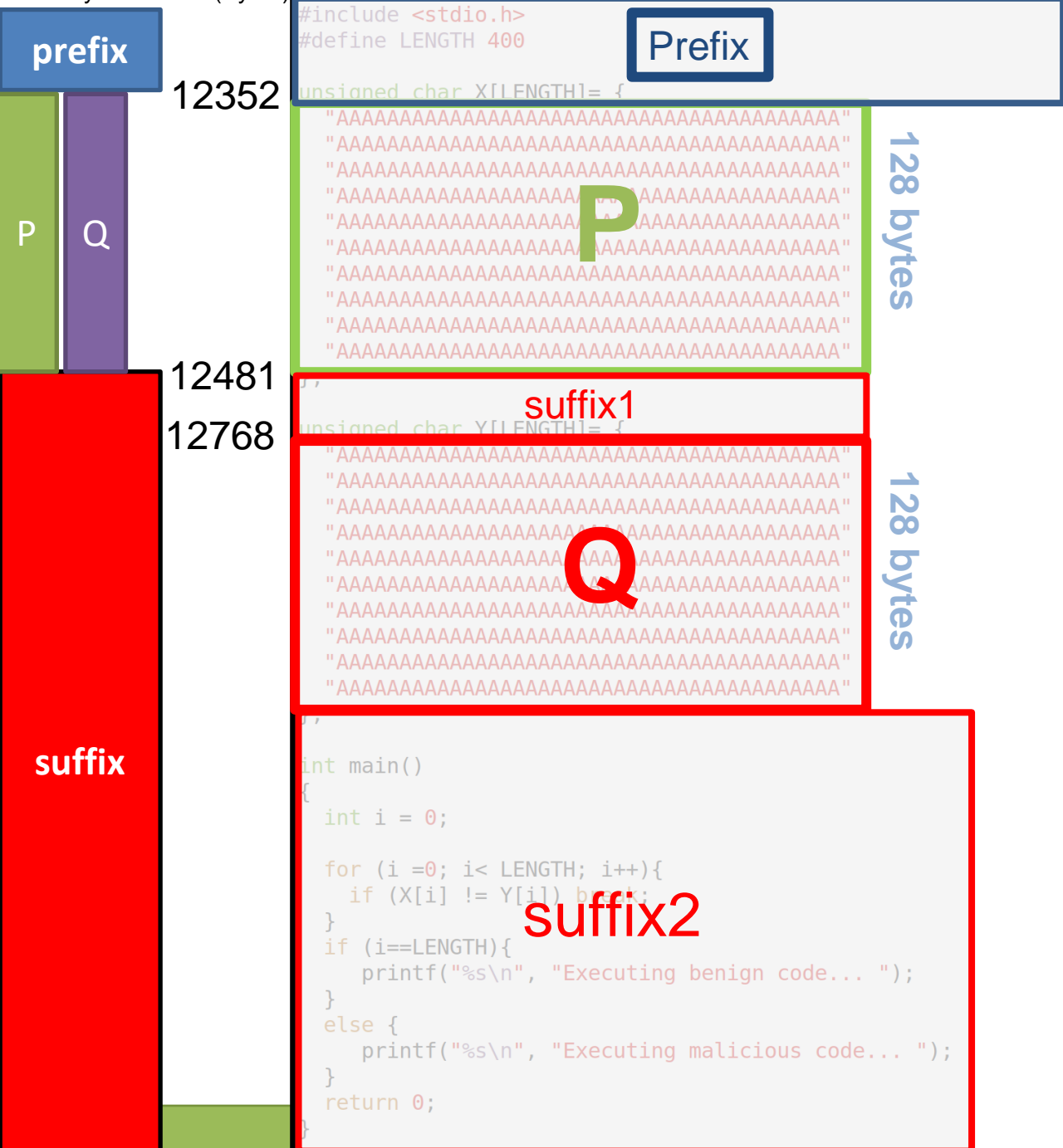
```
[04/17/23]seed@VM:~/.../09_hashing$ md5collgen -p prefix -o out1 out2
MD5 collision generator v1.5
by Marc Stevens (http://www.win.tue.nl/hashclash/)
```

```
Using output filenames: 'out1' and 'out2'
Using prefixfile: 'prefix'
Using initial value: fe5a2a34aa864ea9110d0e2fa43e0327
```

```
Generating first block: .....
Generating second block: S10.....
Running time: 4.60817 s
```

