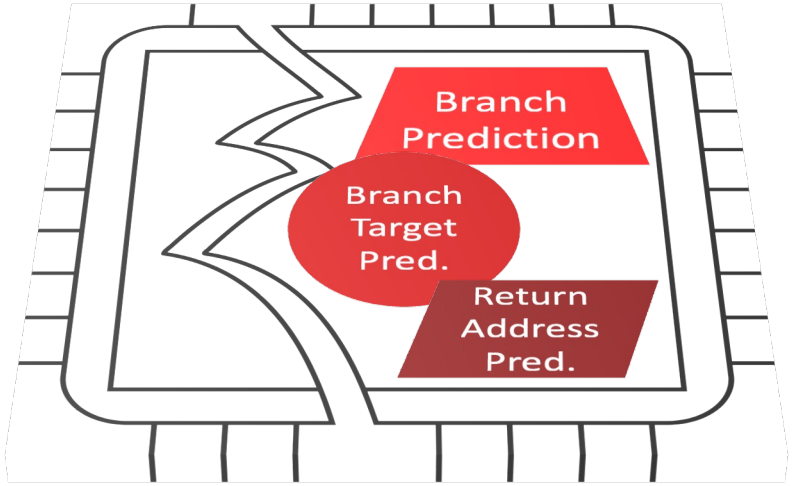
GadgetMeter: Quantitatively and Accurately Gauging the Exploitability of Speculative Gadgets

Qi Ling, Yujun Liang, Yi Ren, Baris Kasikci, Shuwen Deng

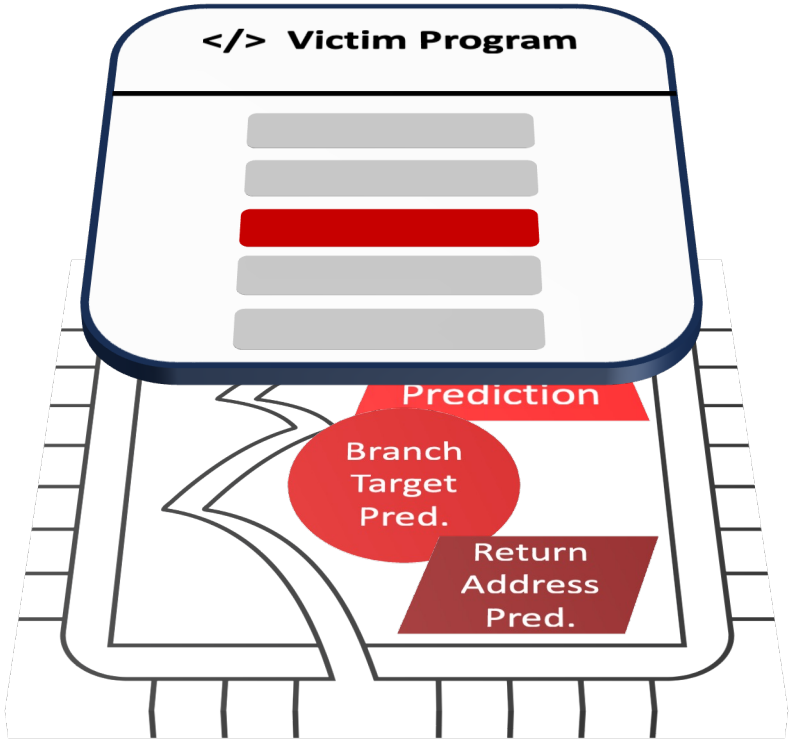
Purdue University, University of Washington, Tsinghua University

