GitHub Secure Open Source Fund















powered by GitHub Sponsors













stripe

Module: Fuzzing



About Me



Antonio Morales GitHub Senior Security Researcher



2nd edition

- This is a completely new version of the fuzzing workshop
- I've removed all C/C++ content and replaced it with JavaScript and Python examples
- After the workshop, I'll share the material from the first version with you



Workshop Format

- This is a hands-on workshop that follows a learning-by-doing approach
- You will learn by tackling challenges, and I'll be here to guide you along the way
- These challenges are designed to be solved using fuzzing, but feel free to use any method you prefer (good luck xD)



Sistem requirements

- Access to a Linux environment is recommended for the best experience
- All demos and exercises will be conducted on Ubuntu 24.04 LTS
- If you don't have access to a Linux machine, you can use GitHub Codespaces



Agenda

- 1. Fuzzing introduction
- 2. Coverage-guided fuzzing
- 3. Exercise 1: Introduction to Java fuzzing
- 4. Exercise 2: Introduction to Javascript fuzzing
- 5. Exercise 3: Introduction to Python fuzzing
- 6. More fuzzing
- 7. The fuzzing workflow
- 8. Challenge
- 9. Next steps for learning
- 10. Homework
- 11. Your questions



1. Fuzzing introduction

- Fuzzing is an automated testing technique that helps find software bugs and security vulnerabilities
- It works by injecting malformed or unexpected inputs into a program to observe its behavior.
- In recent years, its use has consistently increased due to its success in uncovering vulnerabilities



The simplest fuzzer

cat /dev/urandom | ./your_application

- cat reads random data from /dev/urandom and pipes it into your_application
- In 2009, Charlie Miller used this fuzzing technique to discover critical vulnerabilities in iPhones



Smart Fuzzing

Dumb Fuzzing







Smart Fuzzing

- Fuzzing has significantly evolved over the past 15 years
- New fuzzing techniques have made it much more effective
- The release of the AFL fuzzer marked the beginning of the "smart fuzzing" era



GitHub Secure Open Source Fund

AFL is an

Evolutionary
Mutation-based
Coverage-guided
fuzzer



2. Mutation-based fuzzing

- A mutation-based fuzzer starts with valid input samples (the corpus)
- It generates new test cases by applying small, random changes (bit/byte flips, integer tweaks, ...)
- The goal is to trigger unexpected behavior or crashes



2. Evolutionary fuzzing

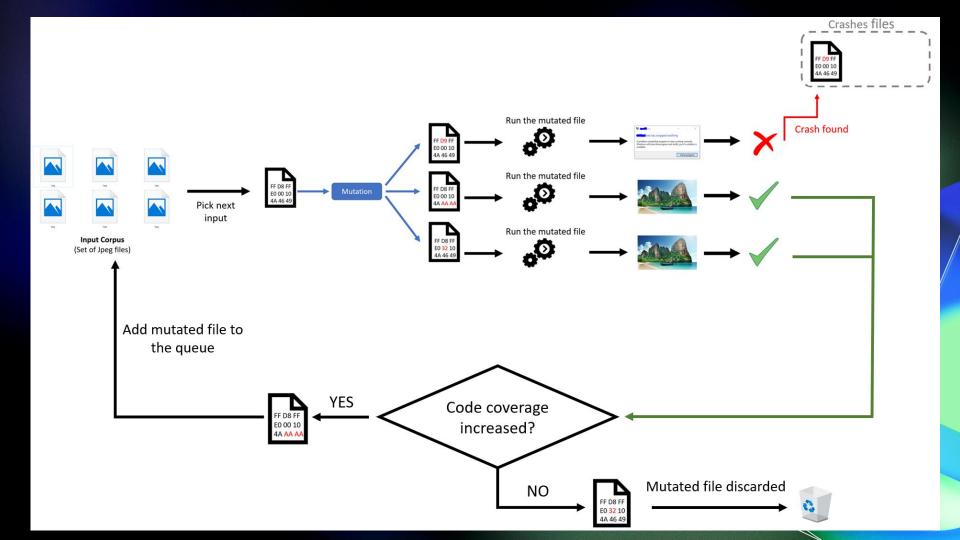
- Based on evolutionary algorithms: inputs evolve over generations according to the principles of natural selection
- It uses some kind of "feedback" from the target program to guide future mutations
- This helps generate inputs that reach deeper and more complex program behavior



