

Modern Fuzzing Research & Engineering

Andrea Fioraldi

[@andrea Fioraldi](https://twitter.com/andrea Fioraldi)



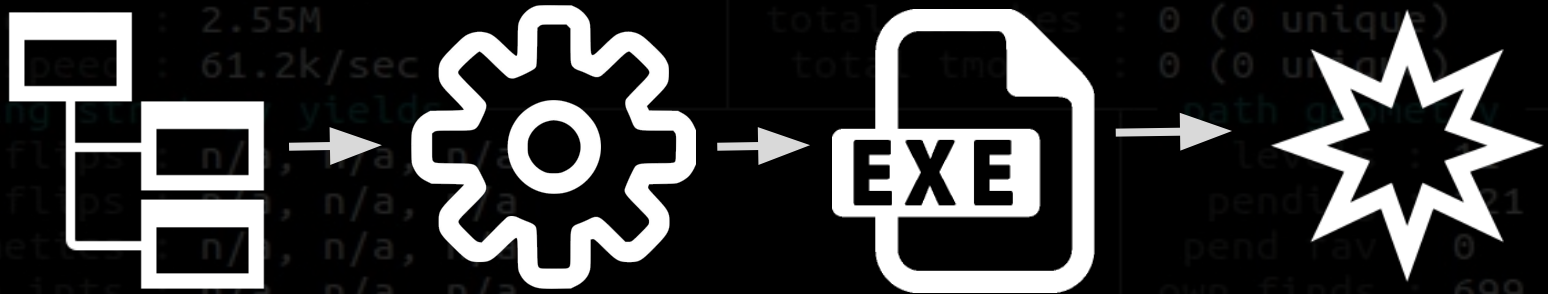
What is Fuzz Testing?

Fuzz Testing, or Fuzzing, is a family of “Software” Testing techniques that involves providing machine-generated inputs to the System Under Test (SUT) in order to satisfy some objectives.

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  raw fuzz : 0 days, 0 hrs, 0 min, 1 sec
last uniq crash : none seen yet
last uniq hang  : none seen yet
stage progress
  now trying : splice 14
  stage execs : 31/32 (96.88%)
  total execs : 2.55M
  exec speed  : 61.2k/sec
  fuzzing strategy yields
    bit flips : n/a, n/a, n/a
    byte flips : n/a, n/a, n/a
    arithmetics : n/a, n/a, n/a
    known ints : n/a, n/a, n/a
    dictionary : n/a, n/a, n/a
  havoc/splice : 506/1.05M, 193/1.44M
    py/custom  : 0/0, 0/0
    trim       : 19.25%/53.2k, n/a
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs  : 0
  map coverage : 0.00%
  crash coverage : 3.30 bits/tuple
  findings in depth
    favored paths : 114 (16.22%)
    new edges on  : 167 (23.76%)
    total crashes : 0 (0 unique)
    total tmouts  : 0 (0 unique)
  path geometry
    levels : 11
    pending : 121
    pend fav : 0
    own finds : 699
    imported  : n/a
  stability : 99.88%
[cpu000: 12%]
```

What is Fuzz Testing?

Fuzz Testing, or Fuzzing, is a family of “Software” Testing techniques that involves providing machine-generated inputs to the System Under Test (SUT) in order to satisfy some objectives.



What is Fuzz Testing?

Machine-generated inputs can be of any kind, beyond the classic definition of “unexpected” (by the way, what does it means?) inputs.

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  raw fuzz : 0 days, 0 hrs, 0 min, 1 sec
last uniq crash : none seen yet
last uniq hang  : none seen yet
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs  : 0
paths timed out : 0 (0.00%)
stage progress
  now trying : splice 14
  stage execs : 31/32 (96.88%)
  total execs : 2.55M
  exec speed  : 61.2k/sec
  findings in depth
    favored paths : 114 (16.22%)
    new edges on  : 167 (23.76%)
    total crashes : 0 (0 unique)
    total tmouts  : 0 (0 unique)
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints  : n/a, n/a, n/a
  dictionary  : n/a, n/a, n/a
  havoc/splice : 506/1.05M, 193/1.44M
  py/custom   : 0/0, 0/0
  trim        : 19.25%/53.2k, n/a
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported  : n/a
  stability : 99.88%
count coverage : 3.30 bits/tuple
[cpu000: 12%]
```

What is Fuzz Testing?

Machine-generated inputs can be of any kind, beyond the classic definition of “unexpected” (by the way, what does it means?) inputs.

Fuzzing is often considered related to Random Testing, a technique that provides inputs sampled uniform independently from the input space (using a specification maybe, so they are not random bytes in general).

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  raw fuzzes, 1 sec
last uniq crash : none seen yet
last uniq hang : none seen yet
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs : 0
paths timed out : 0 (0.00%)
count coverage : 3.30 bits/tuple
  favored paths : 114 (16.22%)
  splice 14
  splice 31/12 (36.80%)
  splice 10/6 (25.75%)
  splice 0/0 (0.00%)
  total inouts : 0 (0 unique)
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
  havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
[cpu000: 12%]
```

What is Fuzz Testing?

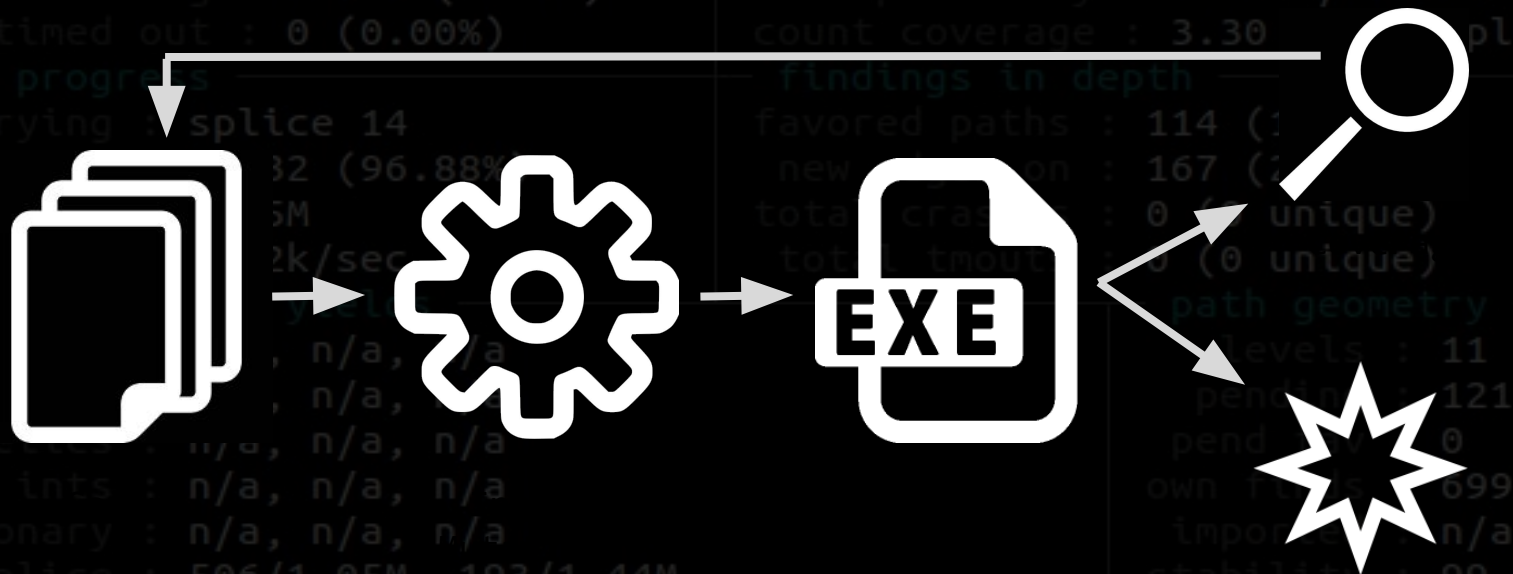
Machine-generated inputs can be of any kind, beyond the classic definition of “unexpected” (by the way, what does it means?) inputs.

Fuzzing is often considered related to Random Testing, a technique that provides inputs sampled uniform independently from the input space (using a specification maybe, so they are not random bytes in general).

But Fuzzing can generate inputs deterministically, or can generate inputs mutating previously generated inputs that makes the sampling from the input space not independent.

Widely discussed, SOTA Fuzzing in 2022

- Feedback-driven, mainly Coverage-guided



Widely discussed SOTA Fuzzing in 2022

- Feedback-driven, mainly Coverage-guided
- Can bypass coverage roadblocks (concolic-aided, taint-assisted, RedQueen, ...)