Spectre Attacks

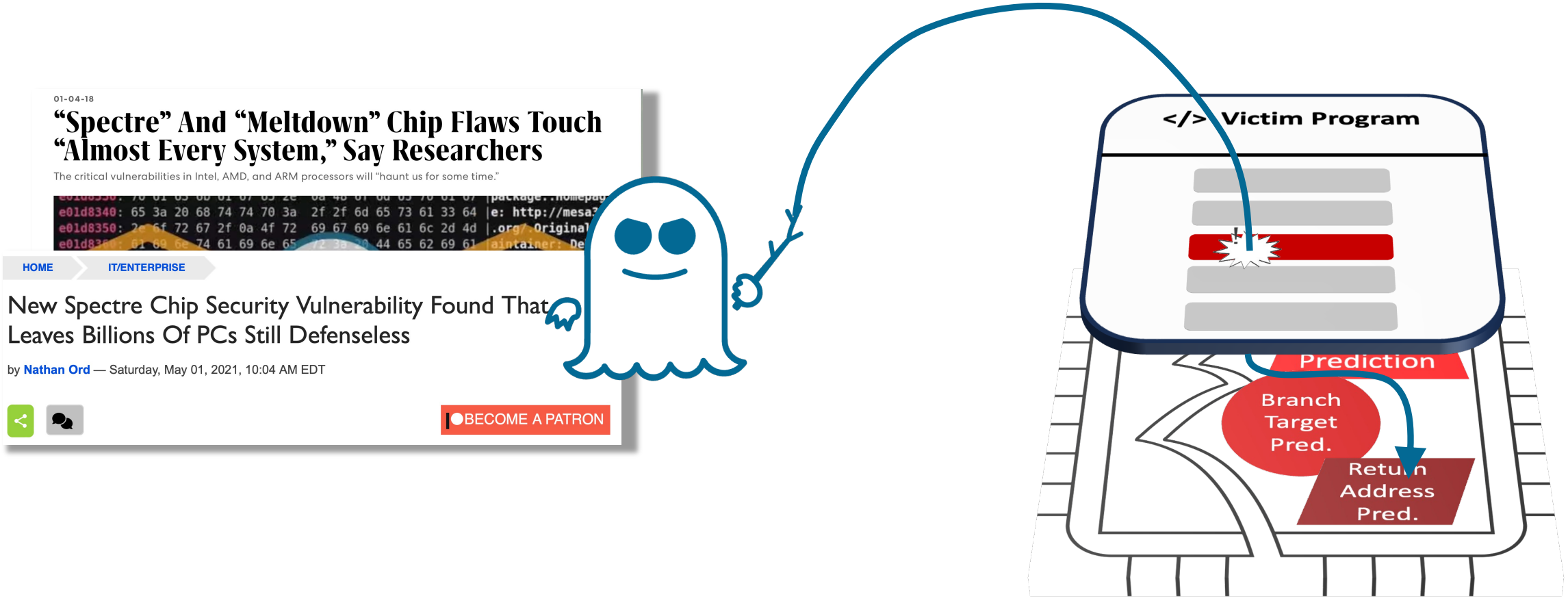
Spectre Attacks



Spectre Attacks



Spectre Attacks



Spectre Attacks

01-04-18

“Spectre” And “Meltdown” Chip Flaws Touch “Almost Every System,” Say Researchers

The critical vulnerabilities in Intel, AMD, and ARM processors will “haunt us for some time.”

