

Find SW vulnerability

via automatic method

WHO AM I

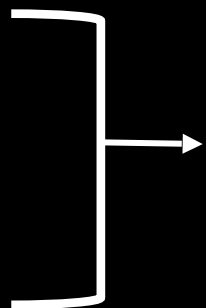
- 정재영 - JJY
- BOB 8기 취약점 분석 트랙
- Reverselab - Pwner
- Bug hunting
 - CVE-2020-12830
 - CVE-2020-7860
 - Lots of KVEs

How to find Bugs?

- Fuzzing
- IDA python
- Monitoring tools

How to find Bugs?

- Fuzzing
- IDA python
- Monitoring tools



자동화

적게 일하고 많이 벌자!!

Fuzzing

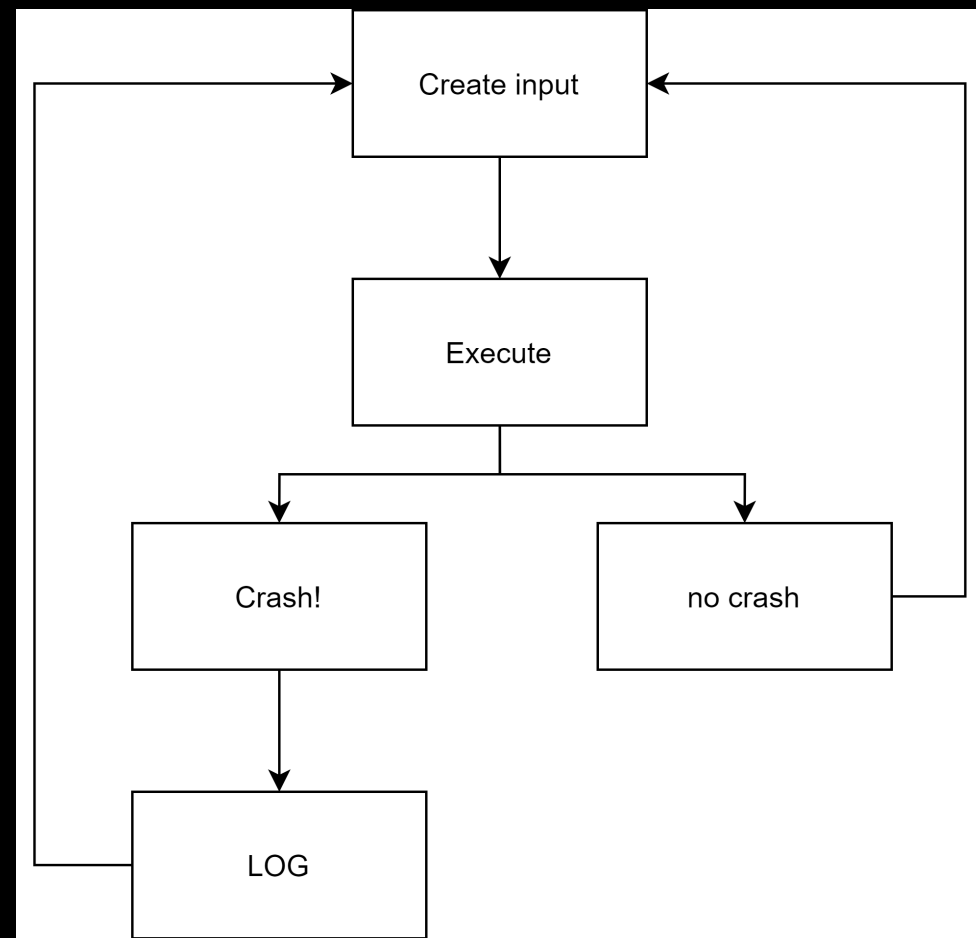
- Fuzzer
 - 랜덤한 인풋을 생성하여 프로그램의 버그를 찾는 도구

Fuzzing

- Black box fuzzer
 - BFF
- Coverage based fuzzer
 - AFL

BFF

- Basic Fuzzing Framework
- Black box fuzzer
- 시드 제공 + 간단한 설정 파일 설정



BFF

- 국내 소프트웨어 32개를 대상으로 퍼저를 (대략 1일 정도) 돌린 결과...

취약점 개수	프로그램 수
0	7
0~10	13
10~100	8
100~150	2
150~200	2

BFF

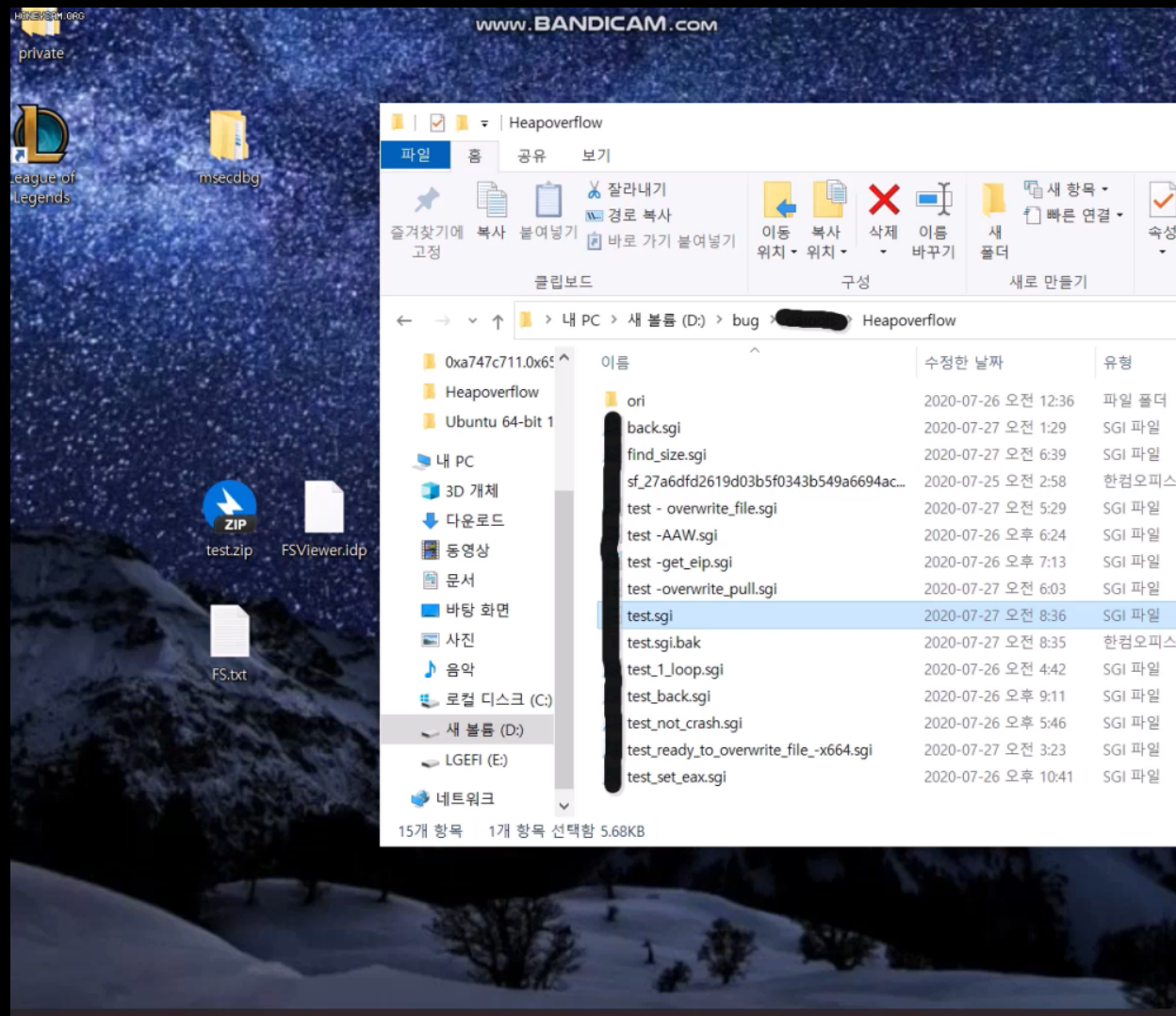
- 국내 소프트웨어 32개를 대상으로 퍼저를 (대략 1일 정도) 돌린 결과...

