

# Automatic Binary Exploitation and Patching using Mechanical [Shell]Phish

Antonio Bianchi

[antonio@cs.ucsb.edu](mailto:antonio@cs.ucsb.edu)



University of California, Santa Barbara (UCSB)



HITCON Pacific  
December 2nd, 2016

# Shellphish

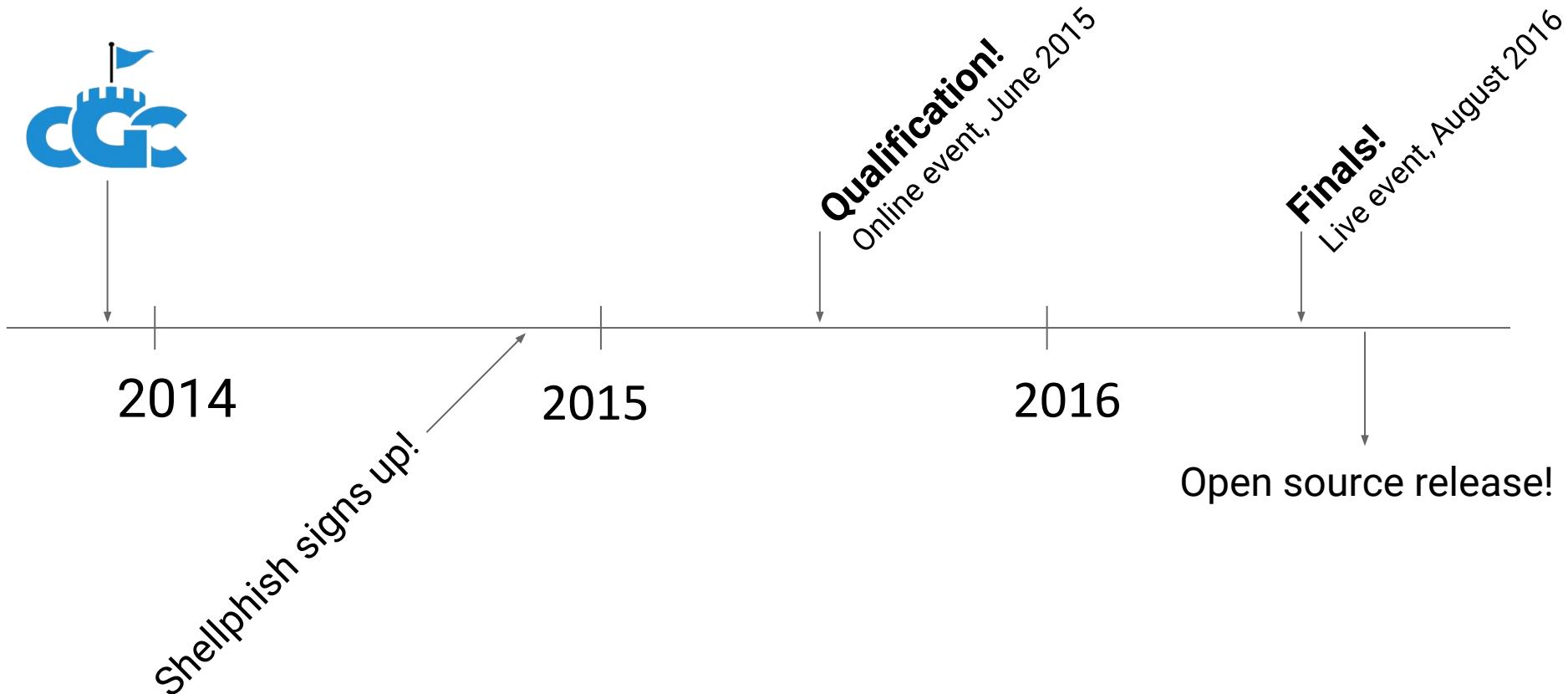


- A team of security enthusiasts
  - Do research in system security
  - Play security competitions (CTF)
  - Mainly students from  
University of California, Santa Barbara
  - More info:
    - “A Dozen Years of Shellphish”  
<https://youtu.be/APY2SsBde1U>

- A fully automated CTF competition
- Organized by DARPA
- No Human intervention



# CGC - Timeline





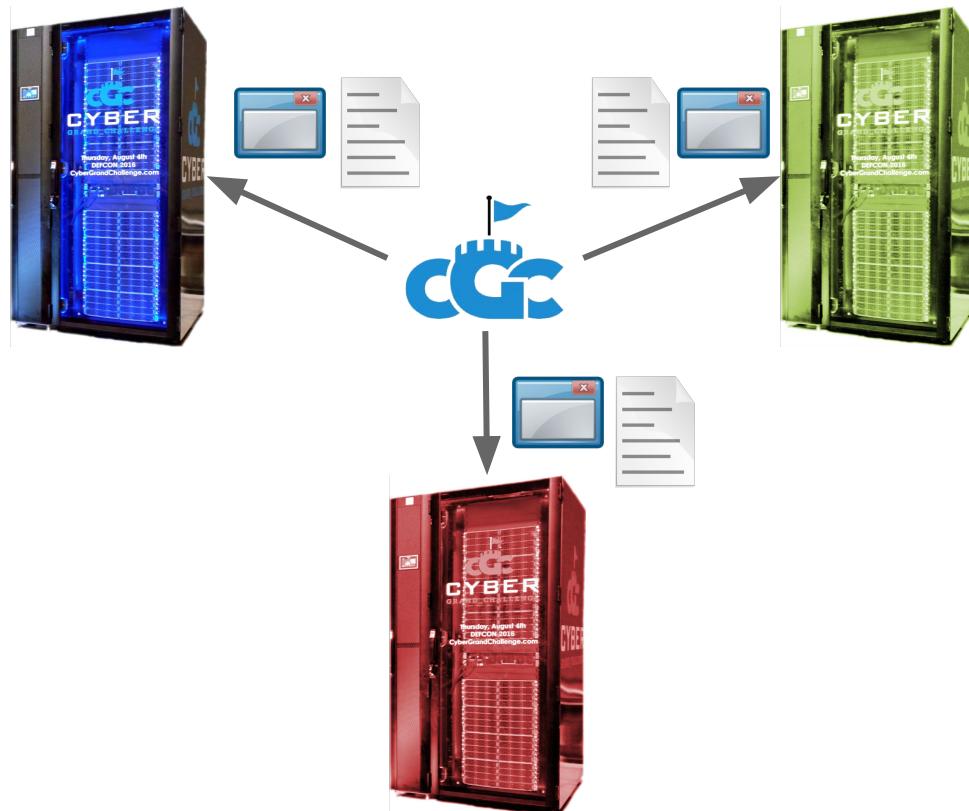
CYBER  
GRAND CHALLENGE

DARPA



# CGC - How the game worked

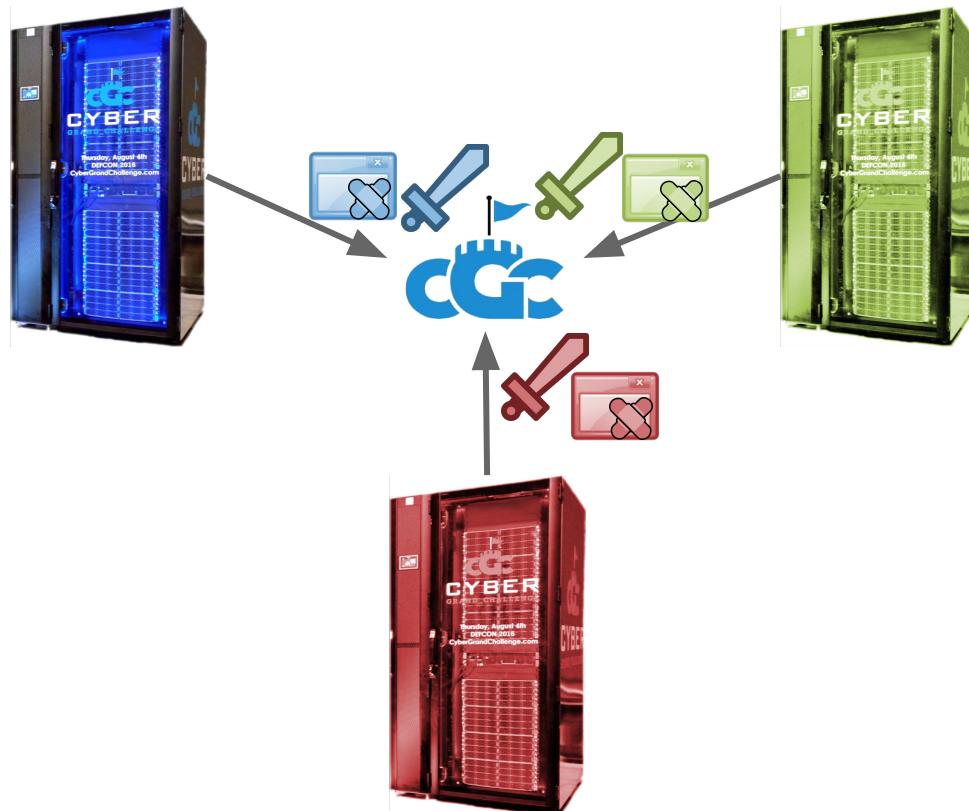
SHELLPHISH



- Round-based game
- Organizers' servers provide:
  - Binaries
    - Linux-like, Intel x86, limited syscalls
    - Console (stdin/stdout)
    - Compiled C programs → no source code
    - Contain one or more vulnerabilities
  - Network traffic
    - Collected during previous rounds