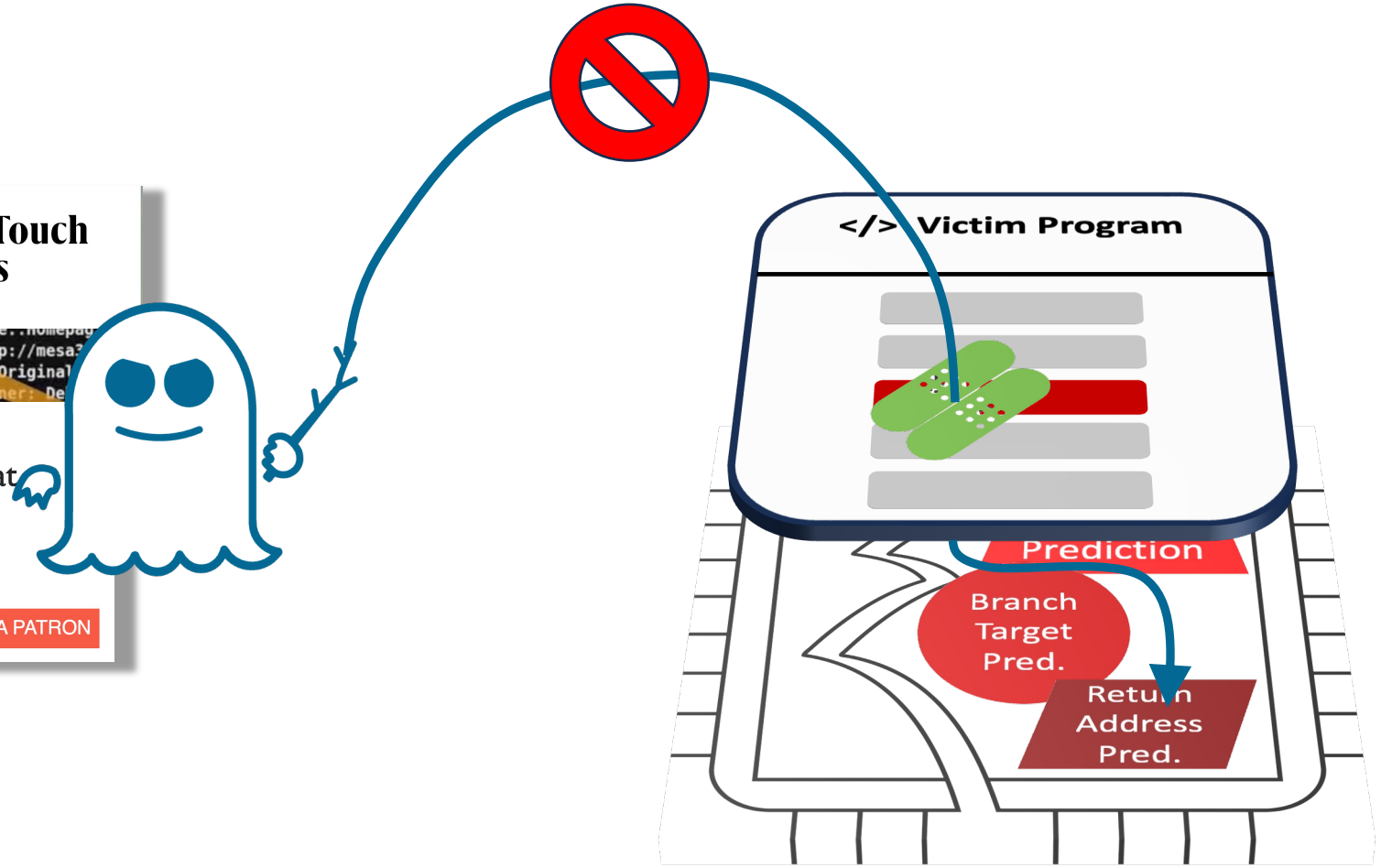
```
e01d8340: 65 3a 20 68 74 74 70 3a  2f 2f 6d 65 73 61 33 64 |e: http://mesa3d
e01d8350: 2e 6f 72 67 2f 0a 4f 72  69 67 69 6e 61 6c 2d 4d |.org/.Original-M
```