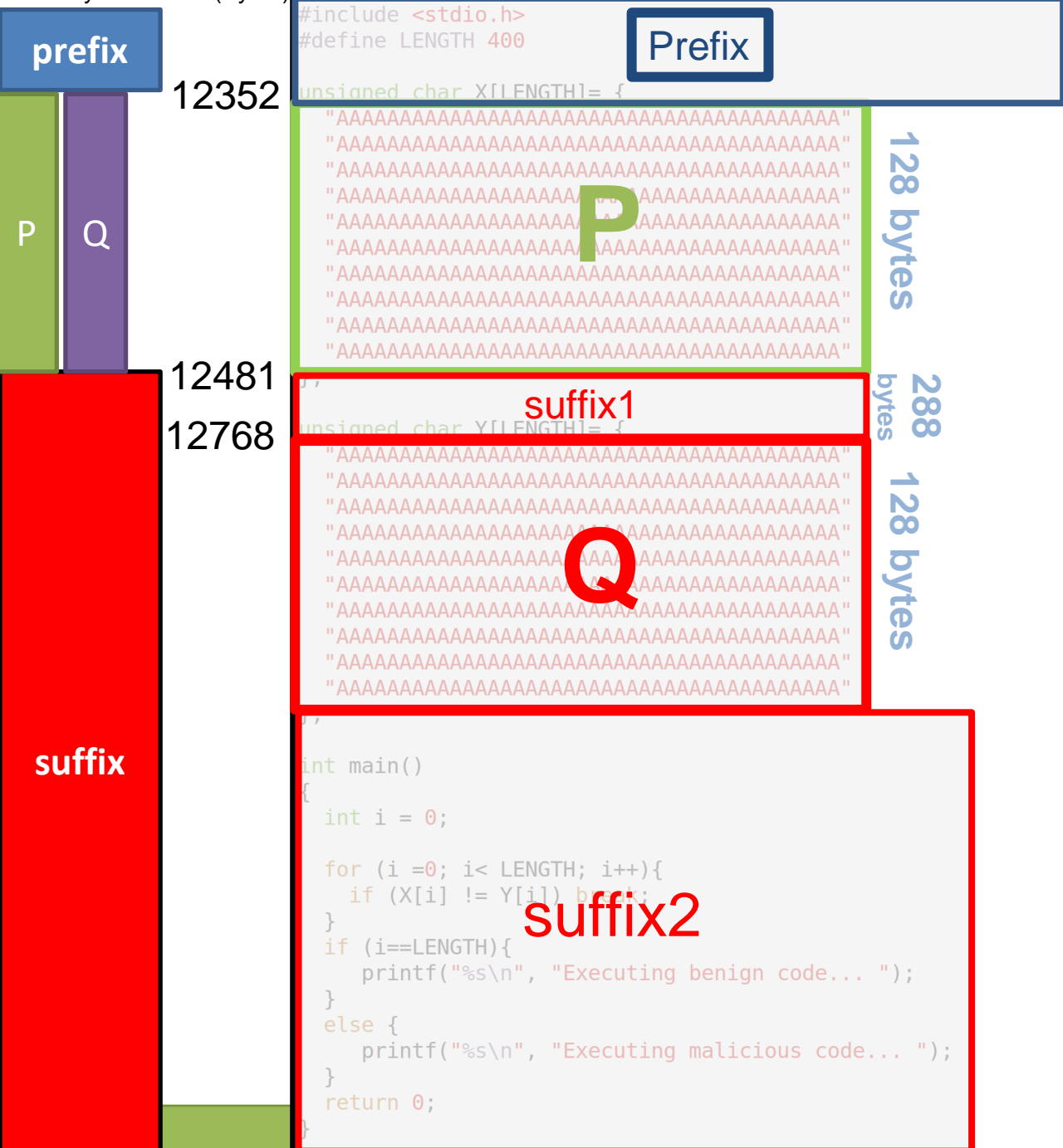
Because we just want P and Q (not the prefix), we will take the final 128 bytes of the output of `out1` and `out2`

```
[04/17/23] seed@VM:~/.../09_hashing$ tail -c 128 out1 > P
[04/17/23] seed@VM:~/.../09_hashing$ tail -c 128 out2 > Q
```



To avoid padding, we had to move the beginning of P up 32 bytes into the array X, so when we insert P/Q into array Y, we need to make sure we also do 32 bytes