```
if (input == 0xabadcafe) {  
    interesting_code();  
}
```

```
overall results  
cycles done : 15  
total time : 703  
uniq crashes : 0  
uniq hangs : 0
```

```
map density : 5.78% / 13.98%  
findings in depth  
favored paths : 114 (16.22%)  
new edges on : 167 (23.76%)  
total crashes : 0 (0 unique)  
total tmouts : 0 (0 unique)
```

```
path geometry  
levels : 11  
pending : 121  
pend fav : 0  
own finds : 699  
imported : n/a  
stability : 99.88%
```

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}  
process timing  
run time : 0 days, 0 hrs, 0 min, 43 sec  
last uniq crash : none seen yet  
last uniq hang : none seen yet  
cycle progress  
now processing : 261*1 (37.1%)  
paths  
stage progress  
now doing splice 14  
stage execs : 31/32 (96.88%)  
total execs : 2.55M  
exec speed : 61.2k/sec  
fuzzing strategy  
bit flips : n/a, n/a, n/a  
byte flips : n/a, n/a, n/a  
arithmetics : n/a, n/a, n/a  
known ints : n/a, n/a, n/a  
dictionary : n/a, n/a, n/a  
havoc/splice : 506/1.05M, 193/1.44M  
py/custom : 0/0, 0/0  
trim : 19.25%/53.2k, n/a
```

```
[cpu000: 12%]
```

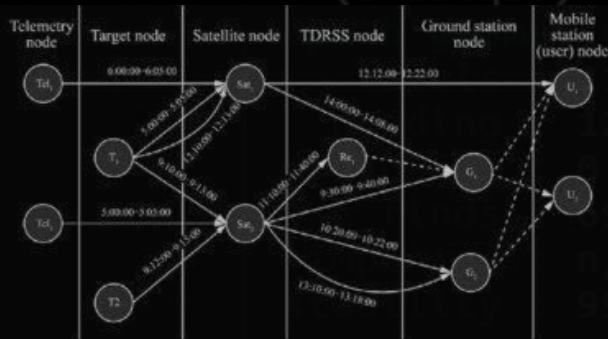

Widely discussed SOTA Fuzzing in 2022

- Feedback-driven, mainly Coverage-guided
- Can bypass coverage roadblocks (concolic-aided, taint-assisted, RedQueen, ...)
- Input models help to fuzz deeper

```
<start> ::= <expr>
<expr> /a, ::= <term> + <expr> | <term> - <expr> | <term>
<term> /a, ::= <term> * <factor> | <term> / <factor> | <factor>
<factor> ::= +<factor> | -<factor> | (<expr>) | <integer> |
<integer>.<integer>
<integer> ::= <digit><integer> | <digit>
<digit> ::= 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9
```

Widely discussed SOTA Fuzzing in 2022

- Feedback-driven, mainly Coverage-guided
- Can bypass coverage roadblocks (concolic-aided, taint-assisted, RedQueen, ...)
- Input models help to fuzz deeper
- Can test network interactions



Widely discussed SOTA Fuzzing in 2022

- Feedback-driven, mainly Coverage-guided
- Can bypass coverage roadblocks (concolic-aided, taint-assisted, RedQueen, ...)
- Input models help to fuzz deeper
- Can test network interactions
- Can fuzz userspace programs, kernel, hypervisors, ...

```
american_fuzzy_lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  cycles done : 15
  total crashes : 703
  last uniq crash : none seen yet
  last uniq hang : none seen yet
  cycle progress :
  now processing : 261*1 (37.1%)
  pairs :
  stage progress :
  now running : splice 14
  stage execs : 31/32 (96.88%)
  total execs : 2.55M
  exec time : 51.21/51c
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetic : n/a, n/a, n/a
  known ones : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
overall results
  cycles done : 15
  total crashes : 703
  last uniq crash : none seen yet
  last uniq hang : none seen yet
  cycle progress :
  now processing : 261*1 (37.1%)
  pairs :
  stage progress :
  now running : splice 14
  stage execs : 31/32 (96.88%)
  total execs : 2.55M
  exec time : 51.21/51c
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetic : n/a, n/a, n/a
  known ones : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
map density : 5.78% / 13.98%
findings in depth
  favored paths : 114 (16.22%)
  new edges on : 167 (23.76%)
  total crashes : 0 (0 unique)
  total tmouts : 0 (0 unique)
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
[cpu000: 12%]
```


Widely used tools in 2022



LLVM's Libfuzzer



NYX

google/honggfuzz 

Security oriented software fuzzer. Supports evolutionary, feedback-driven fuzzing based on code coverage (SW and HW based)

👤 68 Contributors 🕒 17 Issues ⭐ 3k Stars 🍴 492 Forks

```
american fuzzy lop ++2.65d (1
```

```
process timing
```

```
run time : 0 days, 0 hrs, 0 mi
```

```
last new path : 0 days, 0 hrs, 0 mi
```

```
last uniq crash : none seen yet
```

```
last uniq hang : none seen yet
```

```
cycle progress
```

```
now processing : 261*1 (37.1%)
```

```
paths found out : (0.00%)
```

```
stageless
```

```
now trying : splice 14
```

```
stageless : 31/32 (96.88%)
```

```
total paths : 55M
```

```
exec speed : 61.2k/sec
```

```
fuzzing strategy yields
```

```
bit flips : n/a, n/a, n/a
```

```
byte flips : n/a, n/a, n/a
```

```
arithmetics : n/a, n/a, n/a
```

```
known ints : n/a, n/a, n/a
```

```
dictionary : n/a, n/a, n/a
```

```
havoc/splice : 506/1.05M, 193/1.44M
```

```
py/custom : 0/0, 0/0
```

```
trim : 19.25%/53.2k, n/a
```

We still
miss bugs

Yes, even in heavily-fuzzed
projects in OSS-Fuzz

Still finding these bugs by hand...

Project Zero

News and updates from the Project Zero team at Google

Wednesday, December 1, 2021

This shouldn't have happened: A vulnerability postmortem

Posted by Tavis Ormandy, Project Zero

Introduction

This is an unusual blog post. I normally write posts to highlight some hidden attack surface or interesting complex vulnerability class. This time, I want to talk about a vulnerability that is neither of those things. The striking thing about this vulnerability is just how simple it is. This should have been caught earlier, and I want to explore why that didn't happen.

In 2021, all good bugs need a catchy name, so I'm calling this one "BigSig".

First, let's take a look at the bug, I'll explain how I found it and then try to understand why we missed it for so long.

Still finding these bugs by hand...

Issue 2272: libxml2: heap-buffer-overflow in xmlBufAdd

Reported by fwilhelm@google.com on Tue, Mar 8, 2022, 4:19 PM GMT+1

Project Member

libxml2 is vulnerable to a heap-buffer-overflow when xmlBufAdd is called on a very large buffer:

...

int

```
xmlBufAdd(xmlBufPtr buf, const xmlChar *str, int len) {  
    unsigned int needSize;
```

```
    [..]
```

```
    needSize = buf->use + len + 2; (A)
```

```
    if (needSize > buf->size){
```

```
        [..]
```

```
        if (!xmlBufResize(buf, needSize)){
```

```
            xmlBufMemoryError(buf, "growing buffer");
```

```
            return XML_ERR_NO_MEMORY;
```

```
        }
```

```
    }
```

eam at Google

A vulnerability postmortem

s to highlight some hidden attack surface or interesting
lk about a vulnerability that is neither of those things. The
nple it is. This should have been caught earlier, and I want

In 2021, all good bugs need a catchy name, so I'm calling this one "BigSig".

First, let's take a look at the bug, I'll explain how I found it and then try to understand why we missed it for so long.

Why?