HOME IT/ENTERPRISE

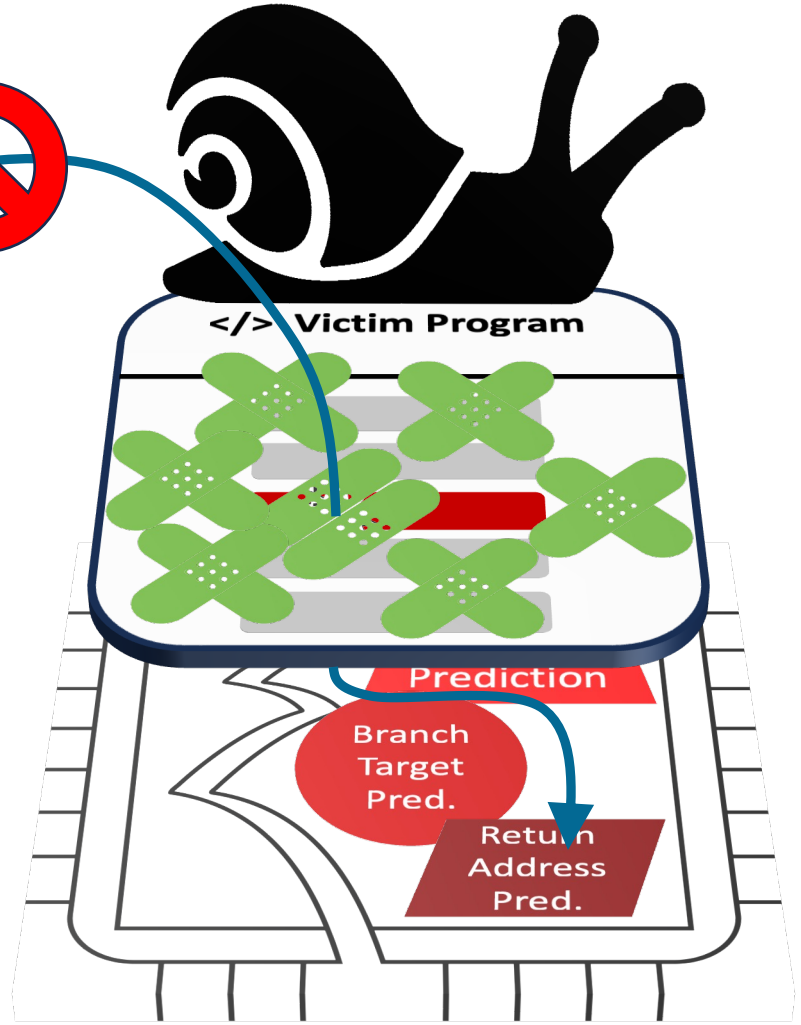
New Spectre Chip Security Vulnerability Found That Leaves Billions Of PCs Still Defenseless