취약점 개수	프로그램 수
0	7
0~10	13
10~100	8
100~150	2
150~200	2

BFF

- Buffer Overflow, Integer Overflow, Double free ...
- Exploit!
 - Heap Logic
 - SEH overwrite
 - ROP

DEMO



Let's go to coverage based fuzzing

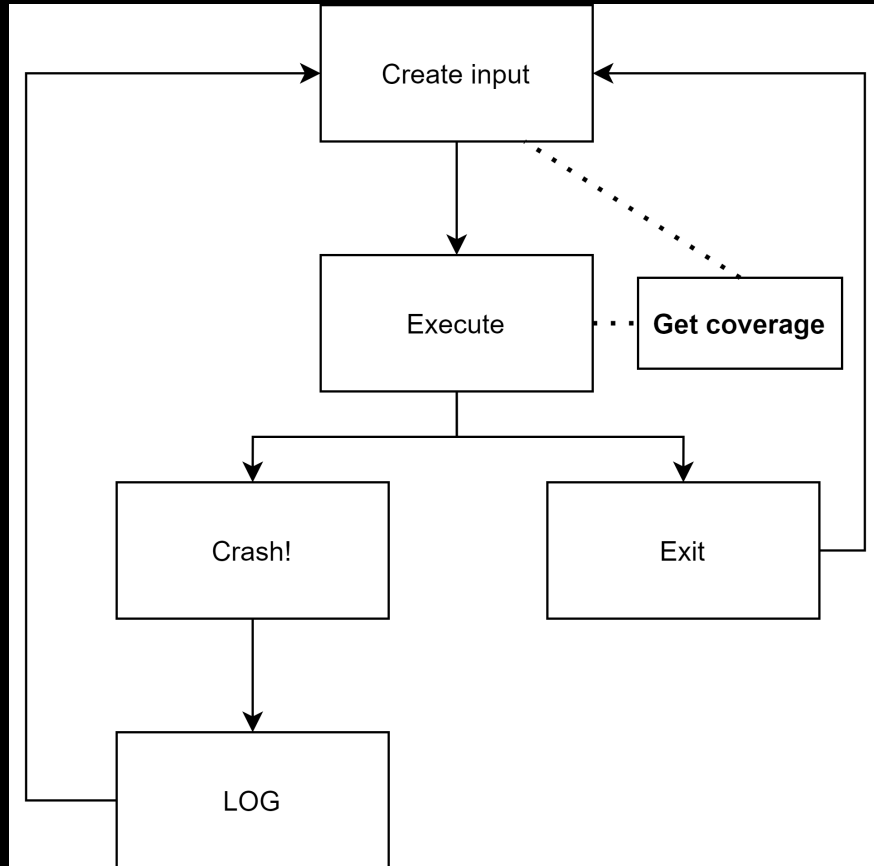
메이저 프로그램의 경우 크래시가 잘 나오지 않음

취약점 개수	프로그램 수
0	7
0~10	13
10~100	8
100~150	2
150~200	2

→ Unexploitable?

Let's go to coverage based fuzzing

- american fuzzy lop



- Code coverage를 이용하여 Input을 smart하게 생성

Hard ...?

- GUI 프로그램을 coverage based fuzzing하기 힘든 이유
 - 종료 시점이 정해져 있지 않음.
 - 어느 시점에 coverage 측정을 끝내야 하지?

Hard ...?

- GUI 프로그램을 coverage based fuzzing하기 힘든 이유
 - 종료 시점이 정해져 있지 않음.
 - 어느 시점에 coverage 측정을 끝내야 하지?

==> 라이브러리에서 **핵심 함수만 가져와서 실행**하는 harness를 만들면
해결 가능

How to find target function

- 파일 or 패킷을 분석하는 함수.
- 공격자가 인자를 마음대로 제어할 수 있는 함수.
- Entry point에 해당하는 함수

Make harness

- Target 함수를 import해서 실행하는 프로그램 만들기

```
typedef void (*lib_func)(wchar_t *,int ,int * );  
int main(int argc,char ** argv){  
    int test;  
    lib_func func = NULL;  
    void * handle = dlopen("censored_lib_path",RTLD_NOW | RTLD_GLOBAL);  
    func = (lib_func ) dlsym(handle, "censored_func_name");  
    wchar_t * ws1;  
    char * st = (char *)calloc(0x100000,1);  
    int fd = open(argv[1],O_RDWR);  
    read(fd,st,0xff000);  
    func(st,0,&test);  
    free(st);  
    return 1;  
}
```

Load library

Make harness

- Target 함수를 import해서 실행하는 프로그램 만들기

```
typedef void (*lib_func)(wchar_t *,int ,int * );
int main(int argc,char ** argv){
    int test;
    lib_func func = NULL;
    void * handle = dlopen("censored_lib_path",RTLD_NOW | RTLD_GLOBAL);
    func = (lib_func ) dlsym(handle, "censored_func_name");
    wchar_t * ws1;
    char * st = (char *)calloc(0x100000,1);
    int fd = open(argv[1],O_RDWR);
    read(fd,st,0xff000);
    func(st,0,&test);
    free(st);
    return 1;
}
```

Load function

Make harness

- Target 함수를 import해서 실행하는 프로그램 만들기

```
typedef void (*lib_func)(wchar_t *,int ,int * );
int main(int argc,char ** argv){
    int test;
    lib_func func = NULL;
    void * handle = dlopen("censored_lib_path",RTLD_NOW | RTLD_GLOBAL);
    func = (lib_func ) dlsym(handle, "censored_func_name");
    wchar_t * ws1;
    char * st = (char *)calloc(0x100000,1);
    int fd = open(argv[1],O_RDWR);
    read(fd,st,0xff000);
    func(st,0,&test);
    free(st);
    return 1;
}
```

Get input

Make harness

- Target 함수를 import해서 실행하는 프로그램 만들기

```
typedef void (*lib_func)(wchar_t *,int ,int * );
int main(int argc,char ** argv){
    int test;
    lib_func func = NULL;
    void * handle = dlopen("censored_lib_path",RTLD_NOW | RTLD_GLOBAL);
    func = (lib_func ) dlsym(handle, "censored_func_name");
    wchar_t * ws1;
    char * st = (char *)calloc(0x100000,1);
    int fd = open(argv[1],O_RDWR);
    read(fd,st,0xff000);
    func(st,0,&test);
    free(st);
    return 1;
}
```