- Fuzzers often tests only the default configuration
- Fuzzers have input length limits
- Code coverage as feedback is not enough

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  new path : 0 days, 0 hrs, 0 min, 1 sec
last uniq crash : none seen yet
last uniq hang  : none seen yet
cycle progress
  now processing : 261*1 (37.1%)
paths
  now solving : splice_14
stage progress
  stage execs : 31/32 (96.88%)
total execs : 2.55M
exec speed : 61.2k/sec
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs : 0
map coverage
  map density : 5.78% / 13.98%
count coverage : 3.30 bits/tuple
findings in depth
  favored paths : 114 (16.22%)
  new edges on : 167 (23.76%)
total crashes : 0 (0 unique)
total tmouts : 0 (0 unique)
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
[cpu000: 12%]
```

Why?

- Fuzzers often tests only the default configuration
- Fuzzers have input length limits
- Code coverage as feedback is not enough (beware of path explosion!)
 - Fioraldi, D'Elia, Balzarotti. "The Use of Likely Invariants as Feedback for Fuzzers"
 - Mantovani, Fioraldi, Balzarotti. "Fuzzing with Data Dependency Information"
 - Herrera, Payer, Hosking. "DATAFLOW - Towards a Data-Flow-Guided Fuzzer"

An Example

```
int wavlike_msadpcm_init (SF_PRIVATE *psf, int blockalign, int samplesperblock)
{
    MSADPCM_PRIVATE *pms;
    unsigned int pmssize;

    ...
    pmssize = sizeof (MSADPCM_PRIVATE) + blockalign + 3 * psf->sf.channels * samplesperblock;
    ;
    ...
    pms->samples = pms->dummysdata; // array in pms
    pms->block = (unsigned char*) (pms->dummysdata + psf->sf.channels * samplesperblock);
    pms->channels = psf->sf.channels;
    pms->blocksize = blockalign;
    ...
}
```

```
overall results
cycles done : 15
total paths : 703
uniq crashes : 0
uniq hangs : 0
```

```
map coverage
map density : 5.78% / 13.98%
count coverage : 3.30 bits/tuple
```

```
findings in depth
new edges on : 167 (23.76%)
total crashes : 0 (0 unique)
```

```
path geometry
levels : 11
pending : 121
pend fav : 0
own finds : 699
imported : n/a
stability : 99.88%
```

[cpu000: 12%]

An Example

```
int wavlike_msadpcm_init (SF_PRIVATE *psf, int blockalign, int samplesperblock)
{
    MSADPCM_PRIVATE *pms;
    unsigned int pmssize;
    ...
    pmssize = sizeof (MSADPCM_PRIVATE) + blockalign + 3 * psf->sf.channels * samplesperblock
    ;
    ...
    pms->samples = pms->dummydata ; // array in pms
    pms->block = (unsigned char*) (pms->dummydata + psf->sf.channels * samplesperblock) ;
    pms->channels = psf->sf.channels ;
    pms->blocksize = blockalign ;
    ...
}
```

An Example

```
static int msadpcm_decode_block (SF_PRIVATE *psf, MSADPCM_PRIVATE *pms)
```

```
{  
    ...  
    sampleindx = 2 * pms->channels ;
```

```
    while (blockindx < pms->blocksize)
```

```
    {  
        bytecode = pms->block [blockindx++] ;
```

```
        pms->samples [sampleindx++] = (bytecode >> 4) & 0x0F ; // heap overflow bug
```

```
        pms->samples [sampleindx++] = bytecode & 0x0F ;  
    }  
}
```

```
overall results
```

```
cycles done : 15  
total paths : 703  
uniq crashes : 0  
uniq hangs : 0
```

```
map coverage
```

```
map density : 5.78% / 13.98%  
count coverage : 3.30 bits/tuple
```

```
findings in depth
```

```
favored paths : 114 (16.22%)  
new edges on : 167 (23.76%)
```

```
total edges : 0 (0 unique)  
total paths : 0 (0 unique)
```

```
path geometry
```

```
levels : 11  
pending : 121  
pend fav : 0  
own finds : 699  
imported : n/a  
stability : 99.88%
```

```
[cpu000: 12%]
```


An Example

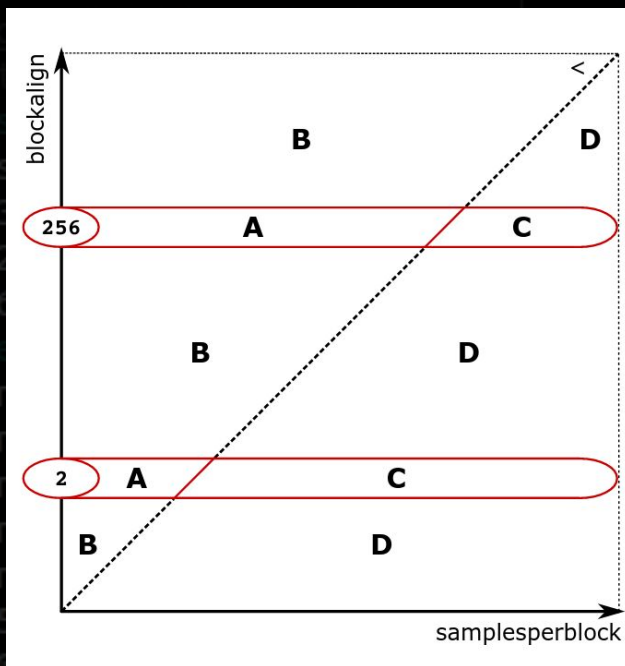
```
static int msadpcm_decode_block (SF_PRIVATE *psf, MSADPCM_PRIVATE *pms)
{
    ...
    sampleindx = 2 * pms->channels ;
    while (blockindx < pms->blocksize)
    {
        bytecode = pms->block [blockindx++] ;
        pms->samples [sampleindx++] = (bytecode >> 4) & 0x0F ; // heap overflow bug
        pms->samples [sampleindx++] = bytecode & 0x0F ;
    }
    ...
}
```

This only happens when the program is in a specific state, characterized by a small allocation size for the pms buffer and a pms->blocksize value sufficiently high to force the loop to write out of the bounds of the array.

However, none of these requirements can be extracted from code coverage, as there are no branches in the program that involve these thresholds

An Example

```
pmssize = sizeof (MSADPCM_PRIVATE) + blockalign + 3 * psf->sf.channels * samplesperblock
```



Invariant	Condition
LI ₁	blockalign ∈ {0,2,256}
LI ₂	blockalign < samplesperblock

Invariant	A	B	C	D
LI ₁	✓	✗	✓	✗
LI ₂	✗	✗	✓	✓

88%

[cpu000: 12%]

Why?

- Fuzzers often tests only the default configuration
- Fuzzers have input length limits
- Code coverage as feedback is not enough
- Harnessing to cover all the code is hard (especially for devs)

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  new path : 0 days, 0 hrs, 0 min, 1 sec
last uniq crash : none seen yet
last uniq hang  : none seen yet
cycle progress
  now processing : 261*1 (37.1%)
  paths          : 306
stage progress
  now solving    : splice_14
stage execs    : 31/32 (96.88%)
total execs   : 2.55M
exec progress  : 1.2k/sec
fuzzing strategy yields
  bit flips   : n/a, n/a, n/a
  byte flips  : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints  : n/a, n/a, n/a
  dictionary  : n/a, n/a, n/a
havoc/splice  : 506/1.05M, 193/1.44M
  py/custom   : 0/0, 0/0
  trim        : 19.25%/53.2k, n/a
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs  : 0
map coverage
  map density : 5.78% / 13.98%
count coverage : 3.30 bits/tuple
findings in depth
  favored paths : 114 (16.22%)
  new edges on : 167 (23.76%)
  total crashes : 0 (0 unique)
  total hangs   : 0 (0 unique)
path geometry
  levels      : 11
  pending    : 121
  pend fav   : 0
  own finds  : 699
  imported   : n/a
  stability  : 99.88%
[cpu000: 12%]
```

Why?

- Harnessing to cover all the code is hard (especially for devs)

- We can generate them automatically

- Ispoglou, Austin, Mohan, Payer. “FuzzGen: Automatic Fuzzer Generation”
- Babić, Bucur, Chen, Ivančić, King, Lemieux, Szekeres, Wang. “FUDGE: Fuzz Driver Generation at Scale”

```
american_fuzzy_lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  new path : 0 days, 0 hrs, 0 min, 1 sec
last uniq crash : none seen yet
last uniq hang : none seen yet
cycle progress
  now processing : 261*1 (37.1%)
paths
  stage progress
  now trying : splice 14
stage execs : 31/32 (96.88%)
total execs : 2.55M
exec speed : 61.2k/sec
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs : 0
map density : 5.78% / 13.98%
count coverage : 3.30 bits/tuple
findings in depth
  favored paths : 114 (10.22%)
  new edges on : 167 (23.76%)
  total crashes : 0 (0 unique)
  total tmouts : 0 (0 unique)
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
[cpu000: 12%]
```

Why?