by Nathan Ord — Saturday, May 01, 2021, 10:04 AM EDT

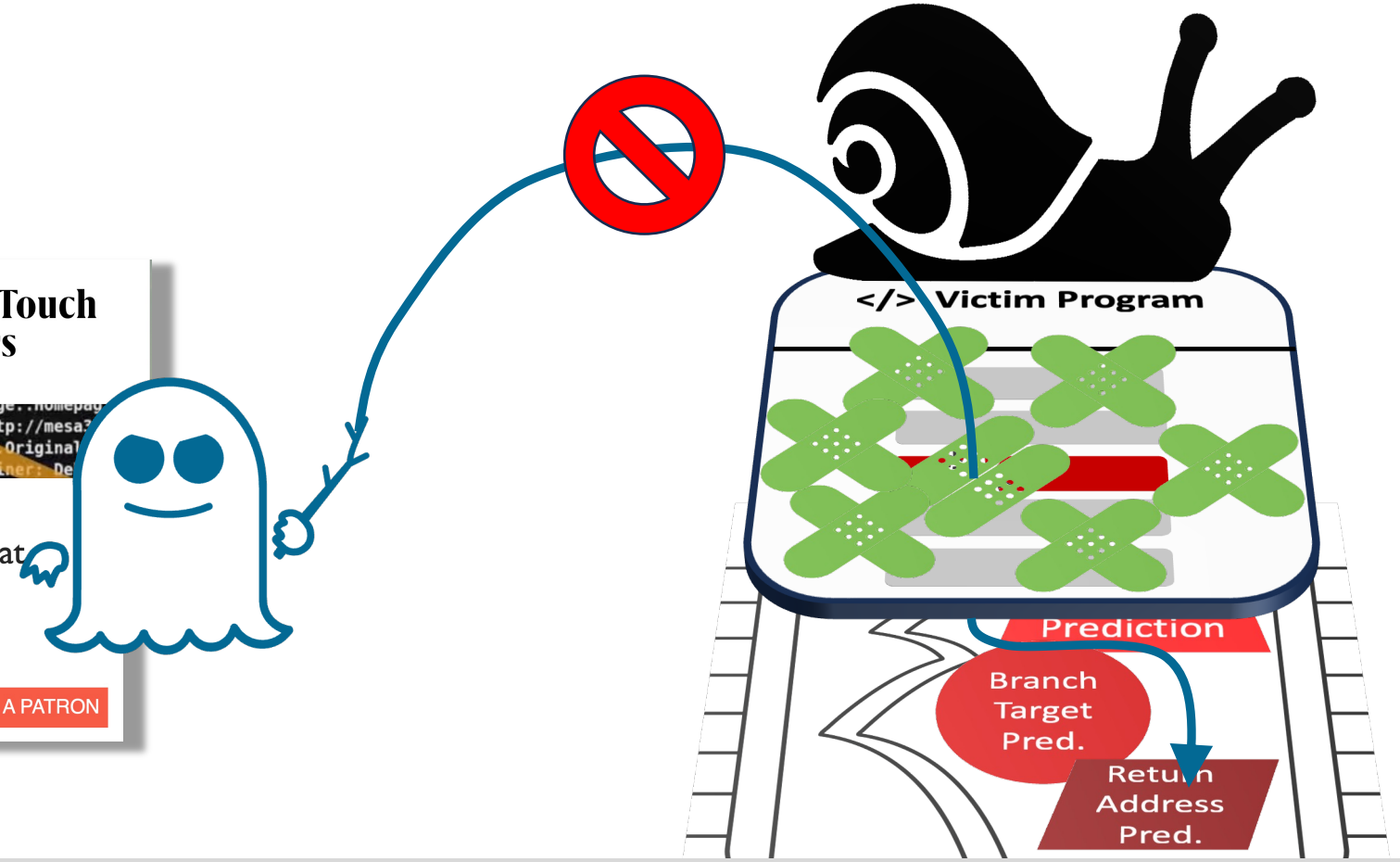
BECOME A PATRON



Spectre Attacks



Spectre Attacks