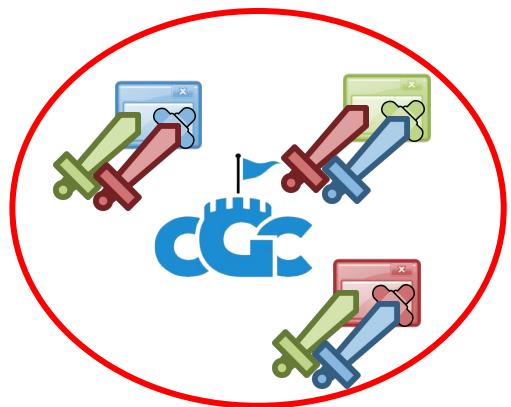
# CGC - How the game worked

SHELLPHISH



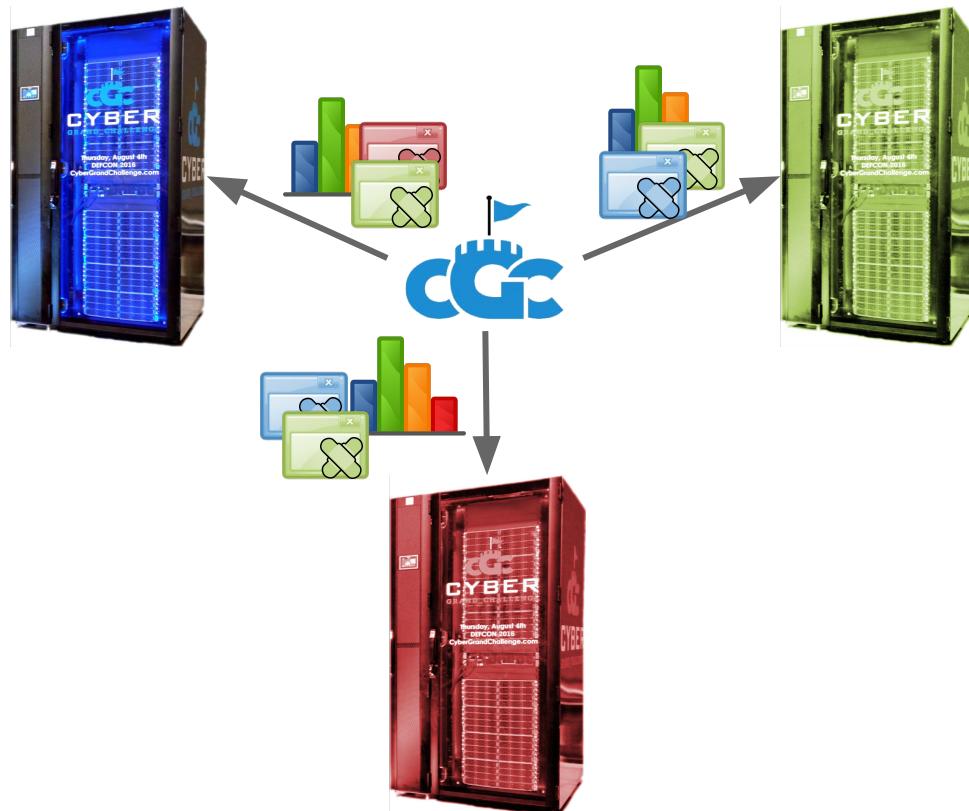
- Teams provide
  - Patched binaries
  - Attacks

# CGC - How the game worked



- Organizers's servers evaluate
  - Attacks vs. Patched binary

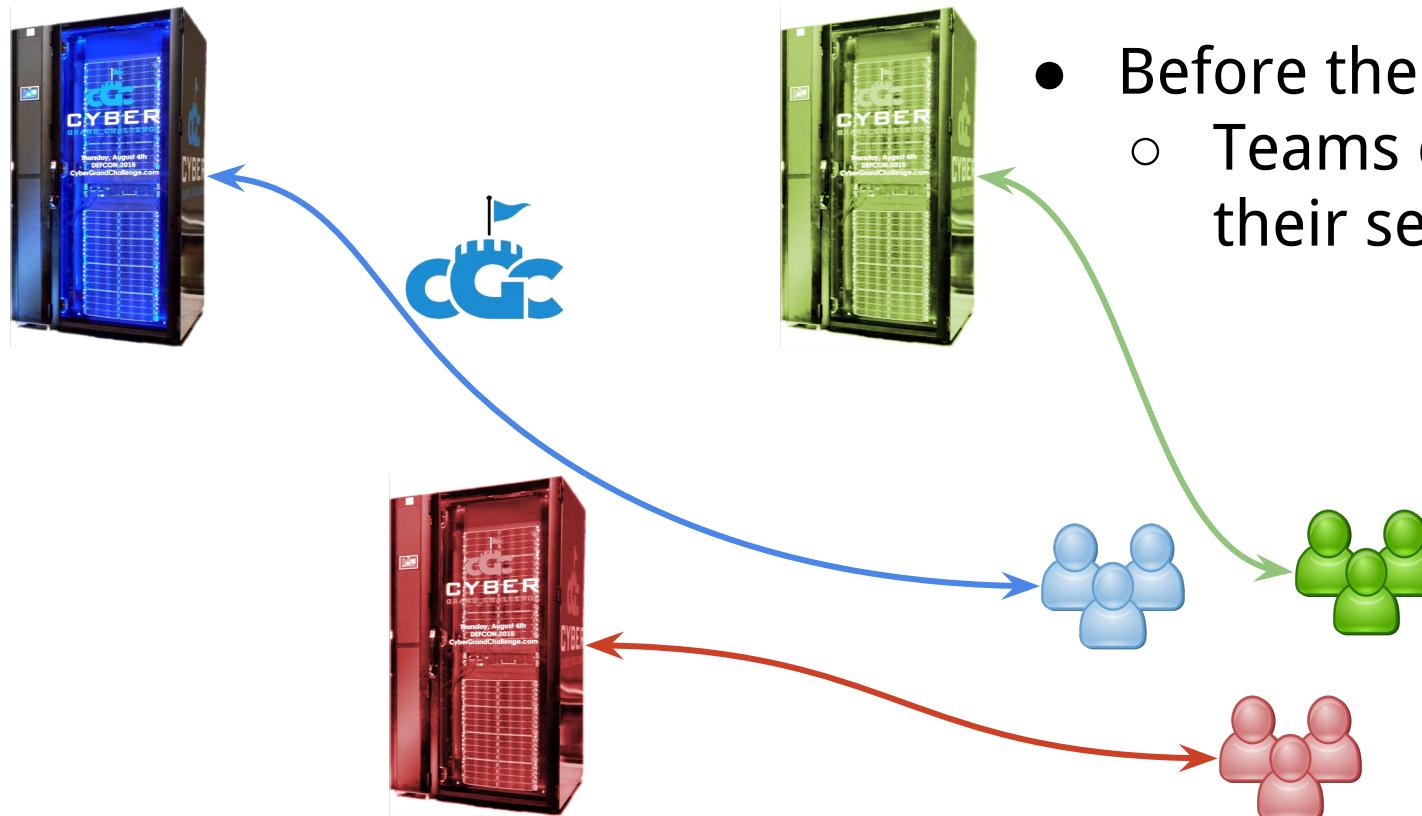
# CGC - How the game worked



- Organizers' servers send back
  - Scores
  - Patched binaries from adversarial teams