2. Coverage-guided fuzzing

- It uses code coverage as feedback to guide the fuzzing process
- Mutations that trigger new execution paths are added to the input queue
- Mutations that don't discover new paths are discarded
- These fuzzers need to "instrument" the code to track coverage





LibFuzzer

- Developed by Google, it's considered the main "rival" to AFL
- It's part of the LLVM project and available on most platforms
- All the fuzzers we'll cover today are based on LibFuzzer



GitHub Secure Open Source Fund

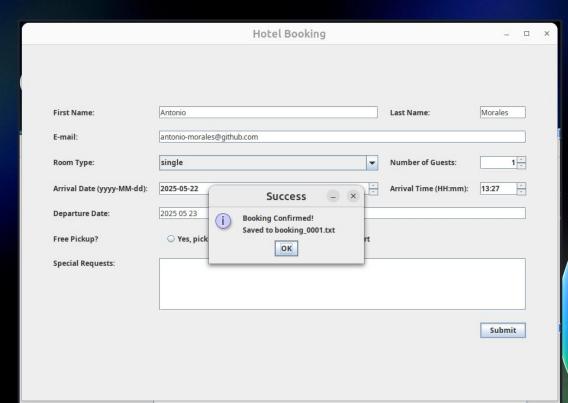
JAZZER



Jazzer

- Jazzer is a coverage-guided fuzzer for Java and other JVM languages
- It's built on libFuzzer
- Developed by Code Intelligence and integrated with Google's OSS-Fuzz
- Available at https://github.com/CodeIntelligenceTesting/jazzer







\$ java --version openjdk 21.0.3 2024-04-16 OpenJDK Runtime Environment (build 21.0.3+9-Ubuntu-1ubuntu1) OpenJDK 64-Bit Server VM (build 21.0.3+9-Ubuntu-1ubuntu1, mixed mode, sharing)

javac BookingFormGUI.java BookingForm.java

java BookingFormGUI

```
import java.io.StringReader;
public class fuzzer {
    // Jazzer expects a public static method named fuzzerTestOneInput
    public static void fuzzerTestOneInput(byte[] data) {
       String s = new String(data);
       try {
         BookingForm.parseBooking(s);
       } catch (ArrayIndexOutOfBoundsException e) {
         throw e;
        } catch (Exception e) {
            // Ignore
```





chmod +x ./jazzer/jazzer

./jazzer/jazzer --cp=out --target_class=fuzzer

./jazzer/jazzer --cp=out --target_class=fuzzer corpus





```
REDUCE cov: 67 ft: 67 corp: 21/665b lim: 4096 exec/s: 343851 rss: 1024Mb L: 45/50 MS: 4 ChangeBit-Custom-EraseBytes-Custom-
#3782366
              REDUCE cov: 67 ft: 67 corp: 21/664b lim: 4096 exec/s: 334294 rss: 1024Mb L: 45/50 MS: 2 EraseBytes-Custom-
#4011532
              pulse cov: 67 ft: 67 corp: 21/664b lim: 4096 exec/s: 349525 rss: 1024Mb
#4194304
#8388608
              pulse cov: 67 ft: 67 corp: 21/664b lim: 4096 exec/s: 335544 rss: 1024Mb
== Java Exception: java.lang.ArrayIndexOutOfBoundsException: Index 11 out of bounds for length 11
       at BookingForm.parseDate(BookingForm.java:195)
       at BookingForm.parseBooking(BookingForm.java:117)
       at fuzzer.fuzzerTestOneInput(fuzzer.java:12)
DEDUP TOKEN: 2812ded32f26d6bb
== libFuzzer crashing input ==
MS: 4 CopyPart-Custom-PersAutoDict-Custom- DE: "ie"-; base unit: 4f5626fcb1b6c23d2b5fa2367f599fe08094c2ba
,0x33,0x3b,0x66,0x61,0x6c,0x73,0x65,0x3b,
n;M;l@m;suite;1;2025-05-21T21:48;ie 05 2323;false;
artifact prefix='./'; Test unit written to ./crash-e7651f9f08c6dc8be1f2eb8d97f310037d5cd54f
Base64: bjtNO2xAbTtzdWlOZTsxOzIwMjUtMDUtMjFUMjE6NDq7aWUqMDUqMjMvMztmYWxzZTs=
reproducer path='.'; Java reproducer written to ./Crash e7651f9f08c6dc8be1f2eb8d97f310037d5cd54f.java
```