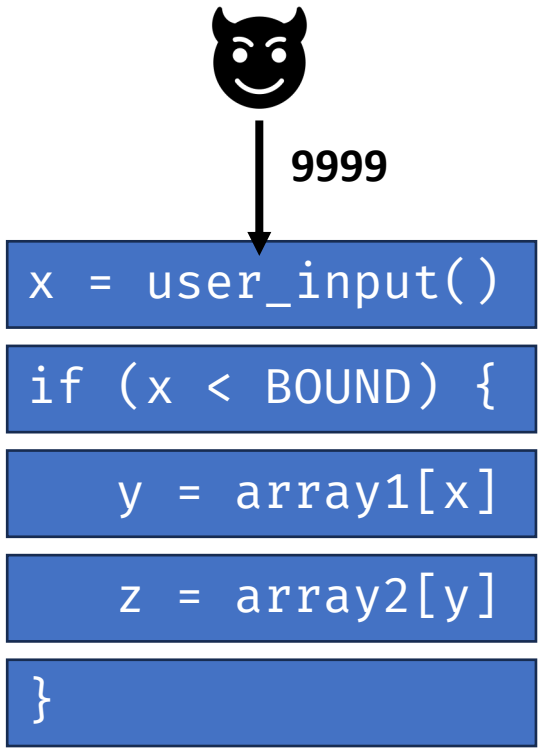


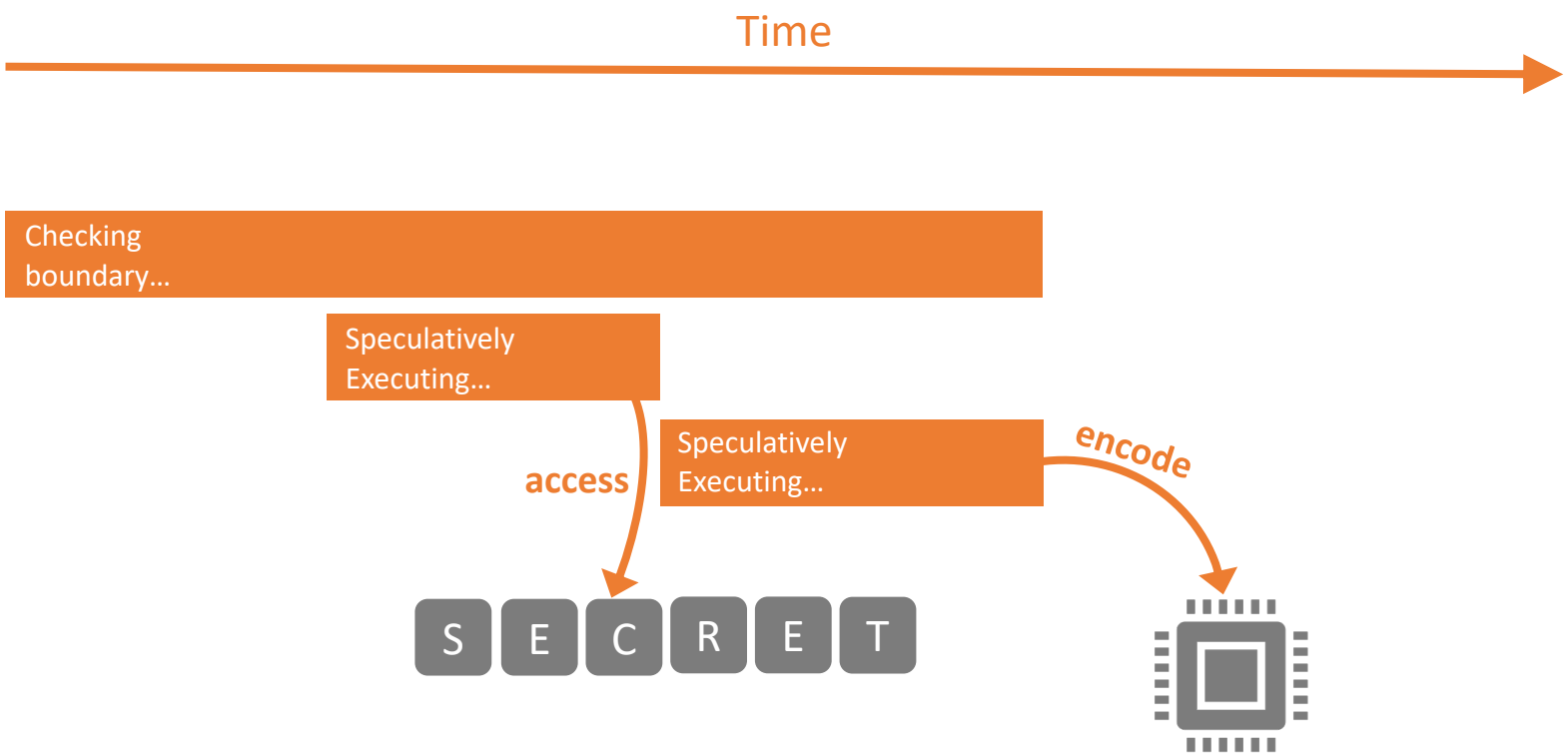
Accurate detection of Spectre gadgets.

What is a Spectre gadget?