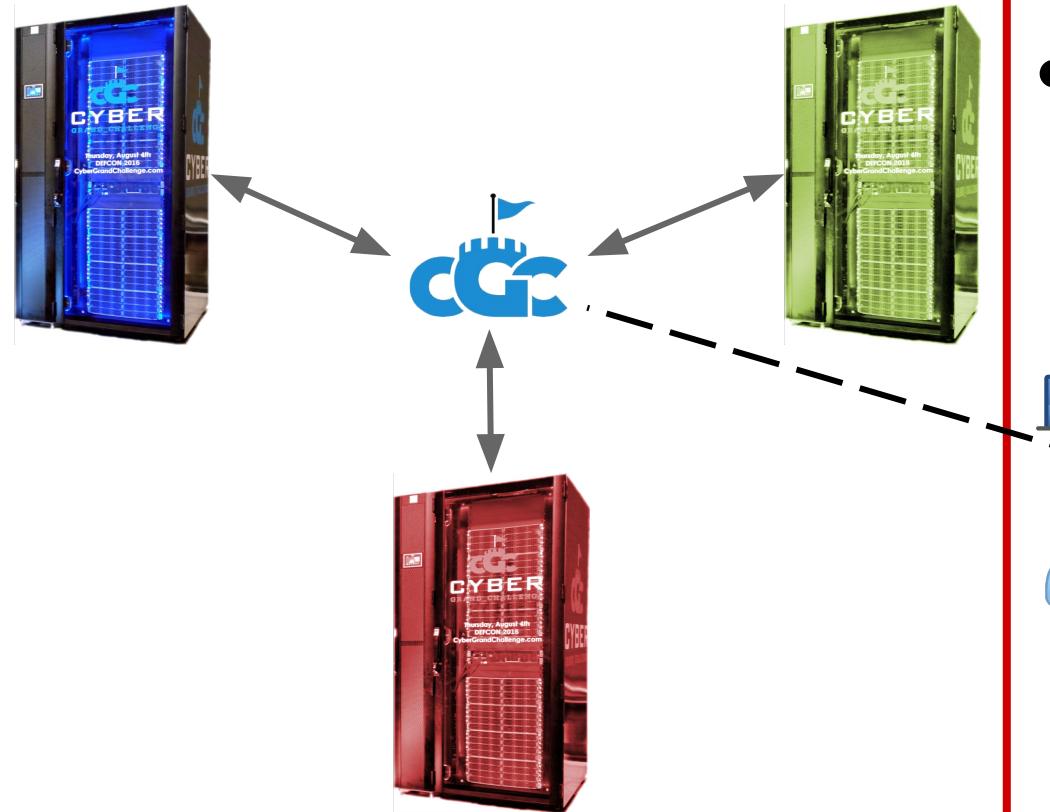
# CGC - How the game worked

SHELLPHISH



# CGC - How the game worked

SHELLPHISH



- During the game (10h)
  - “Certified air gap”
  - Scores are the only data exiting



GRAND CHALLENGE

NINE THE  
AIR GAP

DARPA

DATA OUT

POWER

POWER

# How to play?



**Automatic Binary  
Exploitation**

**Automatic Binary  
Patching**

**Infrastructure**

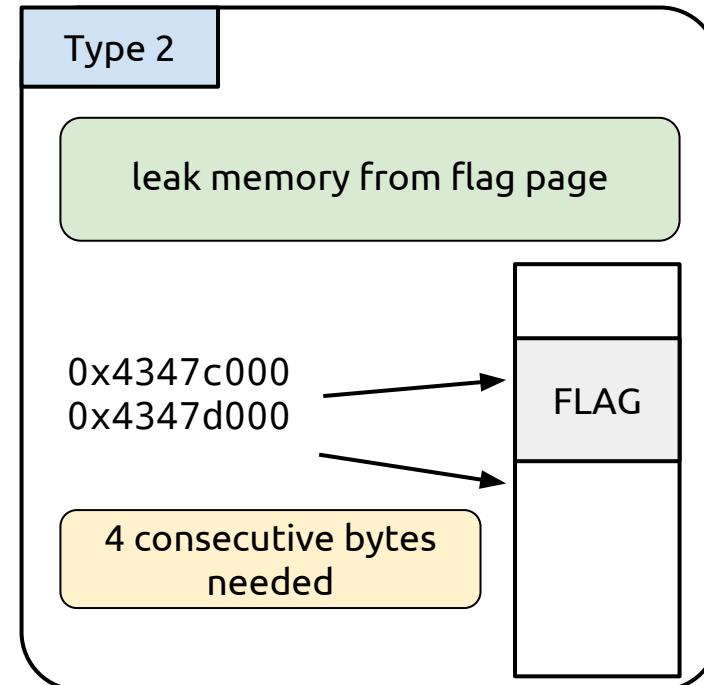
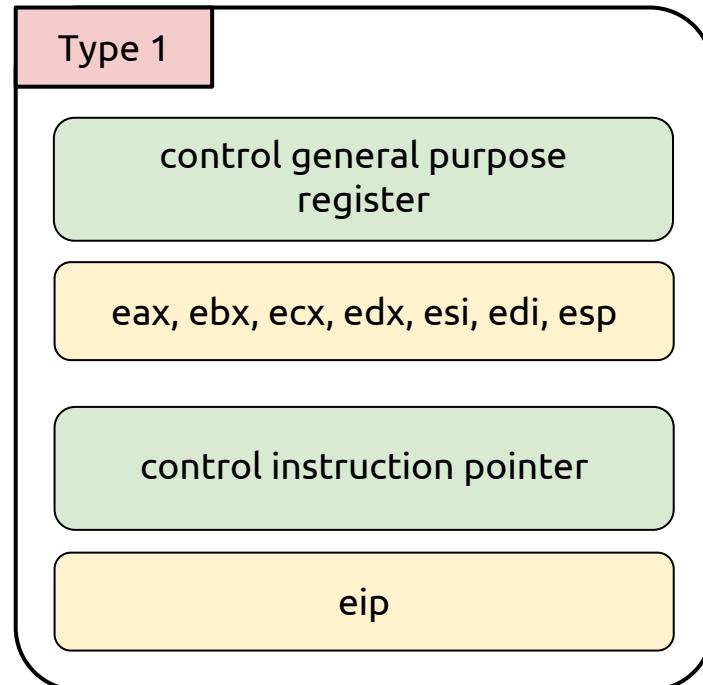
# How to play?

