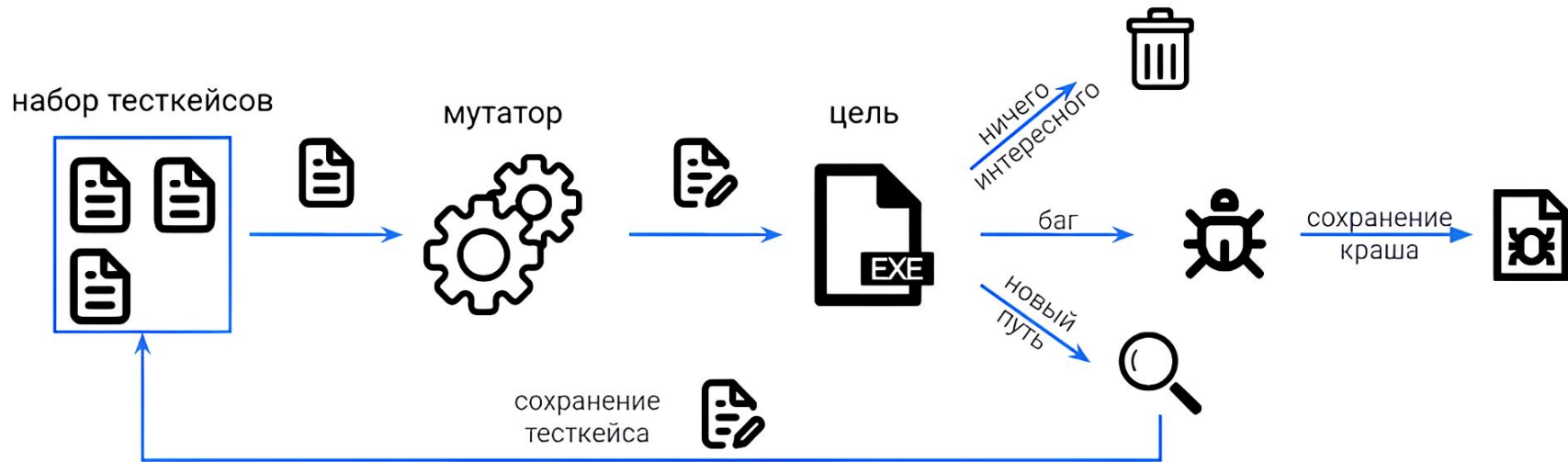


PHP fuzzing workshop

Структурная схема серого фаззера



Важность фаззинга managed языков

Фаззинг способен находить не только memory corruption C/C++. Применяя подход фаззинга для управляемых языков можно обнаружить следующие проблемы:

- Зависания и таймауты
- Состояние гонки
- Неопределенное поведение
- Логические ошибки
- Утечки памяти
- ...

Цель для воркшопа






<https://github.com/smalot/pdfparser>

 pdfparser Public

Watch 81 Fork 553 Star 2.5k

master 2 Branches 71 Tags

Add file <> Code

 k00ni	pull_request_template.md: Set path to CONTRIBUTING.md (#760) ... ✓	Oddcc54 · last month	 452 Commits
 .github	pull_request_template.md: Set path to CONTRIBUTING.md (...)	last month	
 dev-tools	Fixes failing tests; reorganized test files (PHPUnit 10.x related...)	2 years ago	
 doc	Update CustomConfig.md (#729)	8 months ago	

About

PdfParser, a standalone PHP library, provides various tools to extract data from a PDF file.

 Readme

 LGPL-3.0 license

 Activity

▲ - -

Подготовка окружения

Собираем docker образ:

```
docker build --network=host --tag=pdfparser_php_workshop_img .
```

Запускаем контейнер:

```
docker run --rm -v "$(pwd)/assets:/root/assets" -it --network=host  
pdfparser_php_workshop_img /bin/bash
```

Применяем патч:

```
cd /root/pdfparser  
git apply ../assets/add_fuzz.patch
```

Подготовка корпуса

В качестве корпуса будем использовать тестовые pdf файлы из исследуемой библиотеки

```
[root@fedora ~]# ls pdfparser/samples/
Document-Word-Landscape-printedaspdf.pdf  Document3_pdfcreator_nocompressed.pdf  XMP_Metadata.pdf
Document1_foxitreader.pdf                 DocumentWithLotsOfObjects.pdf          bugs
Document1_pdfcreator.pdf                  ImproperFontFallback.pdf               corrupted.pdf
Document1_pdfcreator_nocompressed.pdf      InternationalChars.pdf                 grouped-by-generator
Document1_pdfxchange.pdf                  SimpleInvoiceFilledExample1.pdf        not_really_encrypted.pdf
Document2_pdfcreator_nocompressed.pdf      SimpleInvoiceFilledExample2.pdf
```

[root@fedora ~]# █

Написание обертки

Для php-fuzzer необходимо реализовать функцию `setTarget`. В нашем случае обертка вызывает метод `parseContent` класса `Parser`

```
1  <?php declare(strict_types=1);
2
3  ini_set('memory_limit', '-1');
4  require __DIR__.'/../vendor/autoload.php';
5
6  $parser = new \Smalot\PdfParser\Parser();
7
8  $config->setTarget(function (string $input) use ($parser) {
9      try {
10         $pdf = $parser->parseContent($input);
11     }
12     catch (\Exception $e) {
13     }
14 });
15
16 $config->setMaxLen(2048);
17
```

Фаззинг php кода с помощью php-fuzzer

Создаем директорию для найденных падений:

```
cd /root
```

```
mkdir crashes && cd crashes
```

Запускаем фаззинг PDF парсера:

```
cd /root/crashes
```

```
php-fuzzer fuzz /root/pdfparser/tests/fuzz/fuzz_target.php /root/pdfparser/samples/
```


(Опционально) использование словаря

```
cd /root/crashes
```

```
php-fuzzer fuzz /root/pdfparser/tests/fuzz/fuzz_target.php /root/pdfparser/samples/ --dict  
/root/assets/pdf.dict
```

В случае нахождения падения есть возможность минимизировать его и посмотреть на его вывод

```
cd /root/crashes
```

```
php-fuzzer minimize-crash /root/pdfparser/tests/fuzz/fuzz_target.php  
/root/crashes/crash-HASH.txt
```

```
php-fuzzer run-single /root/pdfparser/tests/fuzz/fuzz_target.php minimized-HASH.txt
```

Сбор покрытия

Создаем директорию для покрытия

```
cd /root
```

```
mkdir coverage
```

Генерируем html отчет

```
cd /root
```

```
php-fuzzer report-coverage pdfparser/tests/fuzz/fuzz_target.php pdfparser/samples/  
coverage
```

Визуализация покрытия

Копируем результат покрытия на хост

```
cp -r /root/coverage /root/assets
```

Выходим из контейнера и открываем в браузере index.html