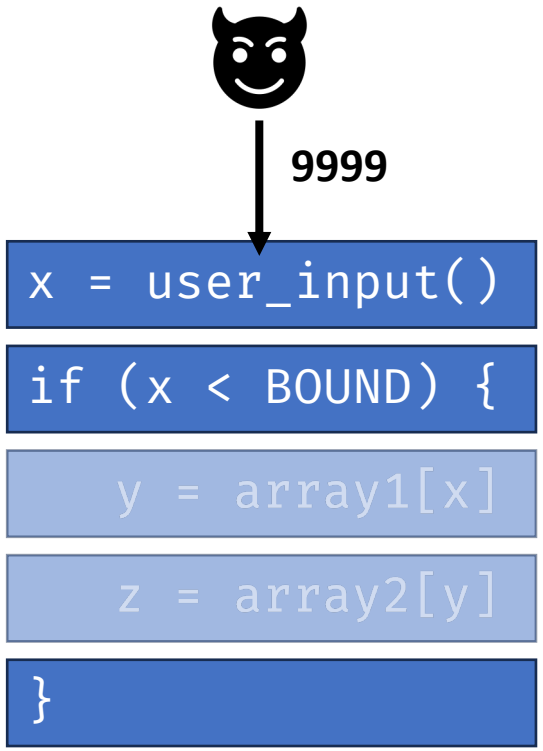
What is a Spectre gadget?

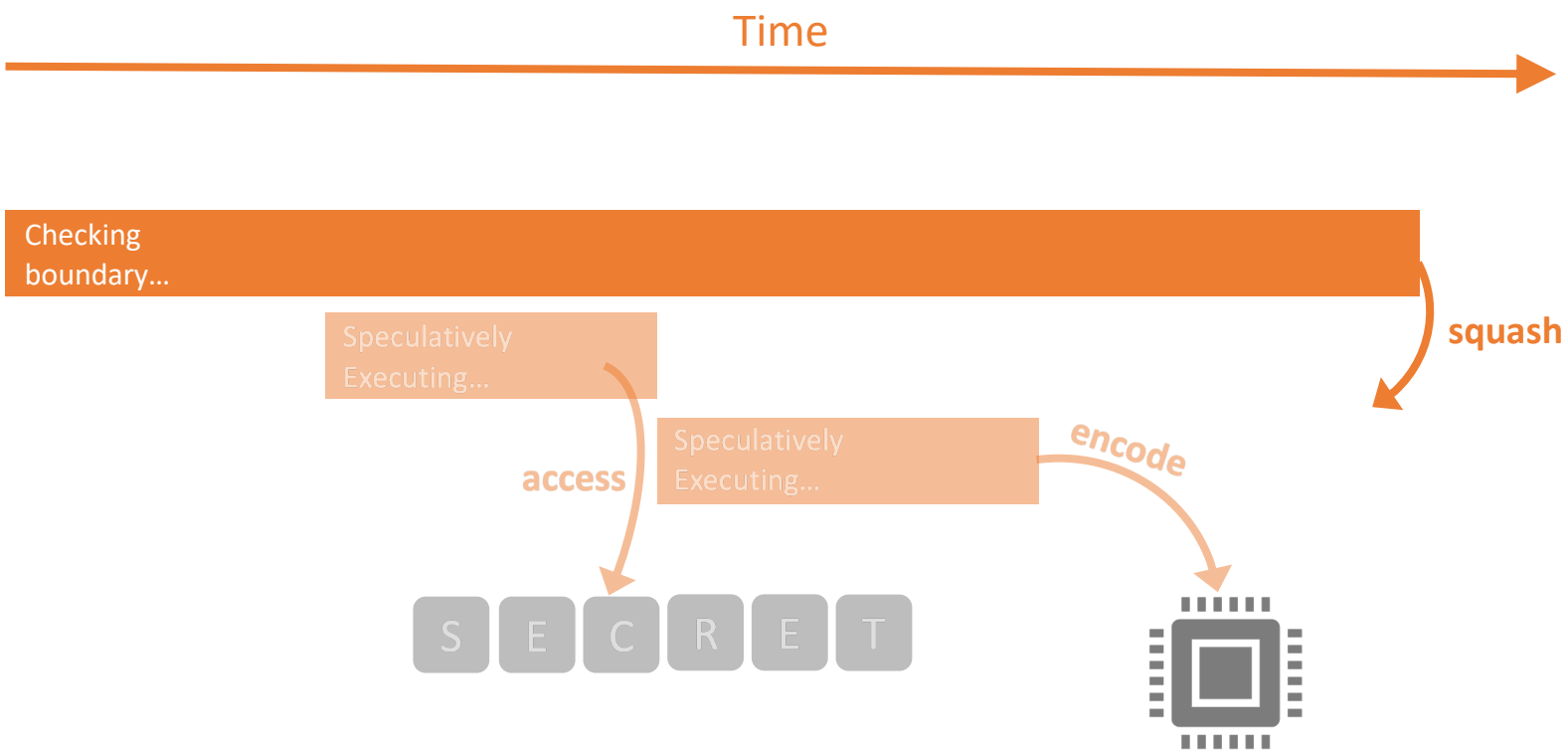
What is a Spectre gadget?