---



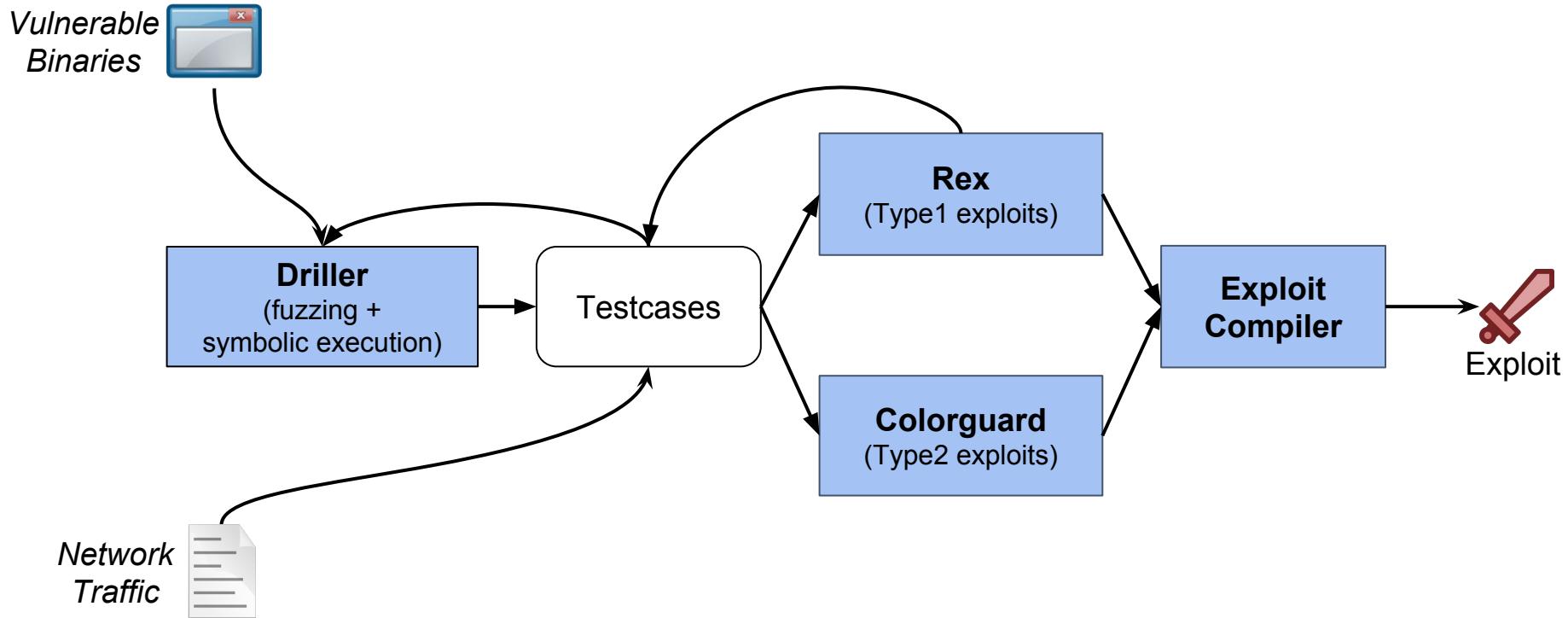
**Automatic Binary  
Exploitation**

# Two types of exploits



- Arbitrary code execution
- Information leak  
(Heartbleed, ASLR base address leak, ...)

# Exploitation pipeline (simplified)



- Execute “the most” of the program →  
Find good inputs to the binary

```
v1 = user_input1()  
v2 = user_input2()  
  
if(v1 < 10){  
    if (v1 == 3){  
        foo()  
    }else if(v1 == 7){  
        bar()  
    }  
}  
}else{  
    if((v1^2 - 19087925*v1)==57263784){  
        function_pointer = v2 + 300  
        function_pointer()  
    }  
}
```

Try many different inputs:  
“1”, “2”, “3”, “4”, “5”, “7”, “8”, ...

# Driller

- Execute “the most” of the program →  
Find good inputs to the binary

```
v1 = user_input1()  
v2 = user_input2()  
  
if(v1 < 10){  
    if (v1 == 3){  
        foo()  
    }else if(v1 == 7){  
        bar()  
    }  
}  
}else{  
    if((v1^2 - 19087925*v1)==57263784){  
        function_pointer = v2 + 300  
        function_pointer()  
    }  
}
```

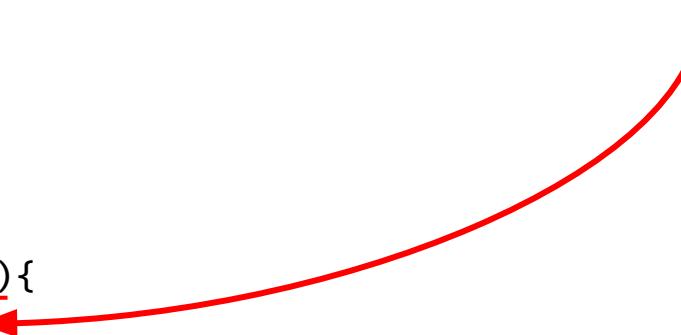
“3” and “7” are “good” testcases:  
they reach new code locations



- Execute “the most” of the program →  
Find good inputs to the binary

```
v1 = user_input1()  
v2 = user_input2()  
  
if(v1 < 10){  
    if (v1 == 3){  
        foo()  
    }else if(v1 == 7){  
        bar()  
    }  
}  
}else{  
    if((v1^2 - 19087925*v1)==57263784){  
        function_pointer = v2 + 300  
        function_pointer()  
    }  
}
```

This is hard to reach randomly



- Execute “the most” of the program →  
Find good inputs to the binary

```
v1 = user_input1()  
v2 = user_input2()
```

```
if(v1 < 10){  
    if (v1 == 3){  
        foo()  
    }else if(v1 == 7){  
        bar()  
    }  
}  
}else{  
    if((v1^2 - 19087925*v1)==57263784){  
        function_pointer = v2 + 300  
        function_pointer()  
    }  
}
```

We can use “symbolic tracing”

Target

- Execute “the most” of the program →  
Find good inputs to the binary

```
v1 = user_input1()  
v2 = user_input2()  
  
if(v1 < 10){  
    if (v1 == 3){  
        foo()  
    }else if(v1 == 7){  
        bar()  
    }  
}  
}else{  
    if((v1^2 - 19087925*v1)==57263784){  
        function_pointer = v2 + 300 ← Target  
        function_pointer()  
    }  
}
```

We can use “symbolic tracing”  
**Constraints:**

- v1 = user\_input1()

- Execute “the most” of the program →  
Find good inputs to the binary

```
v1 = user_input1()  
v2 = user_input2()
```

```
if(v1 < 10){  
    if (v1 == 3){  
        foo()  
    }else if(v1 == 7){  
        bar()  
    }  
}  
}else{  
    if((v1^2 - 19087925*v1)==57263784){  
        function_pointer = v2 + 300  
        function_pointer()  
    }  
}
```

We can use “symbolic tracing”  
**Constraints:**

- $v1 = \text{user\_input1}$
- $\text{not } (v1 < 10)$

Target

- Execute “the most” of the program →  
Find good inputs to the binary

```
v1 = user_input1()  
v2 = user_input2()  
  
if(v1 < 10){  
    if (v1 == 3){  
        foo()  
    }else if(v1 == 7){  
        bar()  
    }  
}else{  
    if((v1^2 - 19087925*v1)==57263784){  
        function_pointer = v2 + 300  
        function_pointer()  
    }  
}
```

We can use “symbolic tracing”  
**Constraints:**

- $v1 = \text{user\_input1}()$
- $\text{not}(v1 < 10)$
- $v1^2 - 19087925*v1 == 57263784$

Target

- Execute “the most” of the program →  
Find good inputs to the binary

```
v1 = user_input1()  
v2 = user_input2()  
  
if(v1 < 10){  
    if (v1 == 3){  
        foo()  
    }else if(v1 == 7){  
        bar()  
    }  
}else{  
    if((v1^2 - 19087925*v1)==57263784){  
        function_pointer = v2 + 300 ← Target  
        function_pointer()  
    }  
}
```

We can use “symbolic tracing”

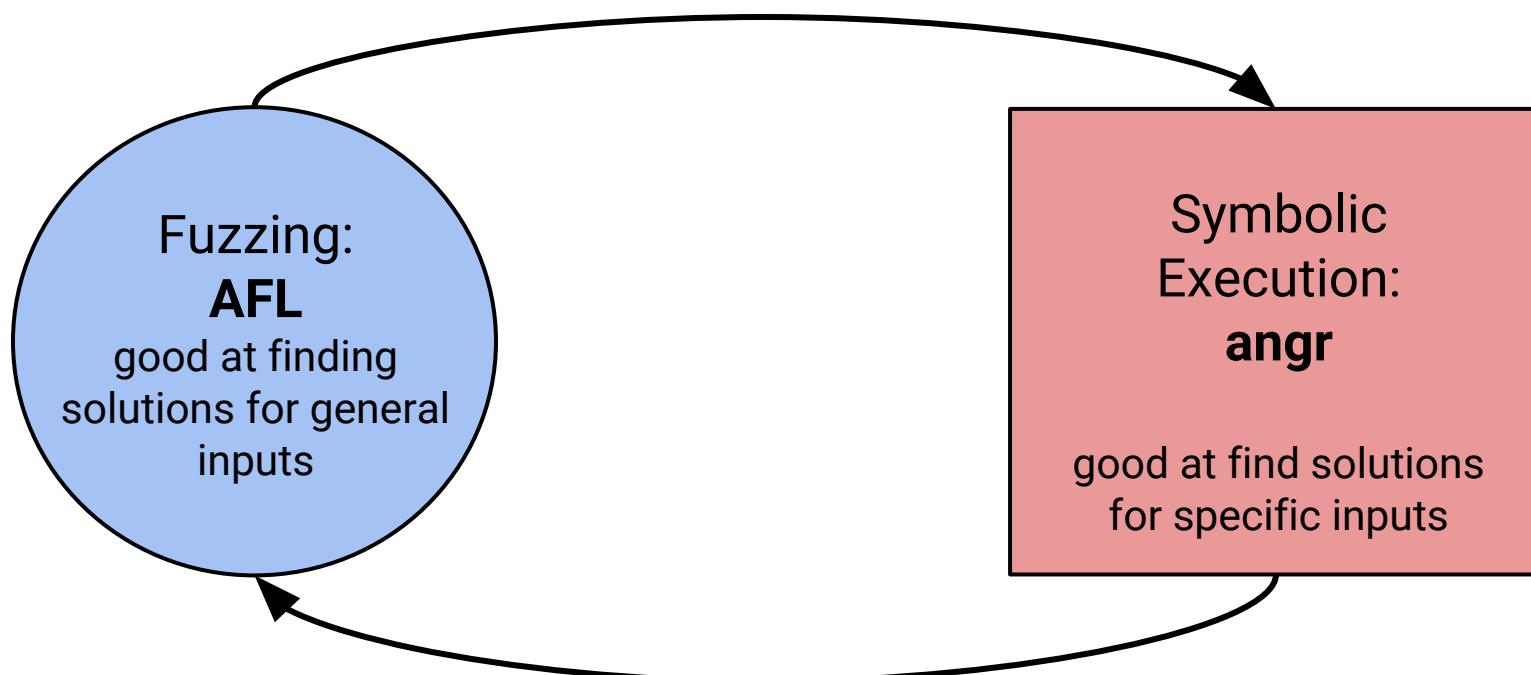
## Constraints:

- $v1 = \text{user\_input1}()$
- $\text{not}(v1 < 10)$
- $v1^2 - 19087925*v1 == 57263784$

## Solve:

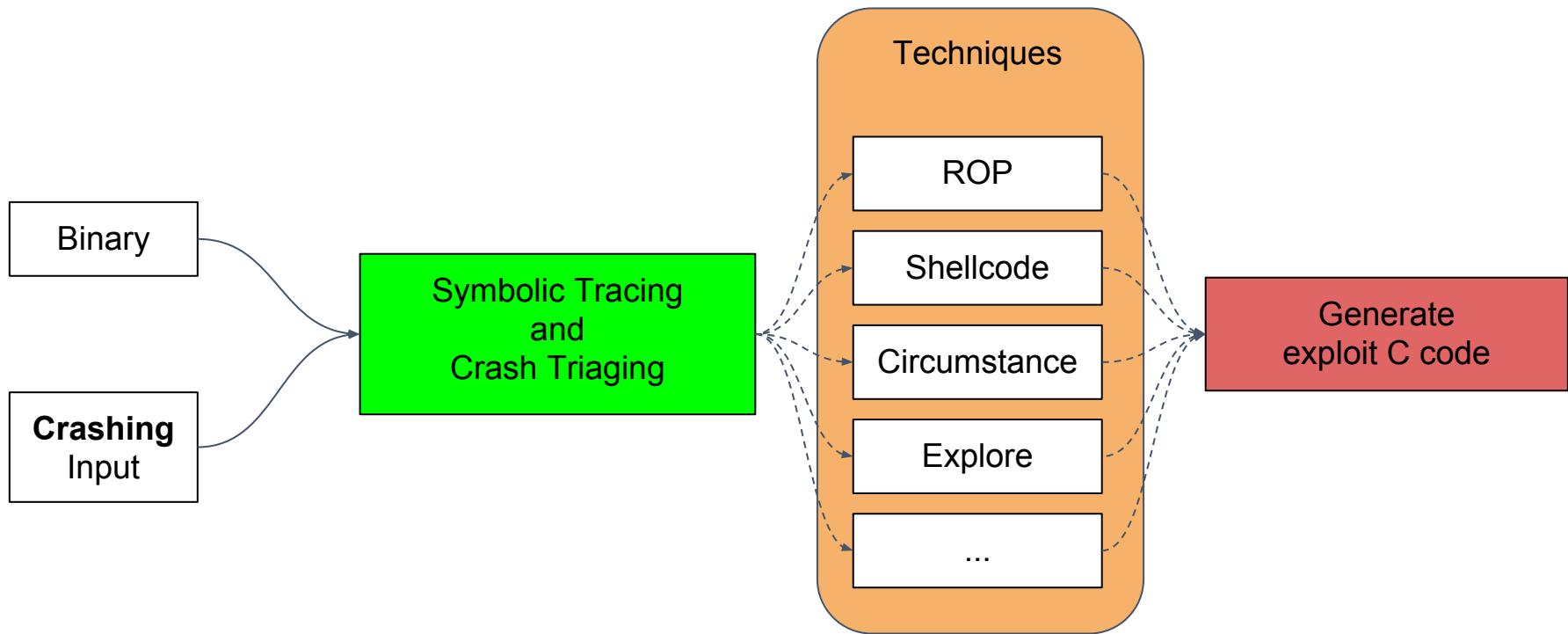
- $\text{user\_input1} = 0x1234238$

# Driller - AFL + angr

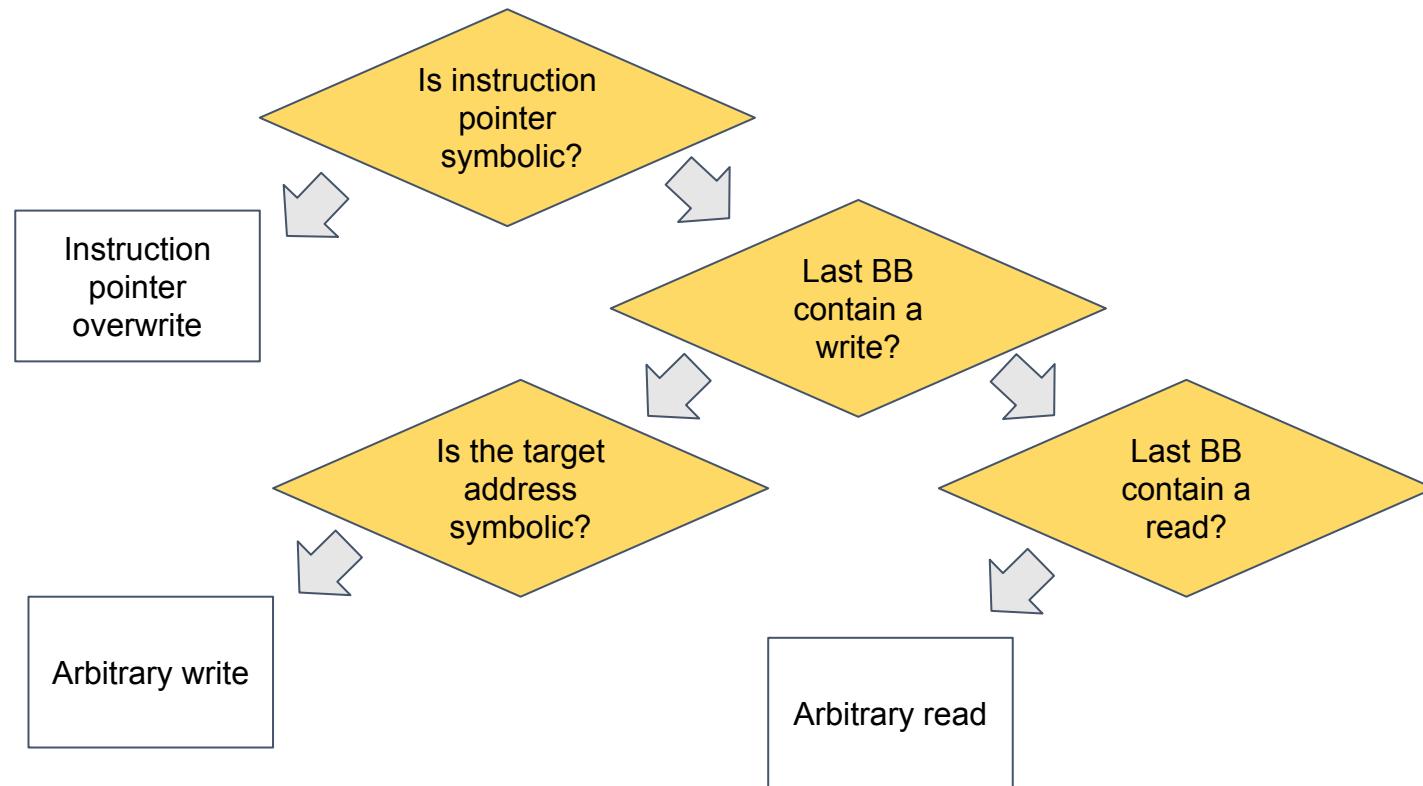


**Driller: Augmenting fuzzing through selective symbolic execution.**

N. Stephens, J. Grosen, C. Salls, A. Dutcher, R. Wang, J. Corbetta, Y. Shoshitaishvili, C. Kruegel, and G. Vigna.  
*at NDSS 2016*



