

### Task 1:

Generate `pre.dat` with `python3 -c "print('9'*64,end=' ')" > pre.dat`

[illegible]

Generate `out1.bin` & `out2.bin`

```
M:~/.../A3$ md5collgen -p pre.dat -o out1.bin out2.bin
MD5 collision generator v1.5
by Marc Stevens (http://www.win.tue.nl/hashclash/)
```

```
Using output filenames: 'out1.bin' and 'out2.bin'
Using prefixfile: 'pre.dat'
Using initial value: b32a9b522add4eb5db55f4eb9f6c9327
```

```
Generating first block: .....
Generating second block: S10.....
Running time: 6.3705 s
```

### Run given commands

```
[03/19/25] seed@VM:~/.../A3$ diff out1.bin out2.bin
Binary files out1.bin and out2.bin differ
[03/19/25] seed@VM:~/.../A3$ md5sum out1.bin
4ba4683e94f0324eb756ealf3435c7c7  out1.bin
[03/19/25] seed@VM:~/.../A3$ md5sum out2.bin
4ba4683e94f0324eb756ealf3435c7c7  out2.bin
```

out1.bin

Out2.bin

[illegible]

1. If the length of the prefix file is not a multiple of 64, MD5's padding ensures the total length (prefix + padding) aligns to a 64-byte boundary before the collision blocks are appended
2. If it is exactly 64 bytes, it fills exactly 1 MD5 block, and the collision blocks start immediately after
3. No they are not completely different.

### First Block (0-63):

- Byte 2: f7 vs. 77.
- Byte 14: 5d vs. dd.
- Byte 21: 7c vs. fc.

### Second Block (64-127):

- Byte 66: 53 vs. d3.
- Byte 78: cb vs. 4b.
- Byte 85: d0 vs. 50.

### Task 2:

We know the files from the previous step have the same md5 value

```
[03/19/25] seed@VM:~/.../A3$ md5sum out1.bin
4ba4683e94f0324eb756ea1f3435c7c7  out1.bin
[03/19/25] seed@VM:~/.../A3$ md5sum out2.bin
4ba4683e94f0324eb756ea1f3435c7c7  out2.bin
```

So we create a new suffix with `python3 -c "print('114514'*10,end='')" > suffix.txt`

Then concatenate with out1.bin and out2.bin using `cat out1.bin suffix.txt > 1T.out` and `cat out2.bin suffix.txt > 2T.out`

```
[03/19/25] seed@VM:~/.../A3$ python3 -c "print('22856'*10,end='')" > suffix.txt
[03/19/25] seed@VM:~/.../A3$ cat out1.bin suffix.txt > 1T.out
[03/19/25] seed@VM:~/.../A3$ cat out2.bin suffix.txt > 2T.out
[03/19/25] seed@VM:~/.../A3$ md5sum 1T.out
87e2308658b694aa02e91392b41b0f2b  1T.out
[03/19/25] seed@VM:~/.../A3$ md5sum 2T.out
87e2308658b694aa02e91392b41b0f2b  2T.out
```

Create and compile `print_array.c`

\*the contents of the array are the hex coding for 'A' 200x

[illegible]

Run `bless print_array`

[illegible]

Offset in bottom right states 12320/16991 so array starts at position 12320. Seeing as this is not divisible by 64, we use the closest starting position to the array that is divisible by 64. - 12288

```
head -c 12288 print_array > prefix and md5collgen -p prefix -o
task3_a.bin task3_b.bin
```

```
[03/20/25]seed@VM:~/.../A3$ head -c 12288 print_array > prefix
[03/20/25]seed@VM:~/.../A3$ md5collgen -p prefix -o task3_a.bin task3_b.bin
MD5 collision generator v1.5
by Marc Stevens (http://www.win.tue.nl/hashclash/)
```

```
Using output filenames: 'task3_a.bin' and 'task3_b.bin'
Using prefixfile: 'prefix'
Using initial value: 66e90aff81a62e7bee2c4e52b49cf84d
```

```
Generating first block: ....
Generating second block: W..
Running time: 4.65445 s
```

Generate Suffix which should be prefix + 128 so we get 12416

```
tail -c +12416 print_array > suffix
```

Run

```
Tail -c 128 task3_a.bin > p
Tail -c 128 task3_b.bin > q
Cat prefix p suffix > task3_1
Cat prefix q suffix > task3_2
```