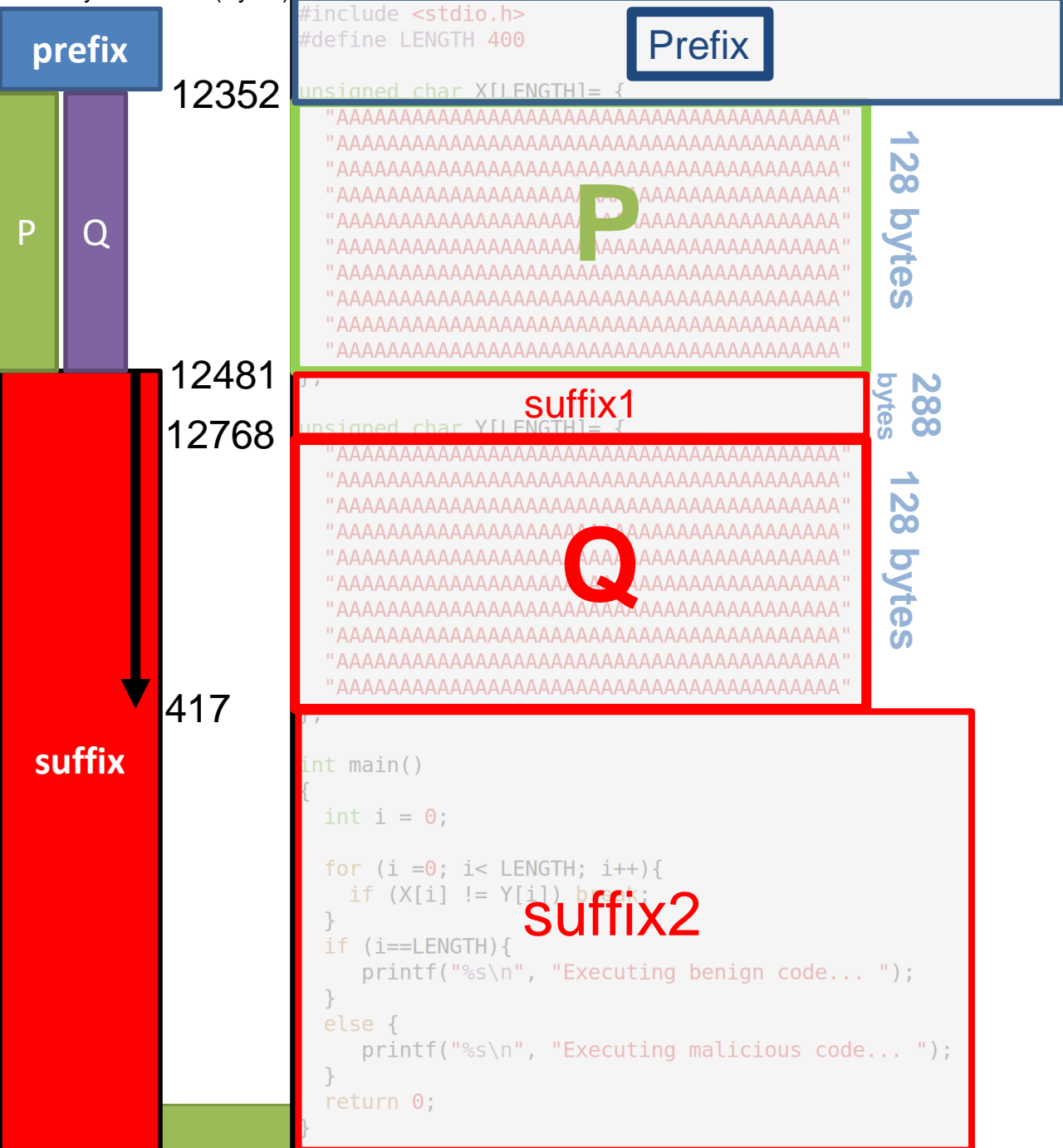
Array Y starts at 12736, but we need to inject at byte $(12736 + 28) = 12768$



To avoid padding, we had to move the beginning of P up 32 bytes into the array X, so when we insert P/Q into array Y, we need to make sure we also do 32 bytes