# Rex – Crash Triaging



# Rex – Symbolic Tracing



- Understand “how” to control the crash

```
v1 = user_input1()
v2 = user_input2()

if(v1 > 10){
    if (v1 == 3){
        foo()
    }else if(v1 == 7){
        bar()
    }
}else{
    if((v1^2 - 19087925*v1)==57263784){
        function_pointer = v2 + 300
        function_pointer() ←
    }
}
```

If this instruction is reached  
the program crashes

# Rex – Symbolic Tracing



- Understand “how” to control the crash

```
v1 = user_input1()
v2 = user_input2()

if(v1 > 10){
    if (v1 == 3){
        foo()
    }else if(v1 == 7){
        bar()
    }
}else{
    if((v1^2 - 19087925*v1)==57263784){
        function_pointer = v2 + 300
        function_pointer() ←
    }
}
```

Using symbolic tracing, we know:

```
instruction_pointer =
function_pointer =
v2 + 300 =
user_input2 + 300
```

# Rex – Symbolic Tracing



- Understand “how” to control the crash

```
v1 = user_input1()
v2 = user_input2()

if(v1 > 10){
    if (v1 == 3){
        foo()
    }else if(v1 == 7){
        bar()
    }
}else{
    if((v1^2 - 19087925*v1)==57263784){
        function_pointer = v2 + 300
        function_pointer() ←
    }
}
```

Using symbolic tracing, we know:

$$\text{instruction\_pointer} = \text{user\_input2} + 300$$

Therefore:

- By controlling the user input  
we control the instruction pointer

# Rex – Symbolic Tracing



- Understand “how” to control the crash

```
v1 = user_input1()
v2 = user_input2()

if(v1 > 10){
    if (v1 == 3){
        foo()
    }else if(v1 == 7){
        bar()
    }
}else{
    if((v1^2 - 19087925*v1)==57263784){
        function_pointer = v2 + 300
        function_pointer() ←
    }
}
```

Using symbolic tracing, we know:

$$\text{instruction\_pointer} = \text{user\_input2} + 300$$

Therefore:

- By controlling the user input  
we control the instruction pointer
- If we want:

$$\text{instruction\_pointer} = X$$

we have to set:

$$\text{user\_input2} = X - 300$$

# Rex – Techniques



- Crashing input → Exploit

Instruction pointer control

Jump to Shellcode

Pivot to ROP chain

“Circumstantial”

Arbitrary Write

Point-to-Data

Explore for Exploit

Arbitrary Read

Point-to-Flag

Point-to-Data

# Rex – Technique: Jump to Shellcode



- We want to place shellcode in buffer and jump to it

```
v1 = user_input1()  
buffer = base64_decode(user_input2())  
//...  
  
function_pointer = v1 + 300  
function_pointer()
```

# Rex – Technique: Jump to Shellcode



- We want to place shellcode in buffer and jump to it

```
v1 = user_input1()  
buffer = base64_decode(user_input2())  
//...  
  
function_pointer = v1 + 300  
function_pointer()
```

Using symbolic tracing, we know:

instruction\_pointer = user\_input1 + 300  
buffer = base64\_decode(user\_input2)

# Rex – Technique: Jump to Shellcode



- We want to place shellcode in buffer and jump to it

```
v1 = user_input1()  
buffer = base64_decode(user_input2())  
//...  
  
function_pointer = v1 + 300  
function_pointer()
```

Using symbolic tracing, we know:

instruction\_pointer = user\_input1 + 300  
buffer = base64\_decode(user\_input2)

We want:

instruction\_pointer = &(buffer)  
buffer = shellcode

# Rex – Technique: Jump to Shellcode



- We want to place shellcode in buffer and jump to it

```
v1 = user_input1()  
buffer = base64_decode(user_input2())  
//...  
  
function_pointer = v1 + 300  
function_pointer()
```

Using symbolic tracing, we know:

instruction\_pointer = user\_input1 + 300  
buffer = base64\_decode(user\_input2)

We want:

instruction\_pointer = &(buffer)  
buffer = shellcode

Therefore:

user\_input1 = &(buffer) - 300  
user\_input2 = base64\_encode(shellcode)

# Colorguard

---



- Memory-leak (Type 2) exploits
- Use symbolic tracing
- Analyze all inputs

# Colorguard



- Memory-leak (Type 2) exploits are also generated using symbolic tracing

```
v1 = user_input1()  
//...  
  
printed_value = array[v1]  
print(printed_value) ←
```

Using symbolic tracing, we know:

```
printed_value = *(&array + v1)  
v1 = user_input1
```

We want:

```
printed_value = flag_page[0]
```

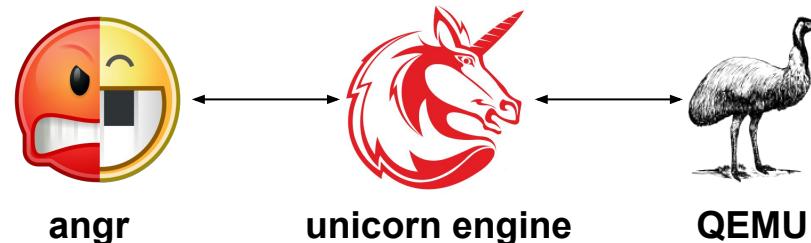
Therefore:

```
user_input1 = (&flag_page) - (&array)
```

# Colorguard – Unicorn Engine



- Every testcase can potentially leak the flag page
- Full symbolic tracing of every testcase is too slow
- angr + unicorn engine (QEMU wrapper)
  - Execute “most” of the code in QEMU