- Harnessing to cover all the code is hard (especially for devs)
 - We can generate them automatically
 - We need introspection of what the fuzzer can cover
 - Fuzz Introspector (<https://github.com/ossf/fuzz-introspector>)

```
american_fuzzy_lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  new path : 0 days, 0 hrs, 0 min, 1 sec
last uniq crash : none seen yet
last uniq hang  : none seen yet
cycle progress
  now processing : 261*1 (37.1%)
paths
  now trying : splice_14
stage execs   : 31/32 (96.88%)
total execs   : 2.55M
exec speed    : 61.2k/sec
fuzzing strategy yields
  bit flips   : n/a, n/a, n/a
  byte flips  : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints  : n/a, n/a, n/a
  dictionary  : n/a, n/a, n/a
havoc/splice  : 506/1.05M, 193/1.44M
  py/custom   : 0/0, 0/0
  trim        : 19.25%/53.2k, n/a
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs  : 0
map density   : 5.78% / 13.98%
count coverage : 3.30 bits/tuple
findings in depth
  avoided paths : 114 (16.22%)
  new edges on  : 167 (23.76%)
  total crashes : 0 (0 unique)
  total tmouts  : 0 (0 unique)
path geometry
  levels       : 11
  pending     : 121
  pend fav    : 0
  own finds   : 699
  imported    : n/a
  stability   : 99.88%
[cpu000: 12%]
```


Why?

- Harnessing to
 - We can generate
 - We need intro
- Fuzz Int

The screenshot shows the Immunity Debugger interface. The main window displays the disassembly of a function named `int64_t sub_1400012b0(void* arg1)`. The disassembly view shows assembly instructions with their corresponding hex values and comments. The Coverage Overview window on the right provides a summary of the function's execution, including coverage percentage, function name, address, blocks hit, instructions hit, function size, and code coverage (CC).

Cov %	Func Name	Address	Blocks Hit	Instr. Hit	Func Size	CC
0.00	sub_140001000	0x140001000	0 / 1	0 / 6	32	1
0.00	sub_140001030	0x140001030	0 / 5	0 / 12	35	4
69.37	sub_140001060	0x140001060	5 / 10	32 / 46	149	6
0.00	sub_140001100	0x140001100	0 / 5	0 / 17	72	4
24.00	sub_140001150	0x140001150	2 / 3	6 / 25	136	1
0.00	sub_1400011e0	0x1400011e0	0 / 5	0 / 15	46	2
0.00	sub_140001210	0x140001210	0 / 3	0 / 9	26	1
0.00	sub_140001230	0x140001230	0 / 3	0 / 10	29	1
0.00	sub_140001250	0x140001250	0 / 5	0 / 23	81	3
88.94	sub_1400012b0	0x1400012b0	17 / 19	191 / 215	1074	9
77.25	sub_1400012e0	0x1400012e0	11 / 13	96 / 122	362	7
84.93	sub_140001930	0x140001930	4 / 10	79 / 93	403	5
59.42	sub_140001b20	0x140001b20	13 / 16	93 / 106	435	8
42.96	sub_140001d10	0x140001d10	11 / 16	69 / 108	446	6
43.75	sub_140001e00	0x140001e00	7 / 16	42 / 96	393	6
88.88	sub_140002080	0x140002080	7 / 8	48 / 54	248	2
88.88	sub_140002160	0x140002160	7 / 8	47 / 53	242	2
56.33	sub_1400022e0	0x1400022e0	11 / 16	82 / 95	436	8
54.43	sub_140002420	0x140002420	8 / 11	43 / 59	359	5
100.00	sub_140002590	0x140002590	4 / 4	42 / 42	322	3
0.00	sub_1400026e0	0x1400026e0	0 / 3	0 / 8	28	1
0.00	sub_140002700	0x140002700	0 / 9	0 / 53	167	5
78.03	sub_1400027c0	0x1400027c0	14 / 19	66 / 82	266	9
35.28	sub_1400028d0	0x1400028d0	4 / 17	24 / 66	247	8
65.35	sub_1400029d0	0x1400029d0	4 / 9	17 / 26	87	3
65.35	sub_140002a60	0x140002a60	4 / 9	17 / 26	87	3
0.00	sub_140002af0	0x140002af0	0 / 1	0 / 17	81	1
100.00	sub_140002b20	0x140002b20	1 / 1	31 / 31	169	1

Selection: 0x1400012b0 to 0x1400012b2 (0x2 bytes) PE * Graph * Options *

stability : 99.88%

[cpu000: 12%]


```
american fuzzy lop ++2.65d (1
process timing
  run time : 0 days, 0 hrs, 0 mi
  last new path : 0 days, 0 hrs, 0 mi
last uniq crash : none seen yet
last uniq hang : none seen yet
cycle pro
now processi
paths timed out : 0 (0.00%)
stage pr
now trying : splice 14
stage execs : 31/32 (96.88%)
total
exec speed : 61.21/sec
fuzzing strategy yields
bit flips : n/a, n/a, n/a
byte flips : n/a, n/a, n/a
arithmetics : n/a, n/a, n/a
known ints : n/a, n/a, n/a
dictionary : n/a, n/a, n/a
havoc/splice : 506/1.05M, 193/1.44M
py/custom : 0/0, 0/0
trim : 19.25%/53.2k, n/a
```

Beyond memory corruption bugs

A SQL injection is not causing
a segfault in your application

Several paths are SOTA

- Differential fuzzing

- Cryptofuzz (<https://github.com/guidovranken/cryptofuzz>)

- Maier, Fäßler, Seifert. “Uncovering Smart Contract VM Bugs Via Differential Fuzzing”

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  total time : 6 days, 6 hrs, 0 min, 11 sec
last uniq crash : none seen yet
last uniq hang : none seen yet
cycle progress
  now processed : 261*1 (37.1%)
paths timed out : 0 (0.00%)
stage progress
  now trying splice 14
stage execs : 31/32 (96.88%)
total execs : 2.55M
exec speed : 61.2k/sec
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
map coverage
  map density : 5.78% / 13.98%
count coverage : 3.30 bits/tuple
  findings in depth
  favored paths : 114 (16.22%)
  new edges on : 167 (23.76%)
total crashes : 0 (0 unique)
total tmouts : 0 (0 unique)
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs : 0
[cpu000: 12%]
```

Several paths are SOTA

- Differential fuzzing

- Cryptofuzz (<https://github.com/guidovranken/cryptofuzz>)
- Maier, Fäßler, Seifert. “Uncovering Smart Contract VM Bugs Via Differential Fuzzing”

- Custom bug detectors

- Handwritten bug detectors, useful for memory safe languages (e.g. Java <https://www.code-intelligence.com/blog/log4j-bug-detectors>)
- Custom sanitizers (e.g. <https://bugs.chromium.org/p/oss-fuzz/issues/detail?id=49053>)
- Mining invariants and automatic insertion of assertions

- Daikon, Purify, 1.93/1.44M

Several paths are SOTA

- Differential fuzzing

- Cryptofuzz (<https://github.com/guidovr>)
- Maier, Fäßler, Seifert. “Uncovering Smi Fuzzing”

- Custom bug detectors

- Handwritten bug detectors, useful for r <https://www.code-intelligence.com/blog>,
- Custom sanitizers (e.g. [https://bugs.chromium.org/p/oss-fuzz/i](https://bugs.chromium.org/p/oss-fuzz/issues/detail?id=1131)),
- Mining invariants and automatic insert:
 - Daikon, Purify, ...



Oliver Chang
@halbecaf

Proof that fuzzing can discover exploitable vulnerabilities that aren't memory corruption! OSS-Fuzz discovered a very interesting command injection vulnerability which was just fixed:

syoyo/tinygltf

#368 Command injection via wordexp call.

5 comments

oliverchang opened on August 16, 2022

github.com

Command injection via wordexp call. · Issue #368 · syoyo/tinygltf

Describe the issue This is a security vulnerability. The wordexp call here allows arbitrary code execution tinymce/tinymce.h Line 2640 in Ofa56e2 int ret = ...

Can we do better?

- Improve invariants mining, the coverage problem causes too many false positive and locally valid constraints unsuitable for fuzzing

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  compile time : 0 days, 0 hrs, 0 min, 1 sec
last uniq crash : none seen yet
last uniq hang : none seen yet
cycle progress
  now processing : 261*1 (37.1%)
paths timed out : 0 (0.00%)
stage progress
  now trying : splice 14
stage execs : 31/32 (96.88%)
total execs : 2.55M
exec speed : 61.2k/sec
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs : 0
  map coverage : 5.78% / 13.98%
  count coverage : 3.30 bits/tuple
findings in depth
  favored paths : 114 (16.22%)
  new edges on : 167 (23.76%)
  total crashes : 0 (0 unique)
  total tmouts : 0 (0 unique)
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
[cpu000: 12%]
```

Can we do better?

- Improve invariants mining, the coverage problem causes too many false positive and locally valid constraints unsuitable for fuzzing
- Build large databases of bug patterns (?)