Execute function

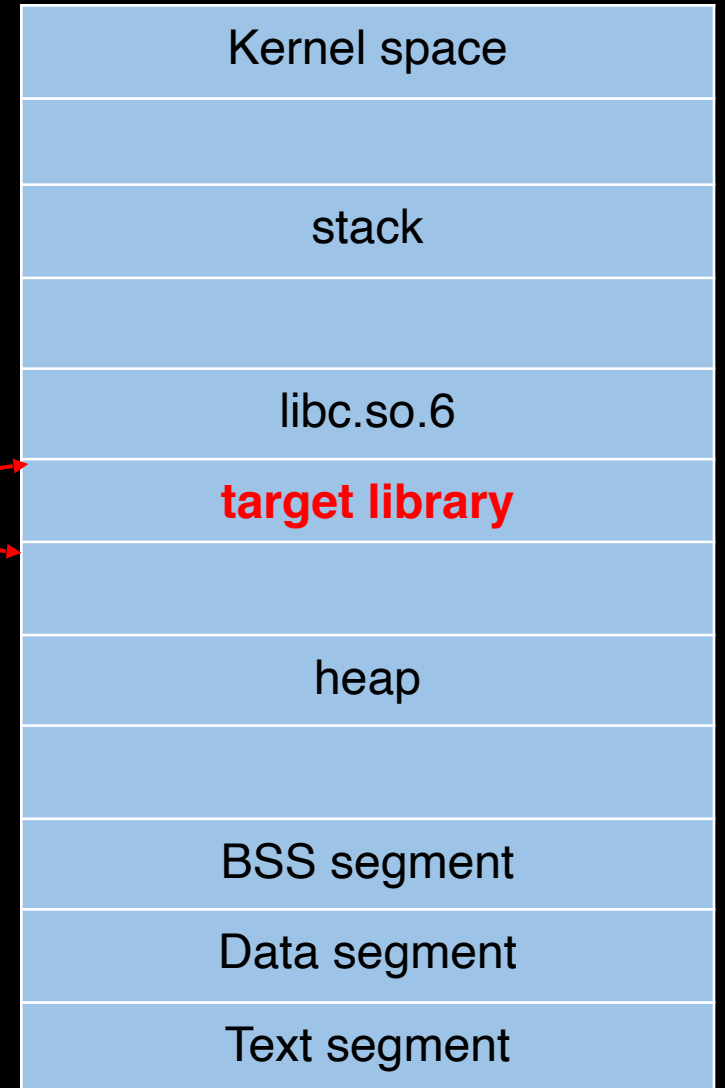
Coverage in AFL qemu mode

- AFL with no source code --> Qemu mode

```
static inline void afl_maybe_log(abi_ulong cur_loc) {  
  
    static __thread abi_ulong prev_loc;  
  
    /* Optimize for cur_loc > afl_end_code, which is the most likely case on  
    Linux systems. */  
  
    if (cur_loc > afl_end_code || cur_loc < afl_start_code || !afl_area_ptr)  
        return;  
  
    /* Looks like QEMU always maps to fixed locations, so ASAN is not a  
    concern. Phew. But instruction addresses may be aligned. Let's mangle  
    the value to get something quasi-uniform. */  
  
    cur_loc = (cur_loc >> 4) ^ (cur_loc << 8);  
    cur_loc &= MAP_SIZE - 1;  
  
    /* Implement probabilistic instrumentation by looking at scrambled block  
    address. This keeps the instrumented locations stable across runs. */  
  
    if (cur_loc >= afl_inst_rms) return;  
  
    afl_area_ptr[cur_loc ^ prev_loc]++;  
    prev_loc = cur_loc >> 1;  
  
}
```

Coverage in AFL qemu mode

```
//patches/afl-qemu-cpu-inl.h
if (getenv("AFL_CODE_START"))
    afl_start_code = strtoll(getenv("AFL_CODE_START"), NULL, 16);
if (getenv("AFL_CODE_END"))
    afl_end_code = strtoll(getenv("AFL_CODE_END"), NULL, 16);
```



Gooooo!

```
HONEYHONEY
afl-fuzz-Q-i ../min_sample-o ../out - [REDACTED] x jyy@ubuntu: ~/bug/

american fuzzy lop ++2.68c [mark] [explore] {0}

process timing
  run time : 0 days, 0 hrs, 14 min, 19 sec
  last new path : 0 days, 0 hrs, 0 min, 3 sec
  last uniq crash : none seen yet
  last uniq hang : none seen yet
cycle progress
  now processing : 0.0 (0.0%)
  paths timed out : 0 (0.00%)
stage progress
  now trying : bitflip 1/1
  stage execs : 6404/455k (1.40%)
  total execs : 11.3k
  exec speed : 9.95/sec (zzzz...)
fuzzing strategy yields
  bit flips : 0/0, 0/0, 0/0
  byte flips : 0/0, 0/0, 0/0
  arithmetics : 0/0, 0/0, 0/0
  known ints : 0/0, 0/0, 0/0
  dictionary : 0/0, 0/0, 0/0
  havoc/splice : 0/0, 0/0
  py/custom : 0/0, 0/0
  trim : 0.00%/1763, n/a

overall results
  cycles done : 0
  total paths : 360
  uniq crashes : 0
  uniq hangs : 0
map coverage
  map density : 7.89% / 13.47%
  count coverage : 1.77 bits/tuple
findings in depth
  favored paths : 9 (2.50%)
  new edges on : 19 (5.28%)
  total crashes : 0 (0 unique)
  total tmouts : 842 (55 unique)
path geometry
  levels : 2
  pending : 360
  pend fav : 9
  own finds : 351
  imported : n/a
  stability : 99.05%

[cpu000: 12%]
```

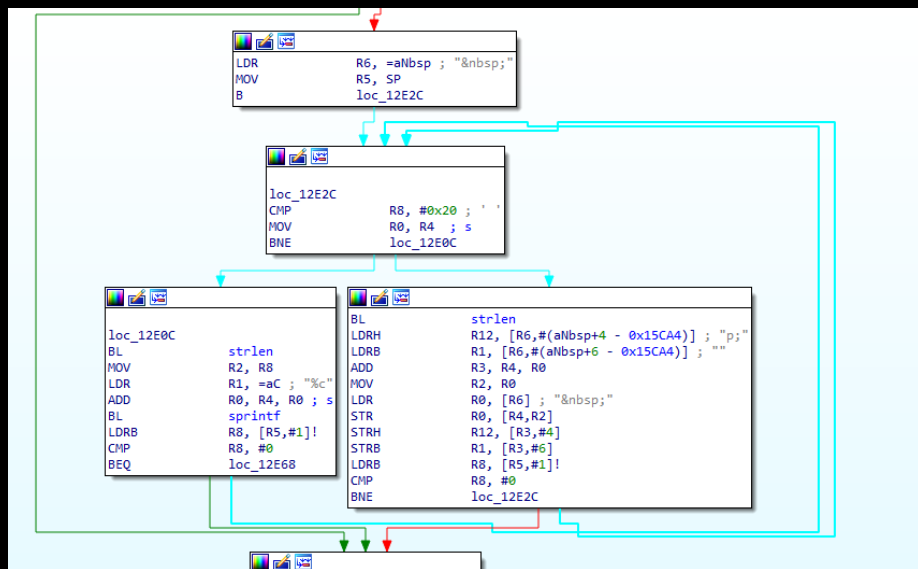
Result

- 1개의 프로그램에서 7개의 취약점 발견
 - Type confusion, Heap Buffer Overflow, OOB, DOS ...

Feedback driven fuzzing is powerfull!