3,000,000 times slower  
than a real CPU

2~5 times slower  
than a real CPU

# How to play?

---



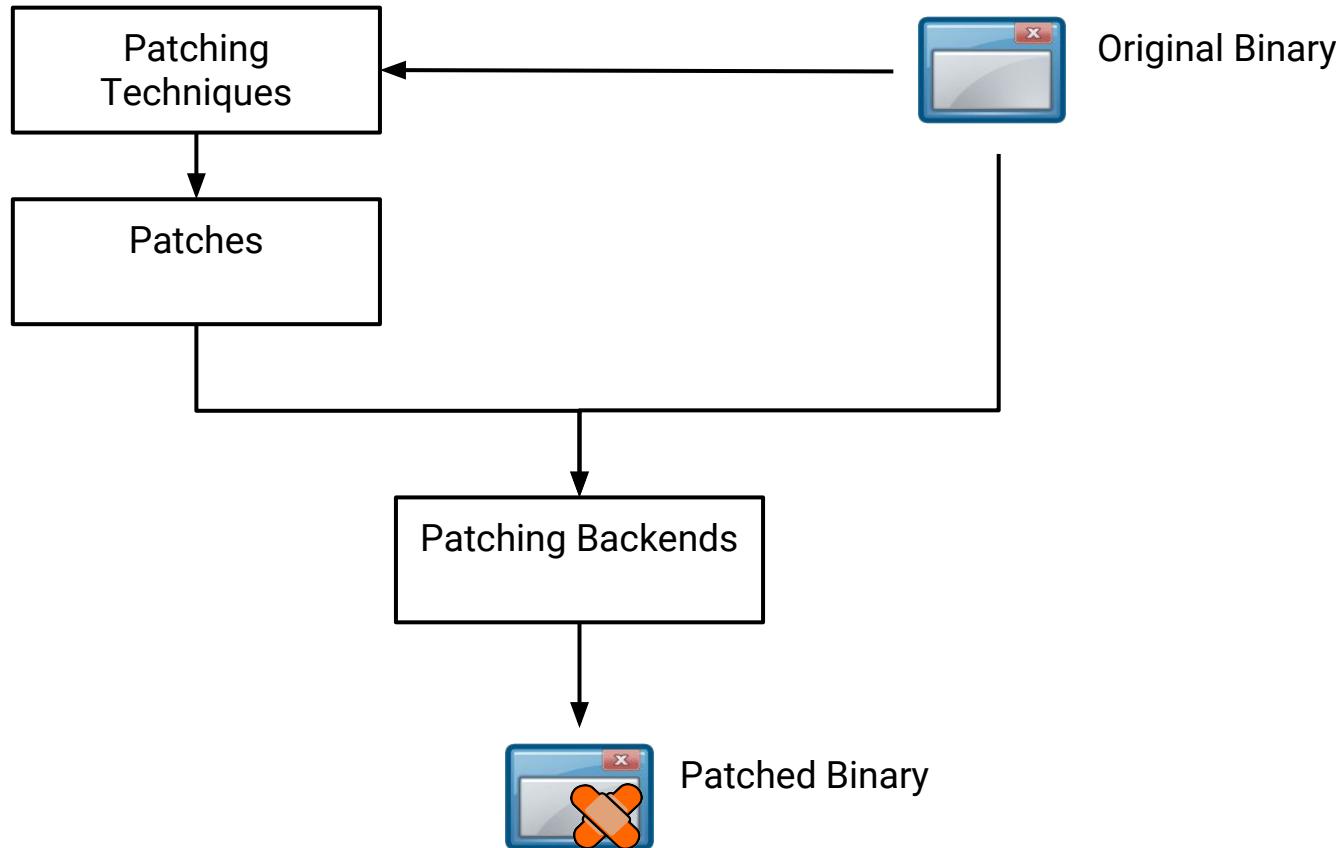
**Automatic Binary  
Patching**

# Automatic Patching

---



- Prevent binary from being exploit
- Preserve binary functionality
- Preserve binary performance
  - speed
  - memory usage
  - disk space
- Prevent analysis from other teams

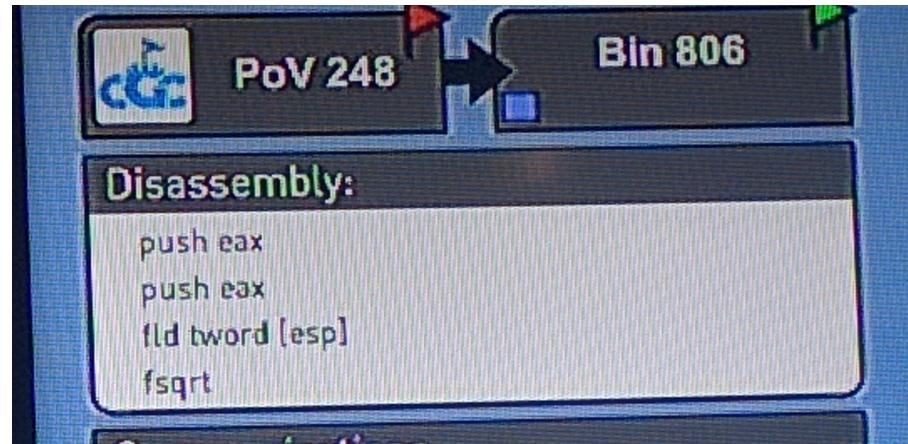


- Defensive Techniques
  - Return pointer encryption
  - Protect indirect calls/jmps
  - Extended Malloc allocations
  - Randomly shift the stack (ASLR)
  - ...

- Adversarial Techniques

- Detect QEMU

```
mov eax, 0x1  
push eax  
push eax  
push eax  
fld TBYTE PTR [esp]  
fsqrt
```



- Backdoor

- ...

- Making the original binary faster →  
Our patches can be slower!
- Optimization Techniques:
  - Constant Propagation
  - Dead Assignment Elimination
  - ...

# Patcherex – Backends

---



- Patching Backends
  - Inject code/data in an existing binary
  - No source code
  - No symbols

# Patcherex – Backends



- How to inject code without breaking functionality?

0x0 : mov eax, 0x11

0x5 : jmp eax

0x7 : mov edx, 0x11223344

0xc : mov ebx, 0x55667788

0x11: mov ecx, ebx



0x0 : mov eax, 0x11

0x5 : jmp eax

0x7: call inserted\_function

0xc : mov edx, 0x11223344

0x11: mov ebx, 0x55667788

0x16: mov ecx, ebx

# Patcherex – Backends



- Detour Backend
  - Try to add code without moving the original one
  - Not always possible
  - Slow (requires a lot of additional jmp instructions)

0x0 : mov eax, 0x11	0x0 : mov eax, 0x11
0x5 : jmp eax	0x5 : jmp eax
0x7 : mov edx, 0x11223344 →	0x7 : jmp out1 → mov edx, 0x11223344
0xc : mov ebx, 0x55667788	0xc : mov ebx, 0x55667788 ↙ call inserted_function
0x11: mov ecx, ebx	0x11: mov ecx, ebx ↘ jmp 0xc

# Patcherex – Backends



- Reassembler Backend
  - Recover original “program symbols”
  - More efficient code
  - (Slightly) less reliable

**Ramblr: Making Reassembly Great Again.**

R. Wang, Y. Shoshtaishvili, A. Bianchi, A. Machiry, J. Grosen, P. Grosen, C. Kruegel, G. Vigna

*In NDSS 2017*

# Patcherex – Backends



0x0 : mov eax, 0x11

0x5 : jmp eax

0x7 : mov edx, 0x11223344 →

0xc : mov ebx, 0x55667788

0x11: mov ecx, ebx

mov eax, \_label1

jmp eax

mov edx, 0x11223344 →

mov ebx, 0x55667788

\_label1:

mov ecx, ebx

mov eax, \_label1

jmp eax

call inserted\_function

mov edx, 0x11223344

mov ebx, 0x55667788

\_label1:

mov ecx, ebx

0x0 : mov eax, 0x16

0x5 : jmp eax

0x7 : call inserted\_function

0xc : mov edx, 0x11223344

0x11: mov ebx, 0x55667788

0x16: mov ecx, ebx

mov eax, \_label1

jmp eax

mov edx, 0x11223344 →

mov ebx, 0x55667788

\_label1:

mov ecx, ebx

call inserted\_function

# How to play?

---



**Infrastructure**

# Infrastructure

SHELL(PHISH)

- Our code had to run for 10 hours on:  
64 servers, 16TB of RAM, 2560 cores
- No human intervention →  
No possibility of failure!
- Extremely hard to test the full system
  - A lot of test cases
  - Testing after every single git push



# Infrastructure

---



- Separate and (mostly) independent *tasks*
- Every task run in a separate container
  - Docker
- Tasks are distributed “transparently” among servers
  - Kubernetes

---

# What Happened?

# Results



- Exploitation
  - 2442 Exploits generated
  - 1709 Exploits for 14/82 challenges with 100% Reliability
  - Longest exploit: 3791 lines of C code
  - crackaddr: 517 lines of C code
  - **Shellphish exploited the most binaries!**
- Defense
  - Only 12/82 services were compromised
  - Second best team in terms of defense points

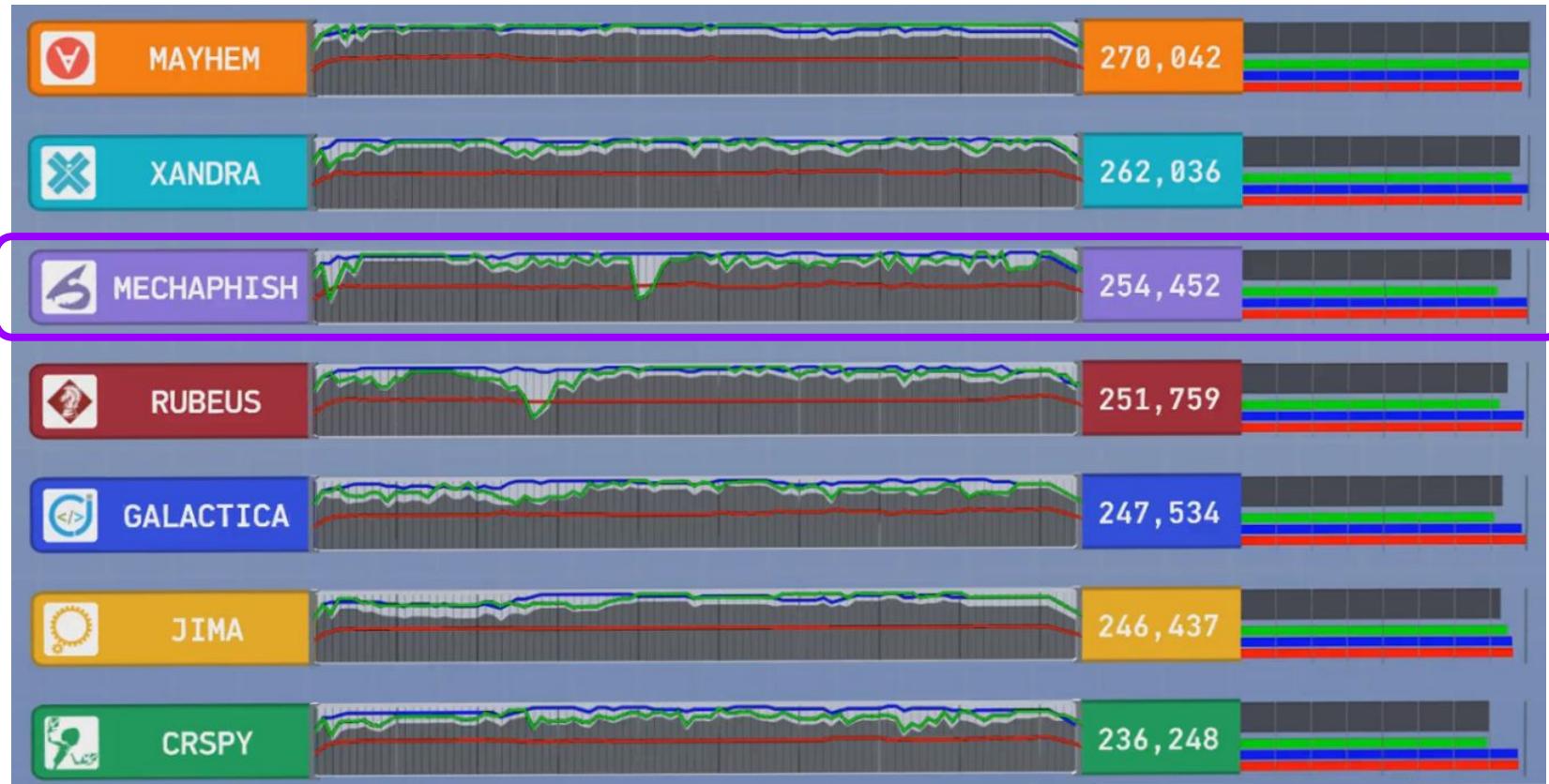
# Third Place! ::\OwO/::