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  time left : 0 days, 0 hrs, 0 min, 1 sec
last uniq crash : none seen yet
last uniq hang  : none seen yet
cycle progress
  now processing : 261*1 (37.1%)
paths timed out : 0 (0.00%)
stage progress
  now trying : splice 14
stage progress
  now trying : splice 14
total execs : 2.55M
exec speed : 61.2k/sec
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs : 0
  map coverage : 5.78% / 13.98%
  count coverage : 3.30 bits/tuple
findings in depth
  favored paths : 114 (16.22%)
  paths on : 167 (23.76%)
  total crashes : 0 (0 unique)
  total tmouts : 0 (0 unique)
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
[cpu000: 12%]
```


Can we do better?

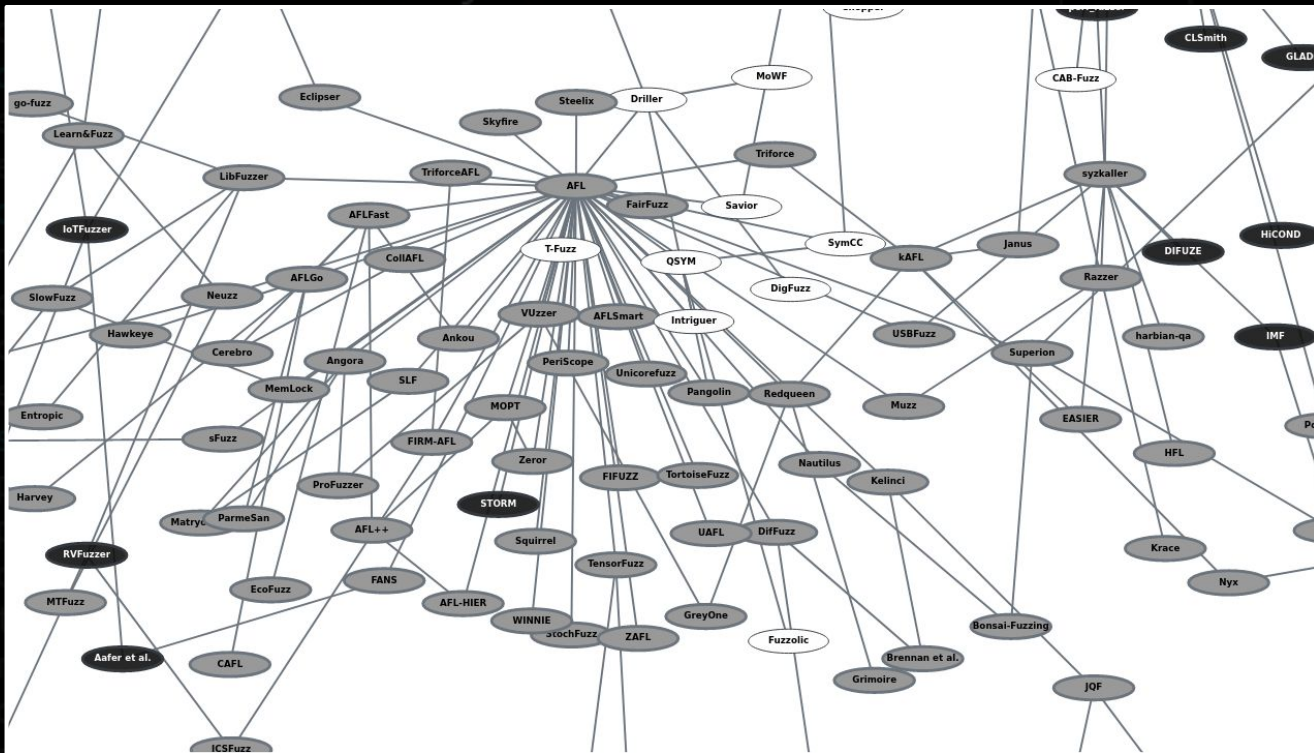
- Improve invariants mining, the coverage problem causes too many false positive and locally valid constraints unsuitable for fuzzing
- Build large databases of bug patterns (?)
- Maybe it's time to start approaching program analysis problems with ML without the “wanna find something to apply this model” bias

```
american fuzzy lop ++2.65d (1
process timing
  run time : 0 days, 0 hrs, 0 mi
  last new path : 0 days, 0 hrs, 0 mi
last uniq crash : none seen yet
last uniq hang : none seen yet
  cycles : 261*1 (57.1%)
paths tried out : 0 (0.00%)
stage execs : 31/32 (96.88%)
  havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
```

Wanna build
a fuzzer and
compare with
the others?

Good luck.

Problem: Fuzzers Fragmentation



From <https://fuzzing-survey.org/>

Cause: Monolithic Codebases

Fuzzers are

⇒ Designed to be tools

⇒ Not designed with code reuse in mind

⇒ Hard to extend

Many fuzzers are incompatible forks of others (usually AFL)

This makes them incompatible with orthogonal techniques

```
american_fuzzy_lop ++2.65d (libpng_harness) [explore] {0}
```

```
process timing
```

```
run time : 0 days, 0 hrs, 0 min, 43 sec
```

```
last uniq crash : none seen yet
```

```
last uniq hang : none seen yet
```

```
cycle progress
```

```
now processing : 261*1 (37.1%)
```

```
path found out : 0 (0.00%)
```

```
stage progress
```

```
now trying : splice 14
```

```
stage execs : 196 (0.00%)
```

```
total execs : 2.55M
```

```
exec speed : 61.2k/sec
```

```
fuzzer yields
```

```
bit flips : n/a, n/a, n/a
```

```
byte flips : n/a, n/a, n/a
```

```
arithmetic : n/a, n/a, n/a
```

```
dictionary : n/a, n/a, n/a
```

```
havoc/splice : 506/1.05M, 193/1.44M
```

```
py/custom : 0/0, 0/0
```

```
trim : 19.25%/53.2k, n/a
```

```
overall results
```

```
cycles done : 15
```

```
total paths : 703
```

```
uniq crashes : 0
```

```
uniq hangs : 0
```

```
map coverage
```

```
map density : 5.78% / 13.98%
```

```
count coverage : 3.30 bits/tuple
```

```
findings in depth
```

```
 favored paths : 114 (16.22%)
```

```
 edges on : 167 (23.76%)
```

```
total crashes : 0 (0 unique)
```

```
total tmouts : 0 (0 unique)
```

```
path geometry
```

```
levels : 11
```

```
pending : 121
```

```
pending : 0
```

```
run finds : 699
```

```
imported : n/a
```

```
stability : 99.88%
```

```
[cpu000: 12%]
```

How to Create a Fuzzer Then?

- Fork an existing fuzzer (the n-th AFL-something)

- Create a custom fuzzer from scratch

```
american_fuzzy_lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  new paths : 0 days, 0 hrs, 0 min, 0 sec
last uniq crash : none seen yet
last uniq hang  : none seen yet
cycle progress
  now processing : 261*1 (37.1%)
  paths done     : 0
stage progress
  now trying : splice 14
  stage execs : 31/32 (96.88%)
  total execs : 2.55M
  exec speed  : 01.2k/sec
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
havoc/splice : 506/1.05M, 193/1.44M
  py/custom  : 0/0, 0/0
  trim       : 19.25%/53.2k, n/a
map coverage
  map density : 5.78% / 13.98%
  3.30 bits/tuple
findings in depth
  favored paths : 114 (16.22%)
  new edges on  : 167 (23.76%)
  total crashes : 0 (0 unique)
  total tmouts  : 0 (0 unique)
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs  : 0
[cpu000: 12%]
```

Custom Fuzzer Engineering Issues

- Lack of code reuse, you will have to spend a lot of time in adapting different techniques from different fuzzers

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  last uniq crash : none seen yet
  last uniq hang : none seen yet
  cycle progress :
    now processing : 261*1 (37.1%)
    paths timed out : 0 (0.00%)
stage progress
  now trying : splice 14
  stage execs : 31/32 (96.88%)
  total execs : 2.55M
  exec speed : 61.2k/sec
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
  havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
overall results
  cycles done : 15
  paths : 703
  uniq crashes : 0
  uniq hangs : 0
  map coverage :
    max density : 5.78% / 13.98%
    count coverage : 3.30 bits/tuple
findings in depth
  favored paths : 114 (16.22%)
  new edges on : 167 (23.76%)
  total crashes : 0 (0 unique)
  total tmouts : 0 (0 unique)
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
[cpu000: 12%]
```


Custom Fuzzer Engineering Issues

- Lack of code reuse, you will have to spend a lot of time in adapting different techniques from different fuzzers
- Reinventing the wheel, you will code the same code to do that same thing that all others do again and again