GitHub Secure Open Source Fund

JAZZER.JS



Jazzer.js

- Jazzer.js is a coverage-guided fuzzer for JavaScript and Node.js applications
- It's built on libFuzzer and uss V8's built-in code coverage feedback
- Developed by Code Intelligence
- Good for fuzzing JS libraries and APIs
- Available at https://github.com/CodeIntelligenceTesting/jazzer.js

Moment.js

- Moment.js is a library to parse, validate, manipulate, and display dates and times in JavaScript
- Source code is available at: github.com/moment/moment
- We'll use Jazzer.js to find a Regular Expression Denial of Service (ReDoS) vulnerability in version 2.15.1

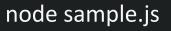


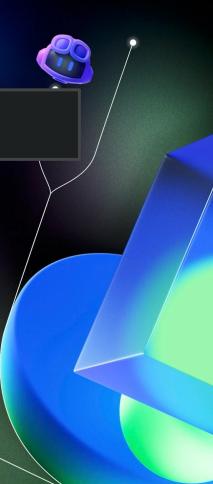


npm init -y

npm install --save-dev @<u>jazzer.js/core</u>

npm install moment@2.15.1





```
const { FuzzedDataProvider } = require('@jazzer.js/core');
const moment = require('moment');
// Set locale to one with vulnerable standalone month parsing
moment.locale('be');
/**
* The fuzz target: Jazzer.js will call this function with arbitrary input.
function fuzz(data) {
  const provider = new FuzzedDataProvider(data);
  const fmtString = provider.consumeRemainingAsString();
 try {
    moment().format(fmtString);
  } catch (e) {
    // Ignore errors; we're interested in performance slowdown
module.exports = { fuzz };
```



\$ npm -v 9.2.0 \$ node -v v18.19.1

npx jazzer fuzzer

GitHub Secure Open Source Fund





```
Dictionary: 4 entries
INFO: Running with entropic power schedule (0xFF, 100).
INFO: Seed: 1735645154
INFO: Loaded 1 modules (512 inline 8-bit counters): 512 [0x7ffff3cff010, 0x7ffff3cff210),
INFO: Loaded 1 PC tables (512 PCs): 512 [0x7fffe9000010,0x7fffe9002010),
INFO: -max len is not provided; libFuzzer will not generate inputs larger than 4096 bytes
INFO: A corpus is not provided, starting from an empty corpus
#2
                           INITED cov: 3 ft: 3 corp: 1/1b exec/s: 0 rss: 142Mb
#65536 pulse cov: 3 ft: 3 corp: 1/1b lim: 652 exec/s: 32768 rss: 158Mb
#131072 pulse cov: 3 ft: 3 corp: 1/1b lim: 1300 exec/s: 32768 rss: 177Mb
#262144 pulse cov: 3 ft: 3 corp: 1/1b lim: 2611 exec/s: 29127 rss: 216Mb
ALARM: working on the last Unit for 7 seconds
                        and the timeout value is 5 (use -timeout=N to change)
MS: 5 InsertRepeatedBytes-ManualDict-InsertRepeatedBytes-CrossOver-ChangeByte- DE: proto -; base unit: adc83b19e793491b1c6ea0fd8b46cd9f32e592fc
8x0, b8x0, b
x0, b8x0, b8
, D8X0, D8X0
0x74,0x6f,0x5f,0x5f,0xcd,0xcd,0xa
 artifact_prefix='./'; Test unit written to ./timeout-e7844097aec01beee0651295dfab04373bee3636
                                 ==273907== ERROR: libFuzzer: timeout after 7 seconds
SUMMARY: LIDFUZZEF: LIMEOUT
```

INFO: Running with entropic power schedule (0xFF, 100).