IDA python

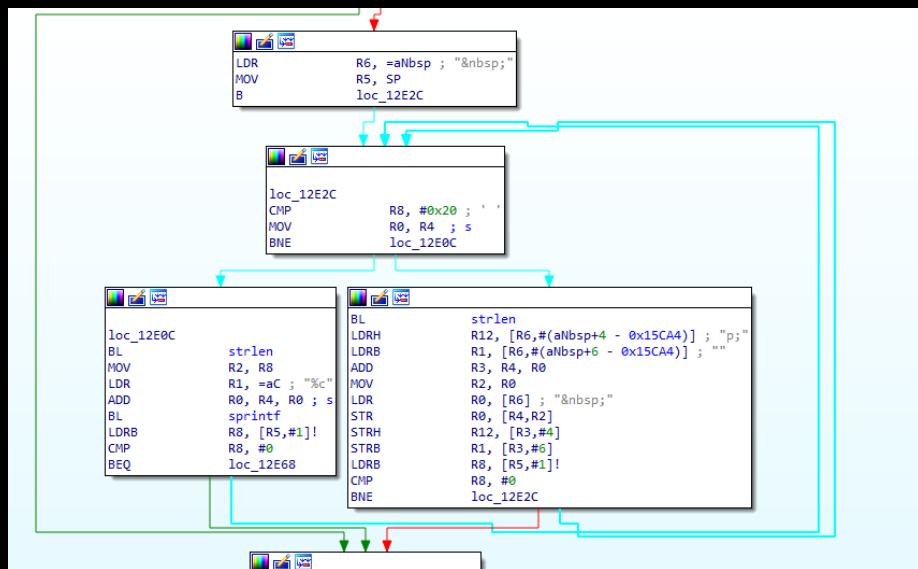
- IDA에서 정적 분석을 위해 지원하는 python 플러그인



```
if ( v9[0] )
{
    v4 = v9;
    do
    {
        while ( v3 != ' ' )
        {
            v5 = strlen(v10);
            sprintf(&v10[v5], "%c", v3);
            v6 = (unsigned __int8)*++v4;
            v3 = v6;
            if ( !v6 )
                return strcpy(a2, v10);
        }
        strcat(v10, "&nbsp;");
        v7 = (unsigned __int8)*++v4;
        v3 = v7;
    }
    while ( v7 );
}
```

IDA python

- IDA에서 정적 분석을 위해 지원하는 python 플러그인



asm

```
if ( v9[0] )
{
    v4 = v9;
    do
    {
        while ( v3 != ' ' )
        {
            v5 = strlen(v10);
            sprintf(&v10[v5], "%c", v3);
            v6 = (unsigned __int8)*++v4;
            v3 = v6;
            if ( !v6 )
                return strcpy(a2, v10);
        }
        strcat(v10, "&nbsp;");
        v7 = (unsigned __int8)*++v4;
        v3 = v7;
    }
    while ( v7 );
}
```

Hex-ray

Fuzzing vs IDA python

- 장점
 - Faster than Fuzzing
 - Can detect Non-Memory-corruption bug
- 단점
 - Hard to make good script
 - Etc...

- hard case

```
v3 = *(_DWORD *) (a1 + 72);
v4 = a3;
v5 = a2;
v6 = a1;
if ( !a3 )
{
    *(_DWORD *) (**(_QWORD **) (a1 + 80) + 40i64) = 40;
    (**(void (**)(void)) (a1 + 80)) ();
}
if ( !*(_BYTE *) (v6 + 40) )
{
    v7 = v4 + v3;
    v8 = *(_QWORD *) (v6 + 64) | ((v5 & ((1i64 << v4) - 1)) << (24 - (unsigned __int8)v7));
    if ( (signed int)v7 >= 8 )
    {
        v9 = (unsigned __int64)v7 >> 3;
        v7 -= 8 * (v7 >> 3);
        do
        {
            **(_BYTE **) (v6 + 48) = BYTE2(v8);
            v10 = (_BYTE *) (*(_QWORD *) (v6 + 48) + 1i64);
            v11 = (*(_QWORD *) (v6 + 56))-- == 1i64;
            *(_QWORD *) (v6 + 48) = v10;
            if ( v11 )
            {
                v12 = *(_QWORD *) (*(_QWORD *) (v6 + 80) + 40i64);
                if ( !(* (unsigned __int8 **) (void)) (v12 + 24)) () )
```

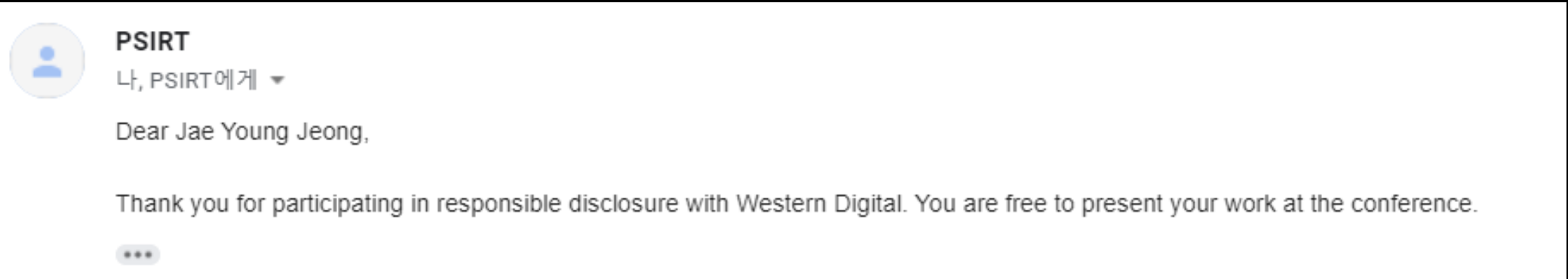
다음 코드에서 유저 인풋 변수와
호출되는 함수 주소를 구하시오 (4점)

IDA python

- 간단한 구조를 가진 프로그램에 유리
- IOT web server, CGI ...
 - Stack buffer overflow
 - command injection
 - etc

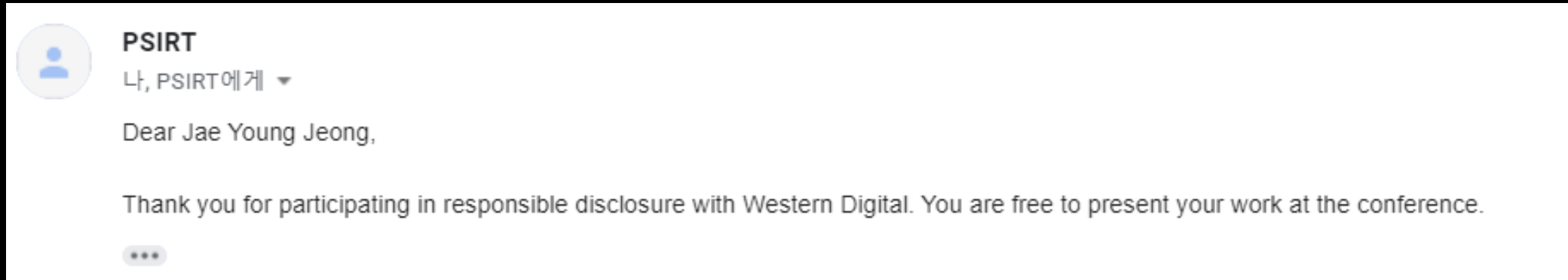
Let's go to the real world!

- Western Digital



Let's go to the real world!