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  last uniq crash : none seen yet
  last uniq hang : none seen yet
  cycle progress : 261*1 (37.1%)
  now processing : splice 14
  paths timed out : 0 (0.00%)
  stage : 11 (100%)
  total execs : 2.55M
  exec speed : 61.2k/sec
  fuzzing strategy yields
    bit flips : n/a, n/a, n/a
    byte flips : n/a, n/a, n/a
    arithmetics : n/a, n/a, n/a
    known ints : n/a, n/a, n/a
    dictionary : n/a, n/a, n/a
    havoc/splice : 506/1.05M, 193/1.44M
    py/custom : 0/0, 0/0
    trim : 19.25%/53.2k, n/a
  overall results
    cycles done : 15
    paths : 703
    uniq crashes : 0
    uniq hangs : 0
  map coverage
    map density : 5.78% / 13.98%
    count coverage : 3.30 bits/tuple
  favored paths : 114 (16.22%)
  paths on : 167 (23.76%)
  total crashes : 0 (0 unique)
  total tmouts : 0 (0 unique)
  path geometry
    levels : 11
    pending : 121
    pend fav : 0
    own finds : 699
    imported : n/a
    stability : 99.88%
[cpu000: 12%]
```

Custom Fuzzer Engineering Issues

- Lack of code reuse, you will have to spend a lot of time in adapting different techniques from different fuzzers
- Reinventing the wheel, you will code the same code to do that same thing that all others do again and again
- Naive design, typically just a mutator

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  last uniq crash : none seen yet
  last uniq hang : none seen yet
  cycle progress : 0 (0.00%)
  now processing : 261*1 (37.1%)
  paths timed out : 0 (0.00%)
  stage : splice 14
  total execs : 2.55M
  havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
  overall results
  cycles done : 15
  paths : 703
  uniq crashes : 0
  uniq hangs : 0
  map coverage
  count coverage : 3.30 bits/tuple
  favored paths : 114 (16.22%)
  paths on : 167 (23.76%)
  total crashes : 0 (0 unique)
  tmouts : 0 (0 unique)
  path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
  [cpu000: 12%]
```

Custom Fuzzer Engineering Issues

- Lack of code reuse, you will have to spend a lot of time in adapting different techniques from different fuzzers
- Reinventing the wheel, you will code the same code to do that same thing that all others do again and again
- Naive design, typically just a mutator
- Scaling, you cannot adapt it easily to multi-core or -machine

LibAFL

README.md

LibAFL, the fuzzer library.

Advanced Fuzzing Library - Slot your own fuzzers together and extend their features using Rust.

LibAFL is written and maintained by Andrea Fioraldi andrea@fioraldi.com and Dominik Maier mail@dmnk.co.

Why LibAFL?

LibAFL gives you many of the benefits of an off-the-shelf fuzzer, while being completely customizable. Some highlight features currently include:



Contributors 26



+ 15 contributors

Languages



Unwatch

26

Unstar

560

Fork

55

Settings

About

Advanced Fuzzing Library - Slot your Fuzzer together in Rust! Scales across cores and machines. For Windows, Android, MacOS, Linux, no_std, ...

fuzzing
afl
afl-fuzz
frida
binary-only
aflplusplus
coverage-guided

Readme

view license

pending :
pend fav : 0
own finds : 699
imported : n/a
stability : 99.88%

[cpu000: 12%]

What?

LibAFL is a library for fuzzers that are

- **Fast** (low IPC, runtime overhead)
- **Scalable** (almost linearly to 200+ cores)
- **Portable** (Android, Windows, MacOS, Linux, Kernels, ...)
- **State-of-the-Art** (Hybrid-, Grammar-, Token-, Feedback-Fuzzing)
- **Multi-instrumentation** (binary-only Frida & Qemu, Clang, Python,...)

And, most importantly, **very extendable** with your own components.

MattGorko/ Tartiflette



Snapshot fuzzing with KVM and LibAFL

2

Contributors

0

Issues

67

Stars

5

Forks



overall results

cycles done : 15

total paths : 703

uniq crashes : 0

uniq hangs : 0

coverage

density : 5.78% / 13.98%

average : 3.30 bits/tuple

in depth

paths : 114 (16.22%)

crashes on : 167 (23.76%)

crashes : 0 (0 unique)

hangs : 0 (0 unique)

path geometry

levels : 11

pending : 121

pending fav : 0

own finds : 699

imported : n/a

stability : 99.88%

[cpu000: 12%]

trim : 19.25%/53.2k, n/a

fuzzing strategy yields

bit flips	: n/a, n/a, n/a
byte flips	: n/a, n/a, n/a
arithmetics	: n/a, n/a, n/a
known ints	: n/a, n/a, n/a
dictionary	: n/a, n/a, n/a
havoc/splice	: 506/1.05M, 193/1.44M
py/custom	: 0/0, 0/0
trim	: 19.25%/53.2k, n/a

MattGorko/ Tartiflette

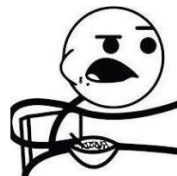


Snapshot of the current state of the project

2
Contributors

epi052/feroxfuzz

A structure-aware HTTP fuzzing library



1

Contributor

1

Issue

0

Stars

0

Forks



overall results
cycles done : 15
total paths : 703
uniq crashes : 0
uniq hangs : 0

coverage
density : 5.78% / 13.98%
average : 3.30 bits/tuple

4 (16.22%)

7 (23.76%)

(0 unique)

(0 unique)

path geometry

levels : 11

pending : 121

end fav : 0

non-finds : 699

imported : n/a

stability : 99.88%

[cpu000: 12%]

MattGorko/ Tartiflette

Snapshot of the current state of the project

2
Contributors

epi052

2
Contributors

57
Issues

62
Stars

6
Forks



tlspuffin/tlspuffin

A symbolic-model-guided fuzzer for TLS



A structure-aware HTTP fuzzing library



1
Contributor

1
Issue

0
Stars

0
Forks



```
american fuzzy lop ++2.65d (1
```

```
process timing
```

```
run time : 0 days, 0 hrs, 0 mi
```

```
last new path : 0 days, 0 hrs, 0 mi
```

```
last uniq crash : none seen yet
```

```
last uniq hang : none seen yet
```

```
cycle progress
```

```
now processing : 261*1 (37.1%)
```

```
paths timed out : 0 (0.00%)
```

Is fuzzer X

better than

Y?

```
stage progress
```

```
now trying : splice 14
```

```
stage execs : 31/32 (96.88%)
```

```
total mem : 5M
```

```
exec speed : 61.2k/sec
```

```
fuzzing strategy fields
```

```
bit flips : n/a, n/a, n/a
```

```
byte flips : n/a, n/a, n/a
```

```
arithmetics : n/a, n/a, n/a
```

```
known ints : n/a, n/a, n/a
```

```
dictionary : n/a, n/a, n/a
```

```
havoc/splice : 506/1.05M, 193/1.44M
```

```
py/custom : 0/0, 0/0
```

```
trim : 19.25%/53.2k, n/a
```

We don't know. Really, we can only speculate about this.

Current benchmarking metrics

- Code coverage over time
- Bugs over time
- Speed
- CVEs found (lol)
- Reached coverage for each fuzz case (not so used, IMO useful to benchmark structured mutators)

```
overall results
cycles done : 15
total paths : 703
uniq crashes : 0
uniq hangs : 0
```

```
map coverage
map density : 5.78% / 13.98%
count coverage : 3.30 bits/tuple
```

```
findings in depth
favored paths : 114 (16.22%)
new edges on : 167 (23.76%)
total crashes : 0 (0 unique)
```

```
path geometry
levels : 11
pending : 121
pend fav : 0
own finds : 699
imported : n/a
stability : 99.88%
```

```
[cpu000: 12%]
```

Standard benchmarks ATM

FuzzBench: Fuzzer Benchmarking As a Service

FuzzBench is a free service that evaluates fuzzers on a wide variety of real-world benchmarks, at Google scale. The goal of FuzzBench is to make it painless to rigorously evaluate fuzzing research and make fuzzing research easier for the community to adopt. We invite members of the research community to contribute their fuzzers and give us feedback on improving our evaluation techniques.

FuzzBench provides:

- An easy API for integrating fuzzers.
- Benchmarks from real-world projects. FuzzBench can use any [OSS-Fuzz](#) project as a benchmark.
- A reporting library that produces reports with graphs and statistical tests to help you understand the significance of results.

To participate, submit your fuzzer to run on the FuzzBench platform by following [our simple guide](#). After your integration is accepted, we will run a large-scale experiment using your fuzzer and generate a report comparing your fuzzer to others, such as AFL and libFuzzer. See [a sample report](#).

Standard benchmarks ATM

Magma: A Ground-Truth Fuzzing Benchmark

Magma is a collection of open-source libraries with widespread usage and a long history of security-critical bugs and vulnerabilities. In light of the need for better fuzzer evaluation, we *front-ported* bugs from previous bug reports to the latest versions of these libraries.