- Third Place!
- Happiness!
- First among University-only teams
- First among unfunded teams

# Results

SHELLPHISH



\$\$\$

---



750,000 \$ +

750,000 \$ =

1,500,000 \$

# What went wrong

---



- Our strategy was not ideal ~~ 囧
  - Patch everything!
  - Score penalty
  - Only 20/82 binaries were exploited in total

# Open source release



- Open source all the code!



# Open source release



- About 100,000 lines of Python code
- [github.com/shellphish](https://github.com/shellphish)
  - Core, independent components: REX, Patcherex, ...
- [github.com/mehaphish](https://github.com/mehaphish)
  - Infrastructure, utilities, and documentation
- [github.com/angr](https://github.com/angr)
  - Binary analysis framework, symbolic execution, ...

# Standing on the shoulders of giants



AFL



# Future Directions

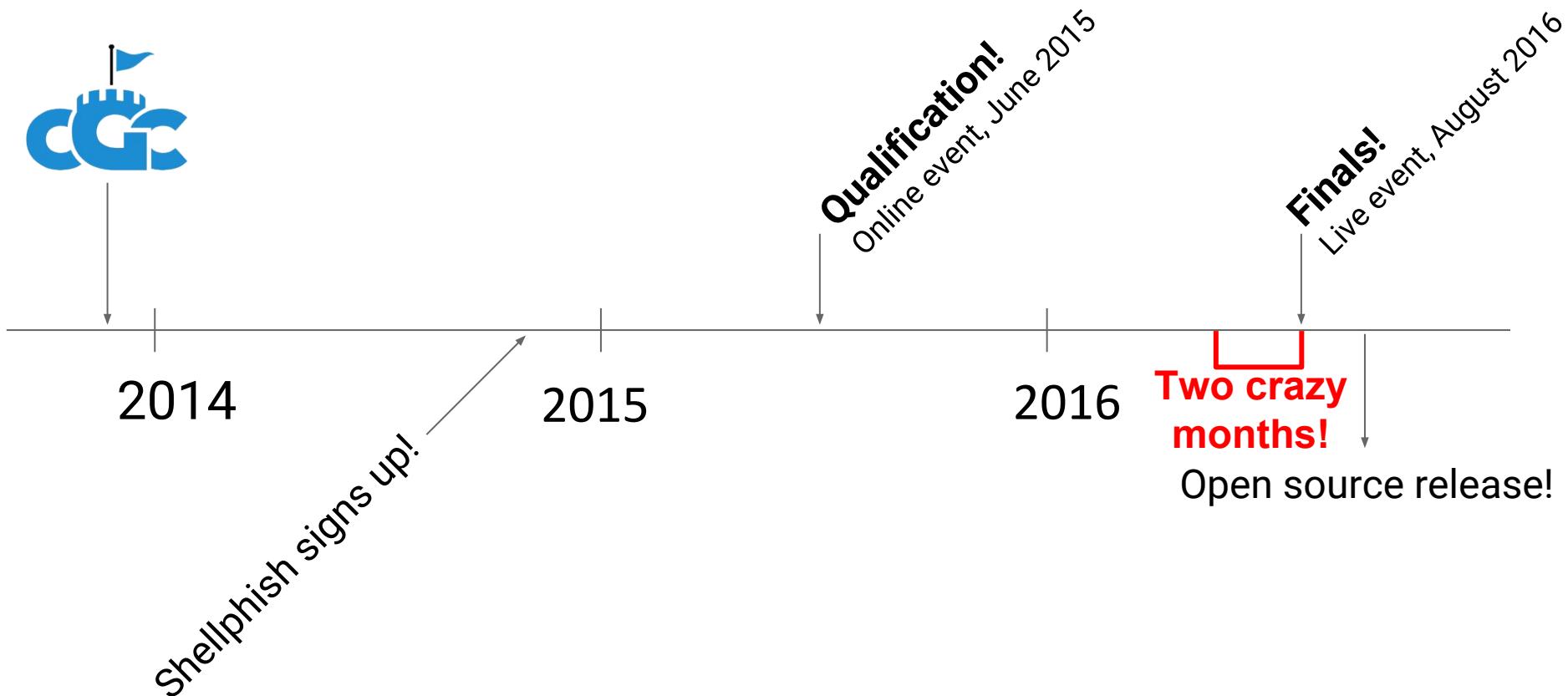
---



- Human-assisted automatic exploitation and defense
- You can contribute
  - Port code to non-CGC architecture
  - Are you a student?  
Looking for an internship?  
Master thesis?  
Wanting to do a PhD?  
Want a free Shellphish Tshirt/sticker?



# CGC - Timeline





WORK | LUNGES

STRETCH  
LIGHT  
MEDIUM  
HEAVY  
FL DEET  
BITFLIP

BAGELS

CRAFT

CHAPTER 3

NRFN3

1

1  
3

1  
3

1

found no  
13+

NRFN\_36  
NRFN\_23  
NRFN\_21  
NRFN\_11  
NRFN\_14  
Demo\_39  
Craft\_70  
Craft\_71  
NRFN\_5

closed C  
closed C

① → Computer  
② → Poll  
③ → 2-1 handshake  
④ → C71

VALID POLL  
NETWORK POLL  
SANITIZER

4

IDS RULES

Syntax

POLL  
CREATOR

AFL  
TEST CASES

IDS UDP SYN  
DATA  
HTTP  
AT  
20000

woundedwoman





BALLY'S



## UNVERIFIED WINNERS



XANDRA  
TECHX



MAYHEM  
FOR ALLSECURE



MECHANICAL  
PHISH  
SHELLPHISH

MACHINICAL PRISM  
TRIPOD RANGE



HESKETH



RE



# Questions?

## References:

- all the technical details: “very soon” published in a “popular security ezine”
- this presentation: [goo.gl/RvDbxS](http://goo.gl/RvDbxS)
- CGC final event show: [youtu.be/n0kn4mDXY6I](https://youtu.be/n0kn4mDXY6I)
- Twitter:@shellphish
- Twitter team: @anton00b - @caovc - @giovanni\_vigna - @jac\_arc - @ltFish\_@machiry\_msdic - @nebirhos - @rhehmot - @zardus
- email: [antonio@cs.ucsb.edu](mailto:antonio@cs.ucsb.edu) - [team@shellphish.net](mailto:team@shellphish.net)
- Github: [github.com/shellphish](https://github.com/shellphish) - [github.com/mecaphish](https://github.com/mecaphish) - [github.com/angr](https://github.com/angr)