- Western Digital



- Find Stack buffer overflow in CGI
 - Sprintf -> too many 1-days

Progress

- Get local variable info
- Get user input info
- Get function arg info
- Check vulnerability

Get local variable info

```
size_t sub_12AB0()
{
    int v0; // r5
    char v2[512]; // [sp+0h] [bp-610h] BYREF
    char v3[512]; // [sp+200h] [bp-410h] BYREF
    char s[528]; // [sp+400h] [bp-210h] BYREF

    cgiFormString((int)"dir", (int)v2, 512);
    cgiFormString((int)"filename", (int)v3, 512);
    fwrite("Pragma: no-cache\r\nCache-Control: no-cache\r\n", 1u, 0x2Bu, (FILE *)cgiOut);
    cgiHeaderContentType("text/xml");
    fwrite("<?xml version=\\"1.0\\" encoding=\\"UTF-8\\"?>", 1u, 0x26u, (FILE *)cgiOut);
    fwrite("<mkdir>", 1u, 7u, (FILE *)cgiOut);
    v0 = sub_1290C(v2, v3);
    sprintf(s, "%s/%s", v2, v3);
    if ( v0 )
    {
        fwrite("<status>error</status>", 1u, 0x16u, (FILE *)cgiOut);
    }
    else
    {
        fwrite("<status>ok</status>", 1u, 0x13u, (FILE *)cgiOut);
        mkdir(s, 0x41FFu);
        chmod(s, 0x41FFu);
        chown(s, 0x1F5u, 0x1F5u);
    }
    return fwrite("</mkdir>", 1u, 8u, (FILE *)cgiOut);
}
```

- V0 -> int
- V2 -> char[512]
- V3 -> char[512]
- s -> char[528]



Get user input info

- 대부분의 IOT web 페이지 실행파일에서는 유저의 input을 parsing하기 위한 함수가 정해져 있음
- 라이브러리 함수를 쓰지 않고 직접 구현한 경우엔 리버싱...

```
cgiFormString("show_file", v42, 2);  
cgiFormString("dir", s, 512);  
cgiFormString("chk_flag", &v40, 4);  
cgiFormString("file_type", &v32, 32);  
cgiFormString("function_id", v37, 12);
```

Get user input info

```
size_t sub_12AB0()
{
    int v0; // r5
    char v2[512]; // [sp+0h] [bp-610h] BYREF
    char v3[512]; // [sp+200h] [bp-410h] BYREF
    char s[528]; // [sp+400h] [bp-210h] BYREF

    cgiFormString((int)"dir", (int)v2, 512);
    cgiFormString((int)"filename", (int)v3, 512);
    fwrite("Pragma: no-cache\r\nCache-Control: no-cache\r\n", 1u, 0x2Bu, (FILE *)cgiOut);
    cgiHeaderContentType("text/xml");
    fwrite("<?xml version=\\"1.0\\" encoding=\\"UTF-8\\"?>", 1u, 0x26u, (FILE *)cgiOut);
    fwrite("<mkdir>", 1u, 7u, (FILE *)cgiOut);
    v0 = sub_1290C(v2, v3);
    sprintf(s, "%s/%s", v2, v3);
    if ( v0 )
    {
        fwrite("<status>error</status>", 1u, 0x16u, (FILE *)cgiOut);
    }
    else
    {
        fwrite("<status>ok</status>", 1u, 0x13u, (FILE *)cgiOut);
        mkdir(s, 0x41FFu);
        chmod(s, 0x41FFu);
        chown(s, 0x1F5u, 0x1F5u);
    }
    return fwrite("</mkdir>", 1u, 8u, (FILE *)cgiOut);
}
```

V2 -> char[512]

V3 -> char[512]

V3:512

V2:512

User inputs

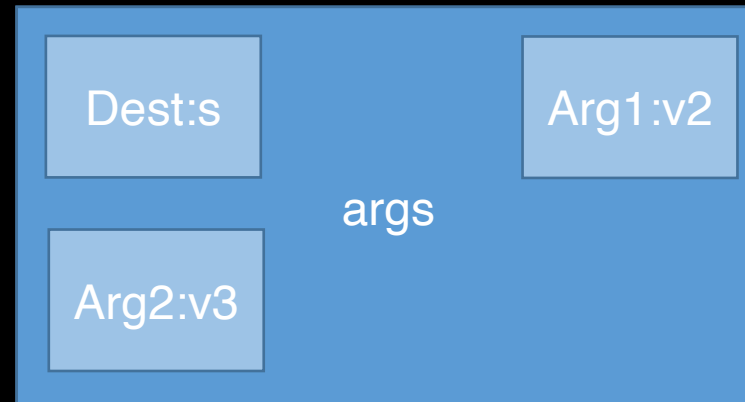
Get function arg info

```
size_t sub_12AB0()
{
    int v0; // r5
    char v2[512]; // [sp+0h] [bp-610h] BYREF
    char v3[512]; // [sp+200h] [bp-410h] BYREF
    char s[528]; // [sp+400h] [bp-210h] BYREF

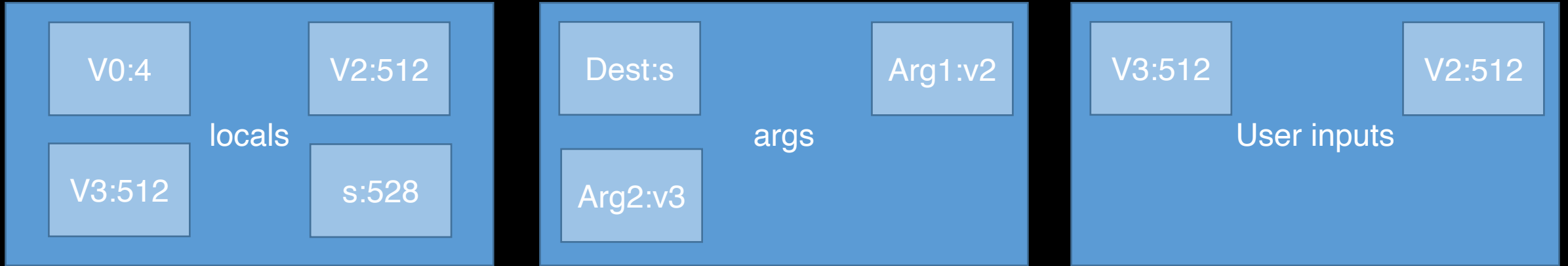
    cgiFormString((int)"dir", (int)v2, 512);
    cgiFormString((int)"filename", (int)v3, 512);
    fwrite("Pragma: no-cache\r\nCache-Control: no-cache\r\n", 1u, 0x2Bu, (FILE *)cgiOut);
    cgiHeaderContentType("text/xml");
    fwrite("<?xml version=\\"1.0\\" encoding=\\"UTF-8\\"?>", 1u, 0x26u, (FILE *)cgiOut);
    fwrite("<mkdir>", 1u, 7u, (FILE *)cgiOut);
    v0 = sub_1290C(v2, v3);
    sprintf(s, "%s/%s", v2, v3);
    if ( v0 )
    {
        fwrite("<status>error</status>", 1u, 0x16u, (FILE *)cgiOut);
    }
    else
    {
        fwrite("<status>ok</status>", 1u, 0x13u, (FILE *)cgiOut);
        mkdir(s, 0x41FFu);
        chmod(s, 0x41FFu);
        chown(s, 0x1F5u, 0x1F5u);
    }
    return fwrite("</mkdir>", 1u, 8u, (FILE *)cgiOut);
}
```