Dictionary: 4 entries

INFO: Seed: 1735645154



npx jazzer fuzzer -- -seed=1735645154

```
INFO: Loaded 1 modules (512 inline 8-bit counters): 512 [0x7ffff3cff010, 0x7ffff3cff210),
INFO: Loaded 1 PC tables (512 PCs): 512 [0x7fffe9000010,0x7fffe9002010),
 INFO: -max len is not provided; libFuzzer will not generate inputs larger than 4096 bytes
INFO: A corpus is not provided, starting from an empty corpus
                           INITED cov: 3 ft: 3 corp: 1/1b exec/s: 0 rss: 142Mb
#65536 pulse cov: 3 ft: 3 corp: 1/1b lim: 652 exec/s: 32768 rss: 158Mb
#131072 pulse cov: 3 ft: 3 corp: 1/1b lim: 1300 exec/s: 32768 rss: 177Mb
#262144 pulse cov: 3 ft: 3 corp: 1/1b lim: 2611 exec/s: 29127 rss: 216Mb
ALARM: working on the last Unit for 7 seconds
                        and the timeout value is 5 (use -timeout=N to change)
MS: 5 InsertRepeatedBytes-ManualDict-InsertRepeatedBytes-CrossOver-ChangeByte- DE: " proto "-; base unit: adc83b19e793491b1c6ea0fd8b46cd9f32e592fc
8x0, b8x0, b
x0, b8x0, b8
. b8x0, b8x0
0x74,0x6f,0x5f,0x5f,0xcd,0xcd,0xa,
 artifact prefix='./'; Test unit written to ./timeout-e7844097aec01beee0651295dfab04373bee3636
==273907== ERROR: libFuzzer: timeout after 7 seconds
SUMMARY: libFuzzer: timeout
```

```
// LOCALES
67
     67
68
     68
          - var MONTHS_IN_FORMAT = /D[oD]?(\[[^\[\]]*\]|\s+)+MMMM?/;
69
          + var MONTHS_IN_FORMAT = /D[oD]?(\[[^\[\]]*\]|\s)+MMMM?/;
     69
            export var defaultLocaleMonths = 'January_February_March_April_May_June_July_August_Sep
70
     70
71
     71
            export function localeMonths (m, format) {
72
     72
                if (!m) {
    @@ -50,7 +50,7 @@ export default moment.defineLocale('lt', {
50
     50
                months : {
                    format: 'sausio vasario kovo balandžio gegužės birželio liepos rugpjūčio rugsėj
51
     51
                    standalone: 'sausis_vasaris_kovas_balandis_gegužė_birželis_liepa_rugpjūtis_rugs
52
     52
53
                    isFormat: /D[oD]?(\[^{\[^{\]}}*\]) \times_{}+MMMM?(MMMM?(\[^{\[^{\[]}}*\]) \times_{}+D[oD]?/
                    isFormat: /D[oD]?(\[[^{[]]*^]]*) + MMMM?(MMM?(\[[^{[]]*^]]*) + D[oD]?/
     53
```

GitHub Secure Open Source Fund





Atheris

- Atheris is a coverage-guided fuzzer for Python and C extension modules
- Developed by Google, it integrates with libFuzzer for feedback-driven fuzzing
- It supports fuzzing both:
 - Pure Python code
 - Native Python extensions (ASAN, UBSAN, etc)
- Available at: https://github.com/google/atheris

Mistune

- "A fast yet powerful Python Markdown parser with renderers and plugins"
- Convert Markdown to HTML with ease:

```
import mistune
mistune.html(your_markdown_text)
```

Available at https://github.com/lepture/mistune



GitHub Secure Open Source Fund

The goal is to find a XSS vulnerability in Mistune using Atheris fuzzer



In order to run Atheris, we need Python 3.11:

curl https://pyenv.run | bash

export PATH="\$HOME/.pyenv/bin:\$PATH"
eval "\$(pyenv init --path)"
eval "\$(pyenv init -)"
eval "\$(pyenv virtualenv-init -)"

exec "\$SHELL"