For each ported bug, we added in-line (source-code-level) instrumentation to collect ground-truth information about bugs **reached** (buggy code executed) and **triggered** (fault condition satisfied by input). This instrumentation allows a monitoring utility to measure fuzzer progress in real time.

Magma also includes the `captain` toolset which facilitates the process of building Magma targets and running campaigns.

Check out a [sample Magma report](#) and read the [paper](#). Questions, comments, and feedback are welcome!



generate a report comparing your fuzzer to others, such as AFL and libFuzzer. See [a sample report](#).

Can we improve?

- More representative bugs
- “Automated Magma”
- Changing often the targets (maybe from OSSFuzz) to avoid overfitting
- Decent synthetic bugs?

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  find time : 0 days, 0 hrs, 0 min, 1 sec
last uniq crash : none seen yet
last uniq hang  : none seen yet
cycle progress  :
now executing   : 261*1 (37.1%)
paths timed out : 0 (0.00%)
stage progress  :
now trying     : splice 14
stage progress  : 13/32 (96.88%)
total execs    : 2.55M
exec progress  :
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
havoc/splice : 506/1.05M, 193/1.44M
  py/custom  : 0/0, 0/0
  trim       : 19.25%/53.2k, n/a
overall results
cycles done : 15
total paths : 703
uniq crashes : 0
uniq hangs  : 0
map coverage
  map density : 5.78% / 13.98%
count coverage : 3.30 bits/tuple
finding dupes
  favored paths : 114 (16.22%)
  new edges on  : 167 (23.76%)
total crashes : 0 (0 unique)
total tmouts  : 0 (0 unique)
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
[cpu000: 12%]
```

Can we improve?

- More representative bugs
- “Automated Magma”
- Changing often the targets (maybe from OSSFuzz) to avoid overfitting
- Decent synthetic bugs?

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  find time : 0 days, 0 hrs, 0 min, 1 sec
last uniq crash : none seen yet
last uniq hang  : none seen yet
cycle progress :
  now completed : 261*1 (37.1%)
  paths timed out : 0 (0.00%)
stage progress :
  now trying : splice 14
stage overfitting/32 (96.88%)
total execs : 2.55M
  execs in libpng_harness
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs : 0
map coverage
  map density : 5.78% / 13.98%
count coverage : 3.30 bits/tuple
finding dupes
  favored paths : 114 (16.22%)
  new edges on : 167 (23.76%)
total crashes : 0 (0 unique)
total tmouts : 0 (0 unique)
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
[cpu000: 12%]
```

Can we improve?

31ST USENIX
SECURITY SYMPOSIUM

ATTEND

PROGRAM

PARTICIPATE

SPONSORS

ABOUT

FIXREVERTER: A Realistic Bug Injection Methodology for Benchmarking Fuzz Testing

Authors:

Zenong Zhang and Zach Patterson, *University of Texas at Dallas*; Michael Hicks, *University of Maryland and Amazon*; Shiya Wei, *University of Texas at Dallas*

Distinguished Paper Award Winner

Abstract:

Fuzz testing is an active area of research with proposed improvements published at a rapid pace. Such proposals are assessed empirically: Can they be shown to perform better than the status quo? Such an assessment requires a benchmark of target programs with well-identified, realistic bugs. To ease the construction of such a benchmark, this paper presents FIXREVERTER, a tool that automatically injects realistic bugs in a program. FIXREVERTER takes as input a bugfix pattern which contains both code syntax and semantic conditions. Any code site that matches the specified syntax is undone if the semantic conditions are satisfied, as checked by static analysis, thus (re)introducing a likely bug. This paper focuses on three bugfix patterns, which we call conditional-abort, conditional-execute, and conditional-assign, based on a study of fixes in a corpus of Common Vulnerabilities and Exposures (CVEs). Using FIXREVERTER we have built REVBUGBENCH, which consists of 10 programs into which we have injected nearly 8,000 bugs; the programs are taken from FuzzBench and Binutils, and represent common targets of fuzzing evaluations. We have integrated REVBUGBENCH into the FuzzBench service, and used it to evaluate five fuzzers. Fuzzing performance varies by fuzzer and program, as desired/expected. Overall, 219 unique bugs were reported, 19% of which were detected by just one fuzzer.

```
american fuzzy lop ++2.65d (1
process timing
  run time : 0 days, 0 hrs, 0 mi
  last new path : 0 days, 0 hrs, 0 mi
last uniq crash : none seen yet
last uniq hang : none seen yet
cycle progress
  now processing : 261*1 (37.1%)
paths timed : 0 (00%)
stage progress
  now trying splice 14
stage : 3 (91.8%)
total exec : 2.55M
exec speed : 61.21/sec
fuzzing
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
```

Hard engineering problems

There's a paper about it, problem solved.



Re-implementing things is hard

- Development cost and maintenance

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  total time : 0 days, 0 hrs, 0 min, 43 sec
last uniq crash : none seen yet
last uniq hang  : none seen yet
cycle progress
  now processing : 261*1 (37.1%)
  paths timed out : 0 (0.00%)
stage progress
  now trying : splice 14
  stage execs : 31/32 (96.88%)
  total execs : 2.55M
  exec speed  : 61.2k/sec
fuzzing strategy yields
  bit flips   : n/a, n/a, n/a
  byte flips  : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints  : n/a, n/a, n/a
  dictionary  : n/a, n/a, n/a
  havoc/splice : 506/1.05M, 193/1.44M
  py/custom   : 0/0, 0/0
  trim        : 19.25%/53.2k, n/a
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs  : 0
map coverage
  map density : 5.78% / 13.98%
  count coverage : 3.30 bits/tuple
findings in depth
  favored paths : 114 (16.22%)
  new edges on : 167 (23.76%)
  total crashes : 0 (0 unique)
  total tmouts  : 0 (0 unique)
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported  : n/a
  stability : 99.88%
[cpu000: 12%]
```

Re-implementing things is hard

- Development cost and maintenance
- Re-evaluate techniques to decide if the improvement worths the effort

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  total time : 0 days, 0 hrs, 0 min, 43 sec
last uniq crash : none seen yet
last uniq hang  : none seen yet
cycle progress :
  now processing : 261*1 (37.1%)
  paths timed out : 0 (0.00%)
stage progress
  now trying : splice 14
  stage execs : 31/32 (96.88%)
  total execs : 2.55M
  exec speed  : 61.2k/sec
fuzzing strategy yields
  bit flips   : n/a, n/a, n/a
  byte flips  : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints  : n/a, n/a, n/a
  dictionary  : n/a, n/a, n/a
  havoc/splice : 506/1.05M, 193/1.44M
  py/custom   : 0/0, 0/0
  trim        : 19.25%/53.2k, n/a
map coverage
  map density : 5.78% / 13.00%
  count coverage : 3.30 bits/tuple
findings in depth
  favored paths : 114 (16.22%)
  new edges on : 167 (23.76%)
  total crashes : 0 (0 unique)
  total tmouts  : 0 (0 unique)
path geometry
  levels      : 11
  pending    : 121
  pend fav   : 0
  own finds  : 699
  imported   : n/a
  stability  : 99.88%
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs  : 0
[cpu000: 12%]
```

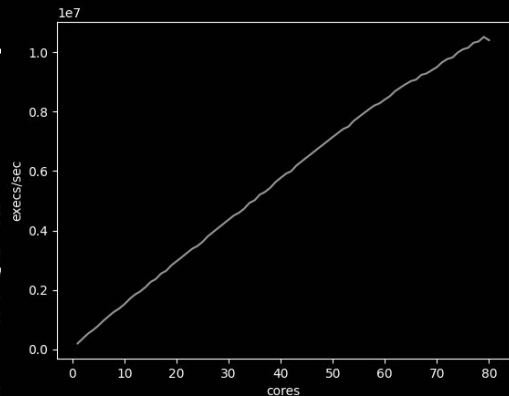

Re-implementing things is hard

- Development cost and maintenance
- Re-evaluate techniques to decide if the improvement worths the effort
- Can we do better simply buying more core?

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  total time : 0 days, 0 hrs, 0 min, 43 sec
last uniq crash : none seen yet
last uniq hang : none seen yet
cycle progress
  now processed : 261*1 (37.1%)
  paths timed out : 0 (0.00%)
stage progress
  now trying : splice 14
  stage : 3 (90%)
total execs : 2.55M
exec speed : 61.2k/sec
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
map coverage
  map density : 5.78% / 13.00%
  count coverage : 3.30 bits/tuple
findings in depth
  favored paths : 114 (16.22%)
  on : 167 (23.76%)
total crashes : 0 (0 unique)
total tmouts : 0 (0 unique)
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
[cpu000: 12%]
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs : 0
```

Re-implementing things is hard

- Development cost and maintenance
- Re-evaluate techniques to decide if the improvement is worth the effort
- Can we do better simply by buying more cores?