Format = "%s/%s"

source = V2/V3

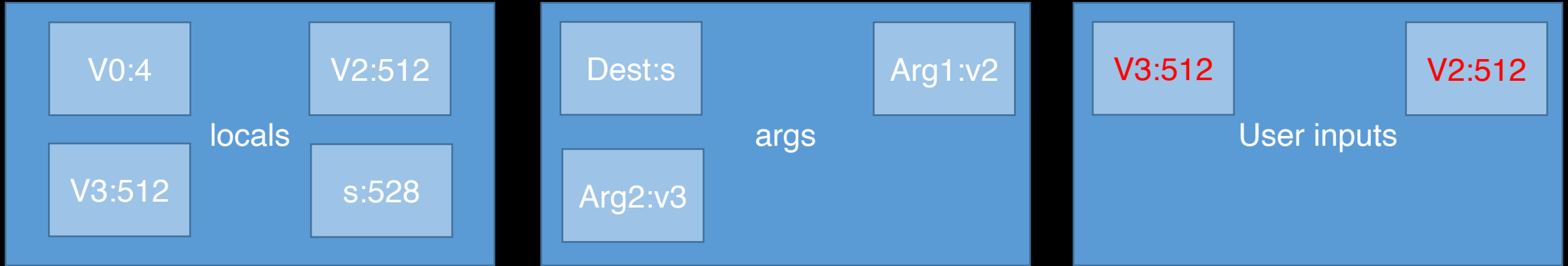


Check vulnerability



IF (dest len < source len) -> Stack buffer overflow

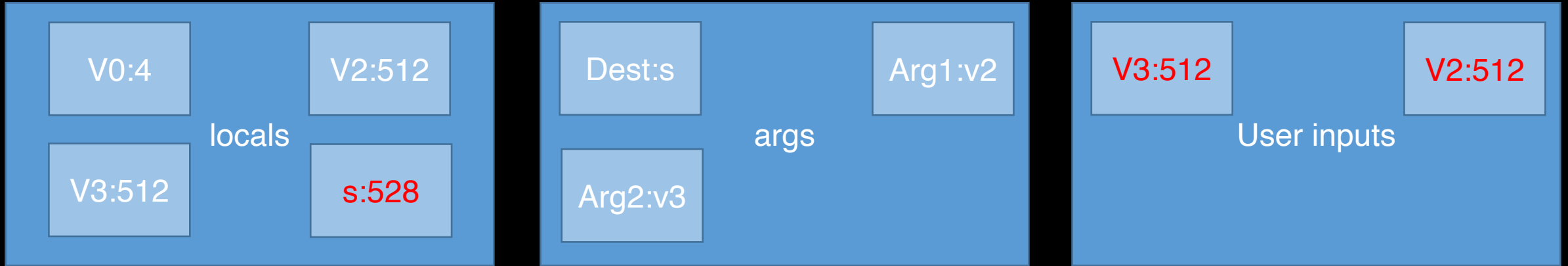
Check vulnerability



IF (dest len < 1024+1) -> Stack buffer overflow

Source = "%s/%s",v2,v3

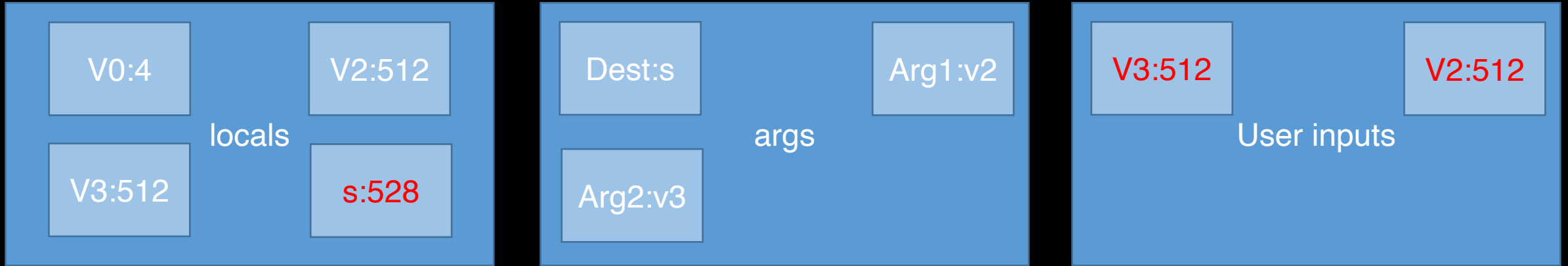
Check vulnerability



IF $(528 < 1024+1)$ -> Stack buffer overflow

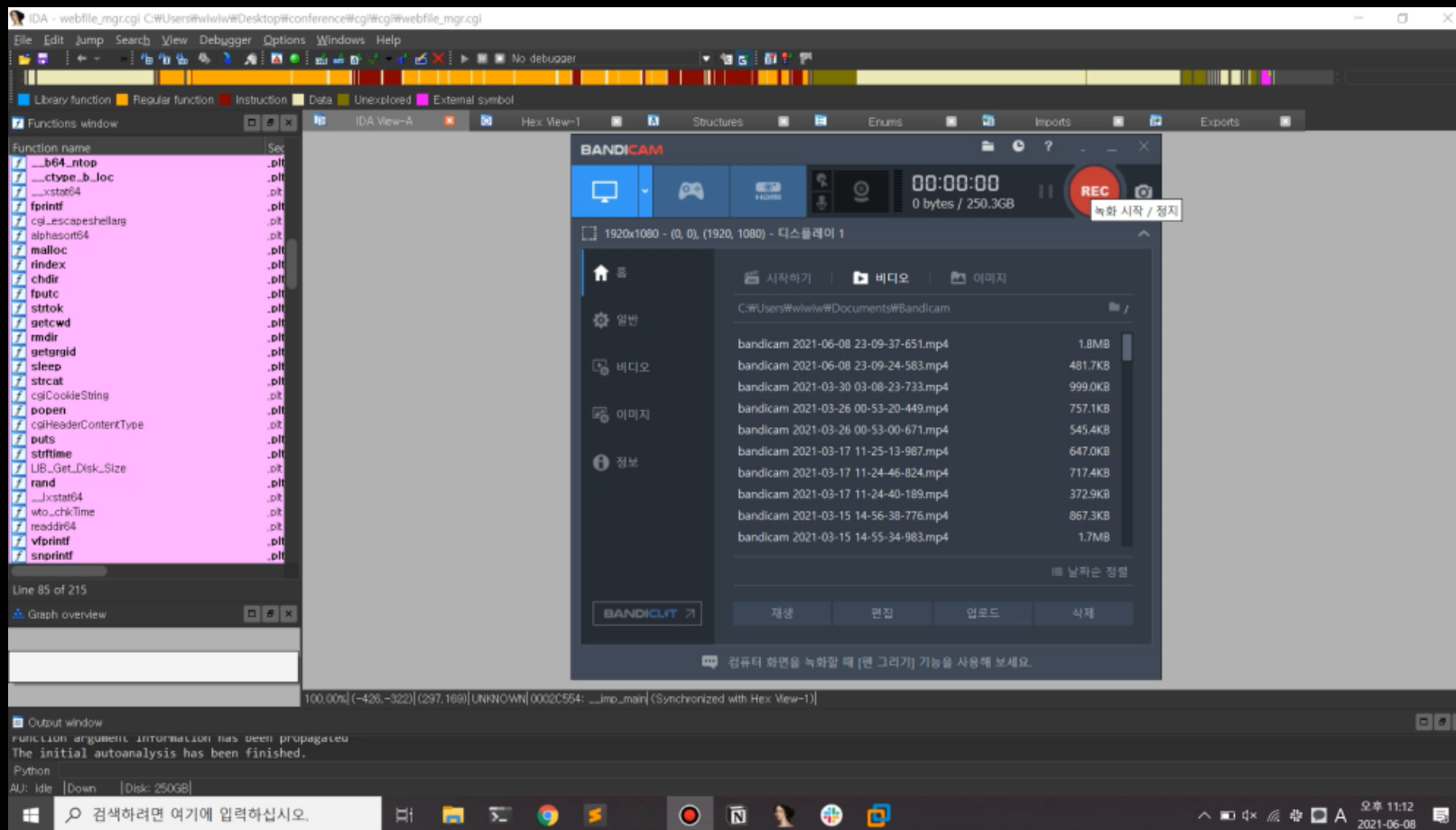
Dest = s:528

Check vulnerability



IF $(528 < 1024 + 1)$ -> **Stack buffer overflow**

DEMO



Monitoring tools

- Windows
 - Process monitor
- Linux
 - ltrace, strace
 - Gdb script