pyenv install 3.11

pyenv global 3.11

python3.11 -m venv venv source venv/bin/activate

pip install atheris

pip install mistune==0.7.4

pip install beautifulsoup4 html5lib

python3 fuzz_mistune.py

Base64: CdvdKCOpCOAj3dvdKCOJ3SMkItskPik6



```
#155204 REDUCE cov: 209 ft: 1047 corp: 246/13150b lim: 357 exec/s: 2675 rss: 54Mb L: 129/334 MS: 3 ChangeBinInt-ChangeBinInt-EraseBytes-
#155505 REDUCE cov: 209 ft: 1047 corp: 246/13131b lim: 357 exec/s: 2681 rss: 54Mb L: 82/334 MS: 1 EraseBytes-
#155576 REDUCE cov: 209 ft: 1047 corp: 246/13112b lim: 357 exec/s: 2682 rss: 54Mb L: 125/334 MS: 1 EraseBytes-
#155933 REDUCE cov: 209 ft: 1047 corp: 246/13075b lim: 357 exec/s: 2642 rss: 54Mb L: 153/334 MS: 2 InsertByte-EraseBytes-
#156809 REDUCE cov: 209 ft: 1047 corp: 246/13071b lim: 365 exec/s: 2657 rss: 54Mb L: 25/334 MS: 1 EraseBytes-
#156863 REDUCE cov: 209 ft: 1057 corp: 247/13390b lim: 365 exec/s: 2658 rss: 54Mb L: 319/334 MS: 4 ChangeByte-CMP-InsertByte-CopyPart- DE
AssertionError: Unescaped quote in <a href="$ ]#$"[$>"> for payload '[]($)\t\x00#][]($\t]#$"[$>):'
Traceback (most recent call last):
 File "/home/antonio/FuzzingWorkshop/exercise3/mistune/fuzz_mistune.py", line 25, in TestOneInput
    raise AssertionError(
AssertionError: Unescaped quote in <a href="$ ]#$"[$>"> for payload '[]($\\t\x00#][]($\\t]#$"[$>):'
==323328== ERROR: libFuzzer: fuzz target exited
SUMMARY: libFuzzer: fuzz target exited
MS: 3 ShuffleBytes-ChangeBinInt-EraseBytes-; base unit: e54ea44c25638e563a5183b02ac98f1c6dd2acd7
0x9,0xdb,0xdd,0x28,0x24,0x29,0x9,0x0,0x23,0xdd,0xdb,0xdb,0xdb,0x28,0x24,0x9,0xdd,0x23,0x24,0x22,0xdb,0x24,0x3e,0x29,0x3a,
\011\333\335($)\011\000#\335\333\335($\011\335#$\"\333$>):
artifact_prefix='./'; Test unit written to ./crash-eefa1b68470b3f5af236af3465b992549f572d2c
```

python3 fuzz_mistune.py -seed=3923903506

```
#155204 REDUCE cov: 209 ft: 1047 corp: 246/13150b lim: 357 exec/s: 2675 rss: 54Mb L: 129/334 MS: 3 ChangeBinInt-ChangeBinInt-EraseBytes-
#155505 REDUCE cov: 209 ft: 1047 corp: 246/13131b lim: 357 exec/s: 2681 rss: 54Mb L: 82/334 MS: 1 EraseBytes-
#155576 REDUCE cov: 209 ft: 1047 corp: 246/13112b lim: 357 exec/s: 2682 rss: 54Mb L: 125/334 MS: 1 EraseBytes-
#155933 REDUCE cov: 209 ft: 1047 corp: 246/13075b lim: 357 exec/s: 2642 rss: 54Mb L: 153/334 MS: 2 InsertByte-EraseBytes-
```

```
#156809 REDUCE cov: 209 ft: 1047 corp: 246/13071b lim: 365 exec/s: 2657 rss: 54Mb L: 25/334 MS: 1 EraseBytes-
#156863 REDUCE cov: 209 ft: 1057 corp: 247/13390b lim: 365 exec/s: 2658 rss: 54Mb L: 319/334 MS: 4 ChangeByte-CMP-InsertByte-CopyPart- DE
AssertionError: Unescaped quote in <a href="$ ]#$"[$>"> for payload '[]($)\t\x00#][]($\t]#$"[$>):'
Traceback (most recent call last):
 File "/home/antonio/FuzzingWorkshop/exercise3/mistune/fuzz_mistune.py", line 25, in TestOneInput
    raise AssertionError(
AssertionError: Unescaped quote in <a href="$ ]#$"[$>"> for payload '[]($\\t\x00#][]($\\t]#$"[$>):'
==323328== ERROR: libFuzzer: fuzz target exited
SUMMARY: libFuzzer: fuzz target exited
MS: 3 ShuffleBytes-ChangeBinInt-EraseBytes-; base unit: e54ea44c25638e563a5183b02ac98f1c6dd2acd7
0x9,0xdb,0xdd,0x28,0x24,0x29,0x9,0x0,0x23,0xdd,0xdb,0xdb,0xdb,0x28,0x24,0x9,0xdd,0x23,0x24,0x22,0xdb,0x24,0x3e,0x29,0x3a,
\011\333\335($)\011\000#\335\333\335($\011\335#$\"\333$>):
artifact_prefix='./'; Test unit written to ./crash-eefa1b68470b3f5af236af3465b992549f572d2c
Base64: CdvdKCOpCOAj3dvdKCOJ3SMkItskPik6
```