Then check if their md5sums are the same with `md5sum task3_1 & md5sum task3_2`

```
[03/20/25]seed@VM:~/.../A3$ tail -c 128 task3_a.bin > p
[03/20/25]seed@VM:~/.../A3$ tail -c 128 task3_b.bin > q
[03/20/25]seed@VM:~/.../A3$ cat prefix p suffix > task3_1
[03/20/25]seed@VM:~/.../A3$ cat prefix q suffix > task3_2
[03/20/25]seed@VM:~/.../A3$ md5sum task3_1
2b780b58af1146503e88c110fa86e514 task3_1
[03/20/25]seed@VM:~/.../A3$ md5sum task3_2
2b780b58af1146503e88c110fa86e514 task3_2
[03/20/25]seed@VM:~/.../A3$ █
```

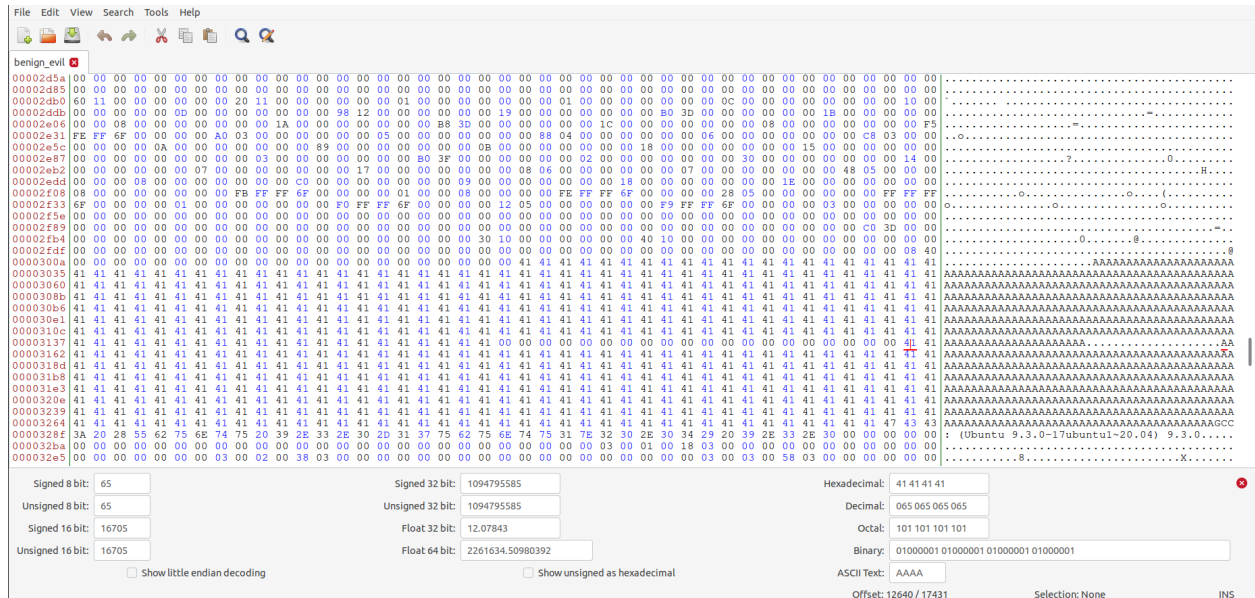
Task 4:

Create `benign_evil.c` then compile

```
1#include <stdio.h>
2#define LEN 300
3
4unsigned char X[LEN] = {
5    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
6    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
7    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
8    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
9    "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
10   "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"};
11
12unsigned char Y[LEN] = {
13   "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
14   "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
15   "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
16   "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
17   "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"
18   "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA"};
19
20int main()
21{
22   for (int i = 0; i < LEN; i++)
23   {
24       if (X[i] != Y[i])
25       {
26           printf("i = %d, X[i] = %.2x, Y[i] = %.2x\n", i, X[i], Y[i]);
27           printf("Malicious\n");
28           return 0;
29       }
30   }
31   printf("Benign\n");
32   return 0;
33}
```

Run `bleess benign_evil` and find where X and Y array starts

First array begins at 12320/17431, Second array begins at 12640/17431



Obtain prefix and suffix from benign\_evil

```
[03/23/25]seed@VM:~/.../A3$ head -c 12320 benign_evil > prefix
[03/23/25]seed@VM:~/.../A3$ tail -c +12448 benign_evil > suffix
```

```
[03/23/25]seed@VM:~/.../A3$ md5collgen -p prefix -o s1 s2
MD5 collision generator v1.5
by Marc Stevens (http://www.win.tue.nl/hashclash/)
```

```
Using output filenames: 's1' and 's2'
Using prefixfile: 'prefix'
Using initial value: ce32bebe2f1f2eefe11981ae5bdd94e3
```

```
Generating first block: ....
Generating second block: W.....
Running time: 5.40198 s
```

Run `tail -c 128 s1 > p` and `tail -c 128 s2 > Q` then generate the 2 halves of the suffix

Y starts from 0xc1(301) in suffix, So if we want to make X=Y, we should replace offset [301,429) with the same P or Q generated above