Monitoring tools

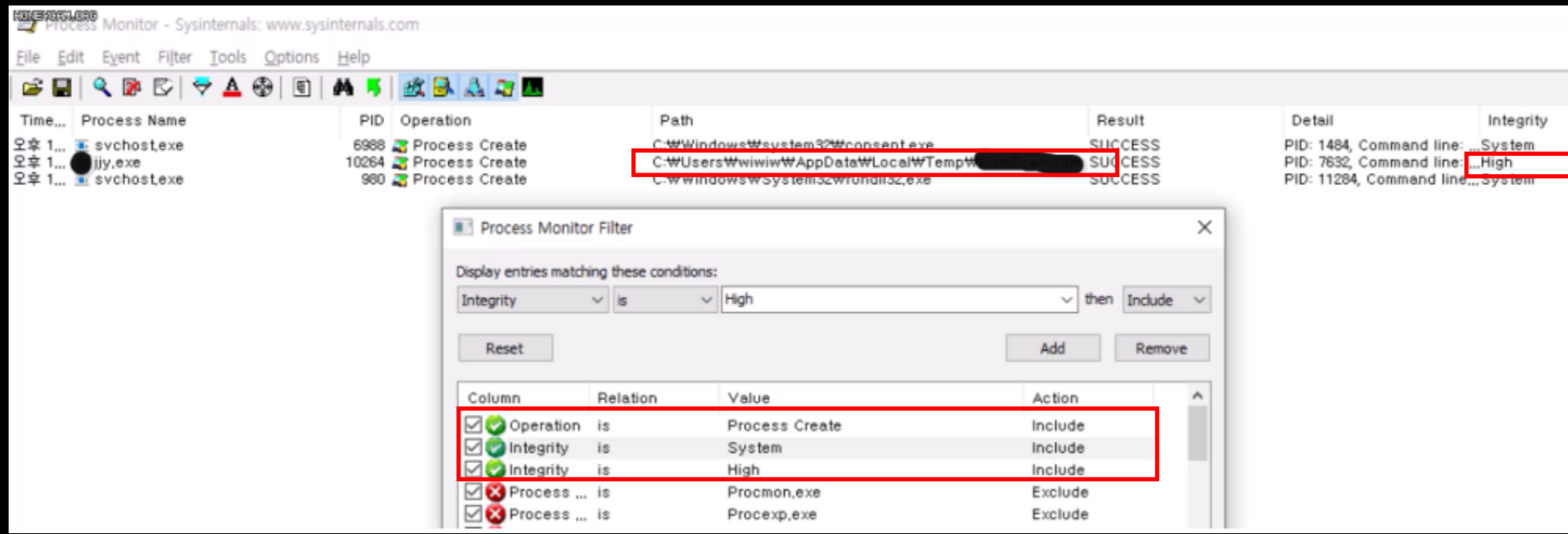
- LPE 취약점을 쉽게 찾을 수 있음.
- Target
 - Root process - linux
 - System(admin) process - windows

Monitoring tools

- 타겟 프로세스가 하는 행위를 관찰
 - Execute file
 - Read file
 - Write file
 - Load Library

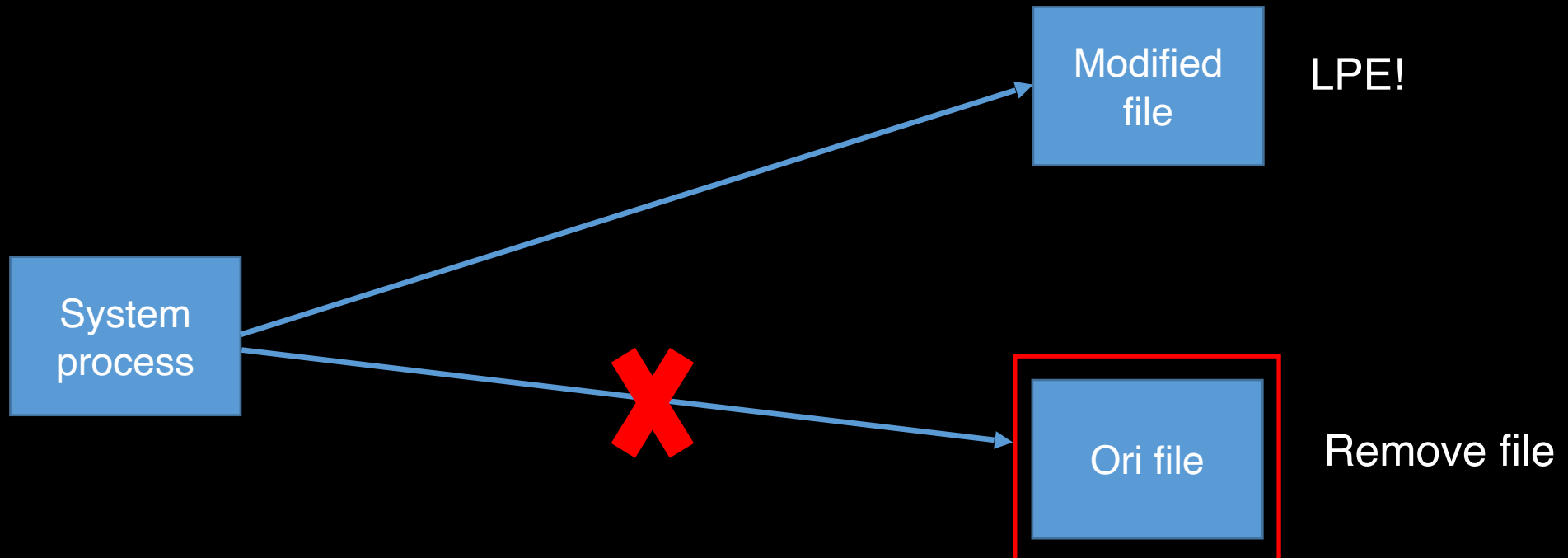
Monitoring tools

- **Execute file** at writeable file



Monitoring tools

- **Execute file** at writeable file
 - **Just replace exe file!**



Monitoring tools

- **Read file** from writeable file
 - 해당 벡터를 공격 벡터로 사용 가능해짐.

[illegible]

Monitoring tools

Write file at writeable folder

Symbolic link attack

Root 권한으로 파일을 쓸 수 있게 됨.