Fix bypassing XSS vulnerability.

Bypasing with newline.

```
01d:
>>> markdown('''[aa](java
... script:alert`1`;)''')
'<a href="java\nscript:alert`1`;">aa</a>\n'
New:
>>> markdown('''[aa](java
... script:alert`1`;)''')
'<a href="">aa</a>\n'
Bypassing with malicious mail address.
Old:
>>> markdown('<junorouse@gmail.com"\nonclick="alert(1);>')
<a href="mailto:junorouse@gmail.com"\nonclick="alert(1);">junorouse@gmail.com"\nonclick="alert(1);</a>\n'
New:
>>> markdown('<junorouse@gmail.com"\nonclick="alert(1);>')
'<a href="mailto:junorouse@gmail.com&quot;\nonclick=&quot;alert(1);">junorouse@gmail.com&quot;\nonclick=&quot;alert(1);</a>\n'
```

6. More fuzzing

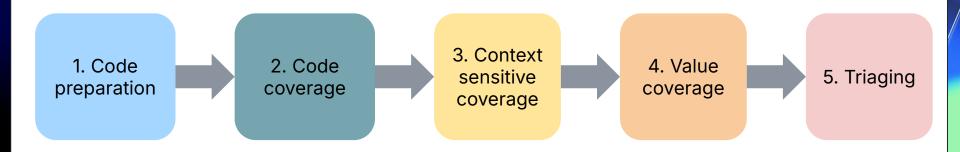
- C/C++: AFL++ fuzzer https://github.com/AFLplusplus/AFLplusplus
- Rust: LibFuzzer + cargo-fuzz
 https://github.com/rust-fuzz/cargo-fuzz
- Go:
 - Go <= 1.17: go-fuzz https://github.com/dvyukov/go-fuzz
 - Go >= 1.18: native fuzzing support

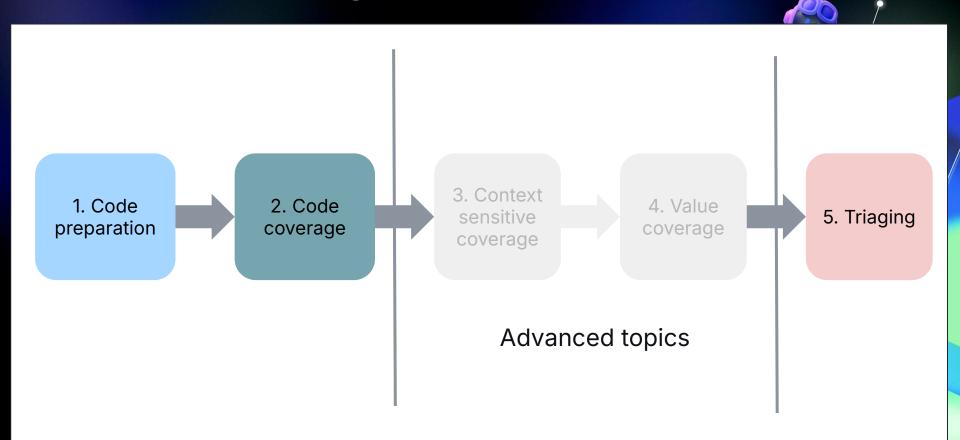


- Fuzzing is an iterative process, similar to any "agile development methodology"
- Thinking of fuzzing as a single-iteration process is one of the biggest mistakes most beginners make
- Next, I will show what I believe is the most efficient "fuzzing process" for obtaining good results in fuzzing