```
1 [ | 1.3% ] 21 [ | 100.0% ] 41 [ | 100.0% ] 61 [ | 100.0% ]
2 [ | 100.0% ] 22 [ | 100.0% ] 42 [ | 100.0% ] 62 [ | 100.0% ]
3 [ | 100.0% ] 23 [ | 100.0% ] 43 [ | 100.0% ] 63 [ | 100.0% ]
4 [ | 100.0% ] 24 [ | 100.0% ] 44 [ | 100.0% ] 64 [ | 100.0% ]
5 [ | 100.0% ] 25 [ | 100.0% ] 45 [ | 100.0% ] 65 [ | 100.0% ]
6 [ | 100.0% ] 26 [ | 100.0% ] 46 [ | 100.0% ] 66 [ | 100.0% ]
7 [ | 100.0% ] 27 [ | 100.0% ] 47 [ | 100.0% ] 67 [ | 100.0% ]
8 [ | 100.0% ] 28 [ | 100.0% ] 48 [ | 100.0% ] 68 [ | 100.0% ]
9 [ | 100.0% ] 29 [ | 100.0% ] 49 [ | 100.0% ] 69 [ | 100.0% ]
10 [ | 100.0% ] 30 [ | 100.0% ] 50 [ | 100.0% ] 70 [ | 100.0% ]
11 [ | 100.0% ] 31 [ | 100.0% ] 51 [ | 100.0% ] 71 [ | 100.0% ]
12 [ | 100.0% ] 32 [ | 100.0% ] 52 [ | 100.0% ] 72 [ | 100.0% ]
13 [ | 100.0% ] 33 [ | 100.0% ] 53 [ | 100.0% ] 73 [ | 100.0% ]
14 [ | 100.0% ] 34 [ | 100.0% ] 54 [ | 100.0% ] 74 [ | 100.0% ]
15 [ | 100.0% ] 35 [ | 100.0% ] 55 [ | 100.0% ] 75 [ | 100.0% ]
16 [ | 100.0% ] 36 [ | 100.0% ] 56 [ | 100.0% ] 76 [ | 100.0% ]
17 [ | 100.0% ] 37 [ | 100.0% ] 57 [ | 100.0% ] 77 [ | 100.0% ]
18 [ | 100.0% ] 38 [ | 100.0% ] 58 [ | 100.0% ] 78 [ | 100.0% ]
19 [ | 100.0% ] 39 [ | 100.0% ] 59 [ | 100.0% ] 79 [ | 100.0% ]
20 [ | 100.0% ] 40 [ | 100.0% ] 60 [ | 100.0% ] 80 [ | 100.0% ]
Mem [ | 7.11G/62.9G ] Tasks: 268, 171 thr; 80 running
Swp [ | 302M/2.00G ] Load average: 57.42 26.66 16.00
Uptime: 6 days, 19:02:49
```

py/custom : 0/0, 0/0

trim : 19.25%/53.2k, n/a

[cpu000: 12%]

Re-implementing things is hard

- Development cost and maintenance
- Re-evaluate techniques to decide if the improvement worths the effort
- Can we do better simply buying more core?
- Lack of production-ready engines for tracing/instrumentation of exotic targets

```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  total time : 0 days, 0 hrs, 0 min, 43 sec
last uniq crash : none seen yet
last uniq hang  : none seen yet
cycle progress
  now processed : 261*1 (37.1%)
  paths timed out : 0 (0.00%)
stage progress
  now trying : splice 14
  stage : 3 (9.09%)
total execs : 2.55M
  execs : 2.55M (100%)
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
  havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
map coverage
  map density : 5.78% / 13.00%
  count coverage : 3.30 bits/tuple
findings in depth
  favored paths : 114 (16.22%)
  on : 167 (23.76%)
total crashes : 0 (0 unique)
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
[cpu000: 12%]
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs : 0
```


Hard targets

```
Command Prompt
(base) c:\work\codes\wtf\targets\hevd>server.bat --address=tcp://192.168.2.41:31337/

over@bubuntu:~/wtf/targets/hevd$ ./fuzz-kvm.sh --address tcp://192.168.2.41:31337/
```

```
memset(ar, 0x0, sizeof(address_range_t));
ar->name = "ld_preload_fuzz.so";
calc_address_range(ar);

if(ar->found){
    printf("[init] ld_preload library mapped at: \t0x%016lx-0x%016lx\n", ar->start, ar->end);
    printf("target region \t0x%016lx-0x%016lx (IP0)\n", ar->ip0_a, ar->ip0_b);
    printf("library region \t0x%016lx-0x%016lx (IP1)\n", ar->ip1_a, ar->ip1_b);

    .loc(sizeof(uint64_t)*3);
    .zeof(uint64_t)*3);

    pt_auto_addr_range_a){
        ip0_a;
        ip0_b;
        5550000000;
        5550000000;

        /*
        :FFFFFFFF000;
        :FFFFFFFF001;

        ranges for IPT tracing even if our target has compile-time instrumentations */
        pt_auto_addr_range_b){
            ip1_a;
            ip1_b;
            ranges[1] = ar->ip1_b;
            ranges[2] = 1;
        }
        else{
            /* fix this later */
            ranges[0] = 0xFFFFFFFFFFFFFFFF001;
            ranges[1] = 0xFFFFFFFFFFFFFFFF002;
            sranges[2] = 1;
        }

        /* submit the address ranges for IPT tracing even if our target has compile-time instrumentations */
        if(!get_harness_state()->afl_mode){
            KAFI_hypercall(HYPERCALL_KAFI_RANGE_SUBMIT, (uintptr_t)ranges);
        }
    }
}
```

Overcl0k / wtf Public



Hard targets

- Usability gap
- Emulation-based fuzzing tools are out-of-date
- We need something like “Step till the break point, put the input in \$rdi, snapshot fuzz from here”

```
american_fuzzy_lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  slow times : 0 days, 0 hrs, 0 min, 1 sec
last uniq crash : none seen yet
last uniq hang : none seen yet
cycle progress : 261*1 (37.1%)
now processing : splice 14
paths timed out : 0 (0.00%)
stage : 261*1 (37.1%)
now trying : splice 14
stage : 261*1 (37.1%)
total execs : 2.55M
exec speed : 61.2k/sec
fuzzing strategy yields
  bit flips : n/a, n/a, n/a
  byte flips : n/a, n/a, n/a
  arithmetics : n/a, n/a, n/a
  known ints : n/a, n/a, n/a
  dictionary : n/a, n/a, n/a
havoc/splice : 506/1.05M, 193/1.44M
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
map coverage
  map density : 5.78% / 13.98%
  count coverage : 3.30 bits/tuple
  favored paths : 114 (16.22%)
  new edges on : 167 (23.76%)
  total crashes : 0 (0 unique)
  total tmouts : 0 (0 unique)
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs : 0
[cpu000: 12%]
```

Ask more about fuzzing at

<https://discord.gg/gCraWct>



```
american fuzzy lop ++2.65d (libpng_harness) [explore] {0}
process timing
  run time : 0 days, 0 hrs, 0 min, 43 sec
  no. of paths : 703
last uniq crash : none seen yet
last uniq hang : none seen yet
cycle progress
  now processing : 261*1 (37.1%)
  map density : 5.78% / 13.98%
  avg : 3.30 bits/tuple
  depth
    as : 114 (16.22%)
    on : 167 (23.76%)
    es : 0 (0 unique)
    ss : 0 (0 unique)
stage progress
  now trying : splice 14
  stage execs : 31/32 (96.87%)
  total execs : 2.55M
  exec speed : 61.2k/sec
  fuzzing strategy yields
    bit flips : n/a, n/a, n/a
    byte flips : n/a, n/a, n/a
    arithmetics : n/a, n/a, n/a
    known ints : n/a, n/a, n/a
    dictionary : n/a, n/a, n/a
  havoc/splice : 506/1.05M, n/a, n/a
  py/custom : 0/0, 0/0
  trim : 19.25%/53.2k, n/a
overall results
  cycles done : 15
  total paths : 703
  uniq crashes : 0
  uniq hangs : 0
path geometry
  levels : 11
  pending : 121
  pend fav : 0
  own finds : 699
  imported : n/a
  stability : 99.88%
[cpu000: 12%]
```

Thanks y'all