Link to '/etc/passwd', '/root/.bashrc'

```
→ test ls -al
total 12
drwxrwxrwx 2 root root 4096 Jun  6 01:40 .
drwxrwxr-x 3 jjy  jjy 4096 Jun  6 01:36 ..
-rw-r--r-- 1 root root   3 Jun  6 01:40 read_only
→ test rm read_only
rm: remove write-protected regular file 'read_only'? y
→ test ls -al
total 8
drwxrwxrwx 2 root root 4096 Jun  6 01:40 .
drwxrwxr-x 3 jjy  jjy 4096 Jun  6 01:36 ..
→ test
```


- **Load Library** from writeable path
 - **DLL Hijacking!!! - windows**

- **Load Library** from writeable path
 - **DLL Hijacking!!! - windows**

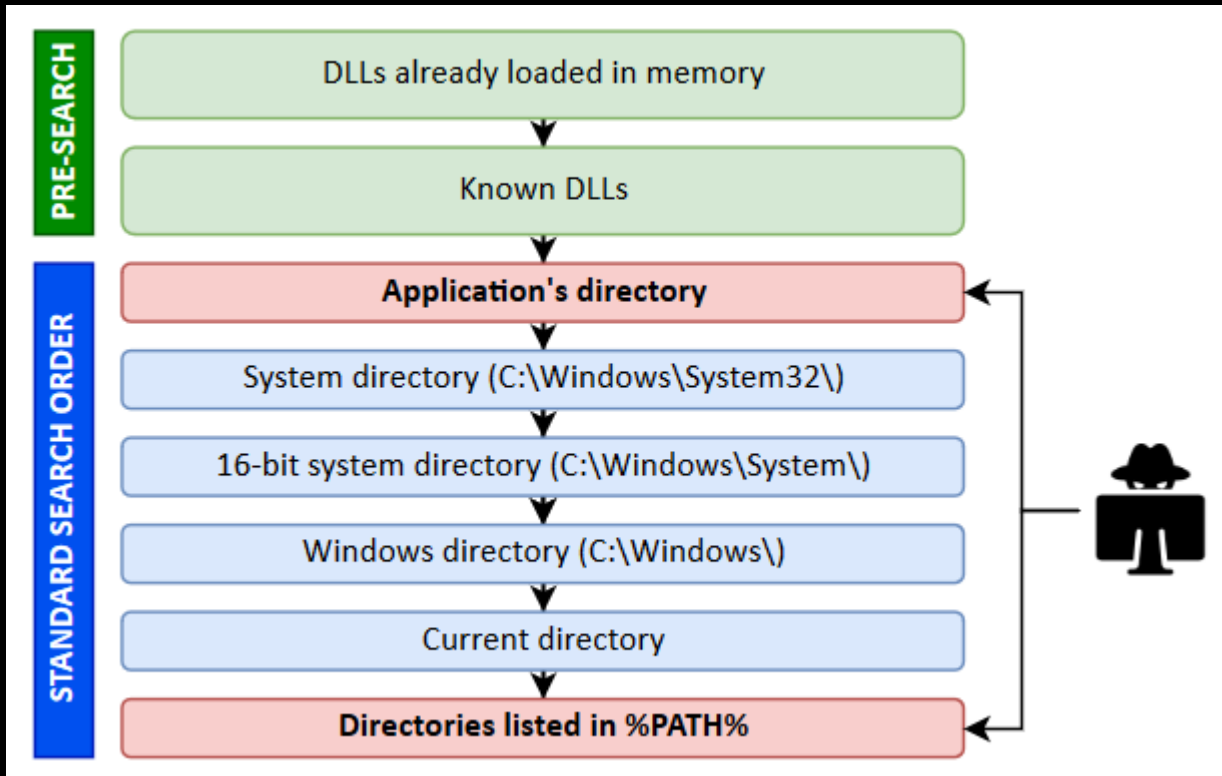
The screenshot displays the Process Monitor application. The main window shows a list of processes and their operations. A filter dialog box is open, showing the following conditions:

Column	Relation	Value	Action
<input checked="" type="checkbox"/> Operation	is	createfile	Include
<input checked="" type="checkbox"/> Path	contains	wiwiw	Include
<input checked="" type="checkbox"/> Integrity	is	System	Include

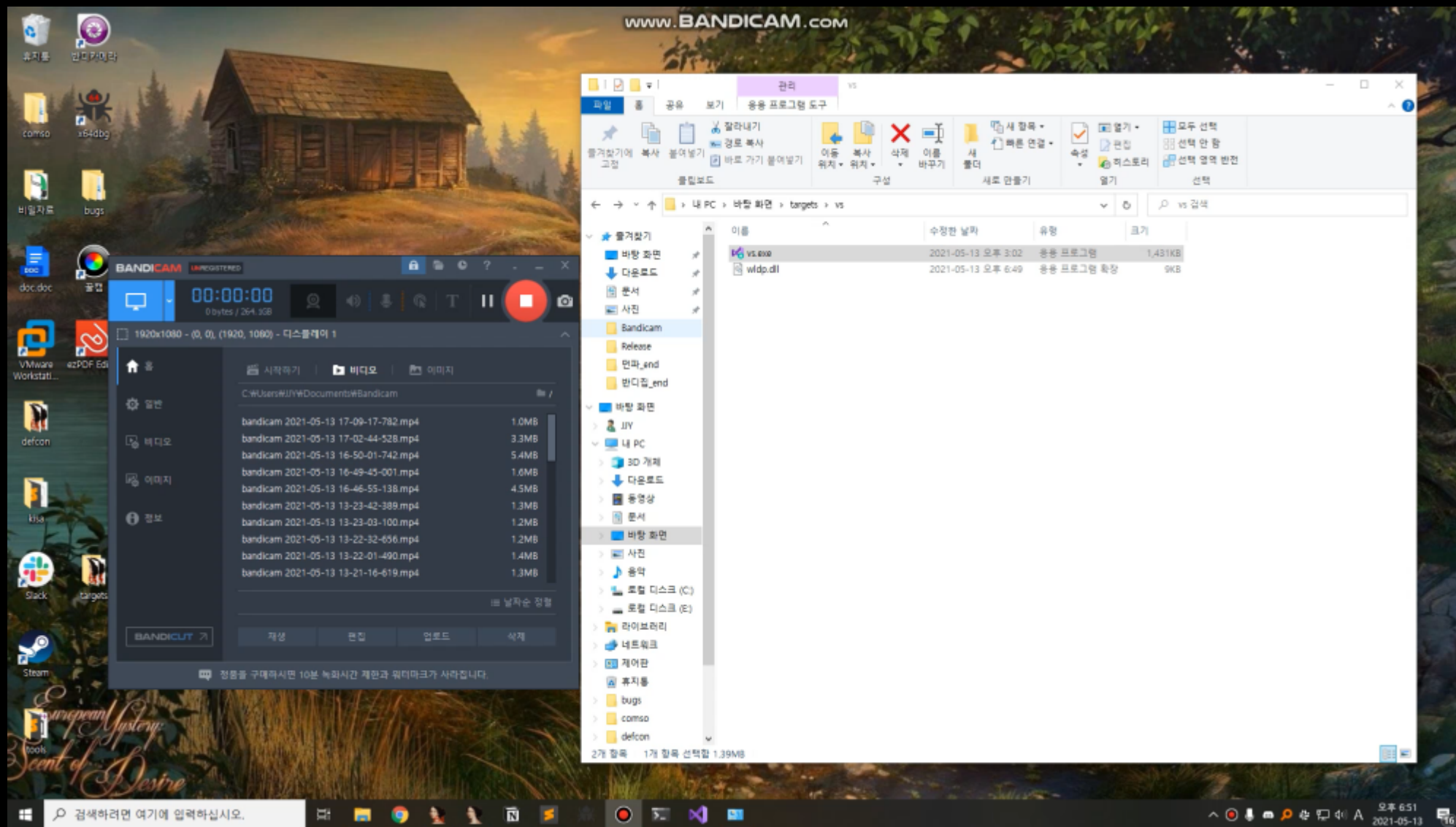
Below the filter dialog, the main window shows a list of processes and their operations. The filter dialog is currently set to display entries matching the conditions: Path excludes dll then Exclude.

Monitoring tools

- **Load Library** from writeable path
- **DLL Hijacking!!! - windows**



DEMO



Find SW vulnerability via automatic method