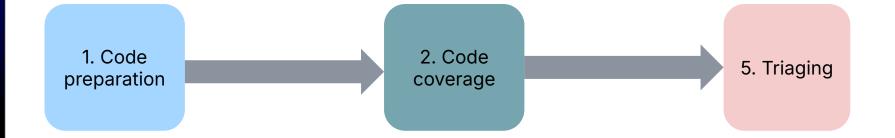




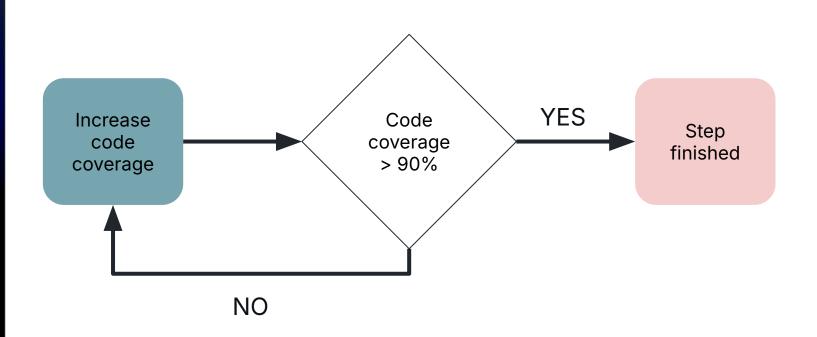














How to increase the code coverage?

- Add new example files
- Use dictionaries
- Use custom mutators [Advanced]
- Use more advanced instrumentation [Advanced]:
 - Compiler transformations (ex. laf-intel)
 - Input-to-State fuzzing (ex. Redqueen)
 - Context-aware fuzzing

5. Triaging:

- Triaging is the process of categorizing bugs
- It involves analyzing crashes or bugs found during testing to determine their root cause, exploitability, and potential risk
- The most straightforward approach is to use a debugger to examine the crash details
- Once we confirm a vulnerability, the last step is reporting it to developer/s



GitHub Secure Open Source Fund

CHALLENGE



Next Steps for Learning



Fuzzing101

https://github.com/antonio-morales/Fuzzing101

A GitHub Security Lab initiative for learning fuzzing and finding vulnerabilities in real software

Todo el material que he mostrado en este workshop pasará a formar parte del repositorio

Exercise No.	Target	CVEs to find	Time estimated	Main topics
Exercise 1	Xpdf	CVE-2019- 13288	120 mins	Afl-clang-fast, Afl-fuzz, GDB
Exercise 2	libexif	CVE-2009-3895, CVE-2012-2836	6 hours	Afl-clang-lto, Fuzz libraries, Eclipse IDE
Exercise 3	TCPdump	CVE-2017- 13028	4 hours	ASan, Sanitizers
Exercise 4	LibTIFF	CVE-2016-9297	3 hours	Code coverage, LCOV
Exercise 5	Libxml2	CVE-2017-9048	3 hours	Dictionaries, Basic parallelization, Fuzzing command-line arguments
Exercise 6	GIMP	CVE-2016-4994, Bonus bugs	7 hours	Persistent fuzzing, Fuzzing interactive applications
Exercise 7	VLC media player	CVE-2019- 14776	6 hours	Partial instrumentation, Fuzzing harness
Exercise 8	Adobe Reader		8 hours	Fuzzing closed-source applications, QEMU instrumentation
Exercise 9	7-Zip	CVE-2016-2334	8 hours	WinAFL, Fuzzing Windows Applications
Exercise 10 (Final Challenge)	Google Chrome / V8	CVE-2019-5847	8 hours	Fuzzilli, Fuzzing Javascript engines

GitHub Secure Open Source Fund

How many of you have a 'fuzzer' folder in your GitHub repository?



Questions?