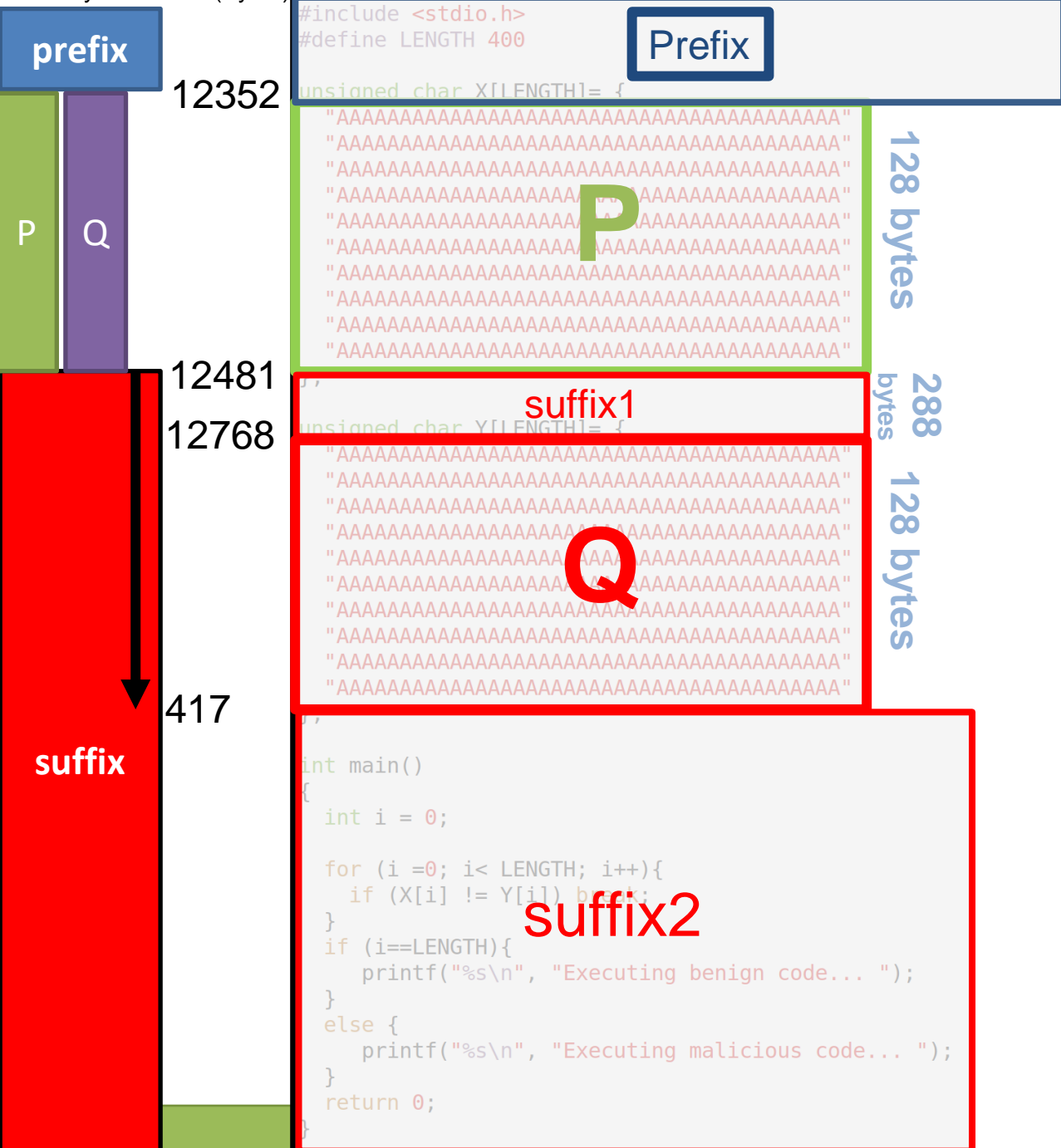
Array Y starts at 12736, but we need to inject at byte $(12736 + 28) = 12768$

Therefore, the size of suffix1 will be $12768 - 12481 = 288$



Now, we put everything together

```
[04/18/23]seed@VM:~/.../09_hashing$ cat prefix P suffix1 P suffix2 > final1
[04/18/23]seed@VM:~/.../09_hashing$ cat prefix Q suffix1 P suffix2 > final2
```



Now, we put everything together

```
[04/18/23] seed@VM:~/.../09_hashing$ cat prefix P suffix1 P suffix2 > final1
[04/18/23] seed@VM:~/.../09_hashing$ cat prefix Q suffix1 P suffix2 > final2
```

Verify that hashes match:

```
[04/18/23] seed@VM:~/.../09_hashing$ md5sum final1 final2
7eb3ea7eae7faa2efbd0ddfa0c7022e76  final1
7eb3ea7eae7faa2efbd0ddfa0c7022e76  final2
```

✓


```
[04/18/23] seed@VM:~/.../09_hashing$ cat prefix P suffix1 P suffix2 > final1
[04/18/23] seed@VM:~/.../09_hashing$ cat prefix Q suffix1 P suffix2 > final2
```

```
[04/18/23]seed@VM:~/.../09_hashing$ md5sum final1 final2
7eb3ea7eaefaa2efbd0ddfa0c7022e76  final1
7eb3ea7eaefaa2efbd0ddfa0c7022e76  final2
```

```
[04/18/23] seed@VM:~/.../09_hashing$ chmod u+x final1 final2
[04/18/23] seed@VM:~/.../09_hashing$ ./final1
Executing benign code...
[04/18/23] seed@VM:~/.../09_hashing$ ./final2
Executing malicious code...
```