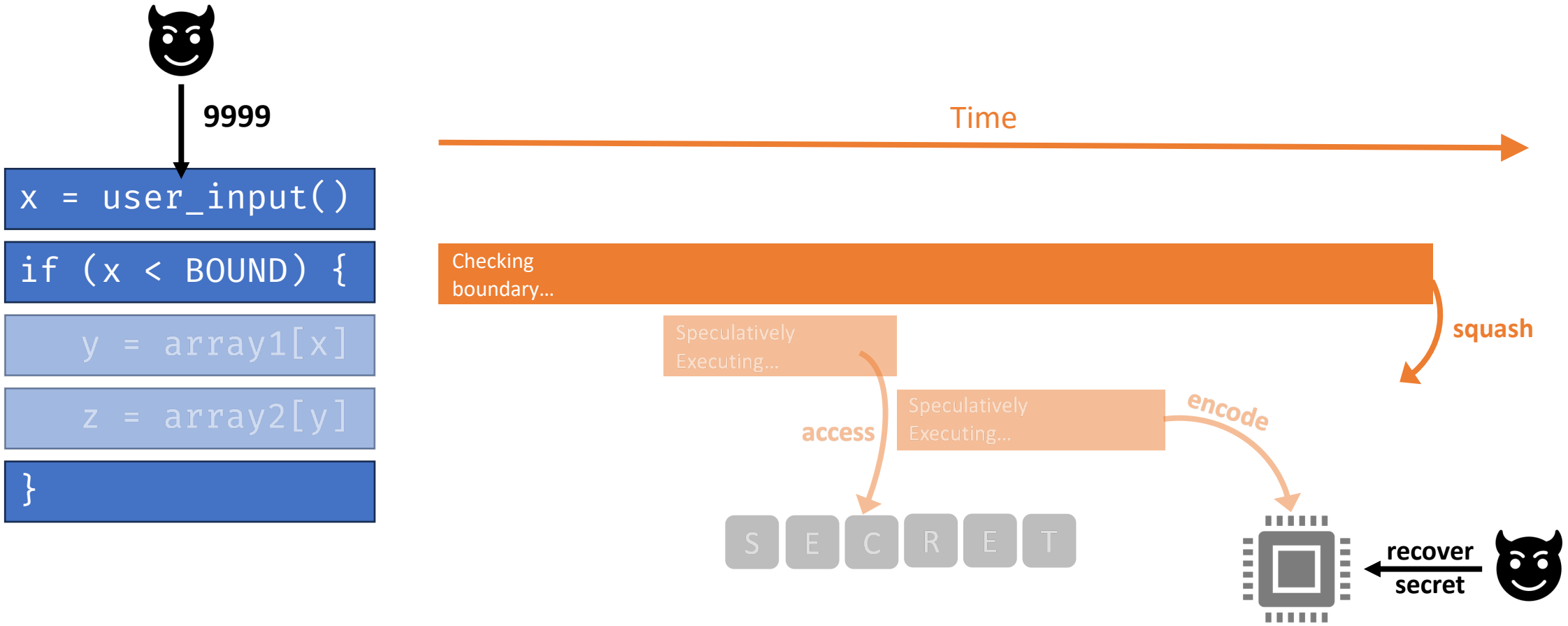


What is a Spectre gadget?





What is a Spectre gadget?



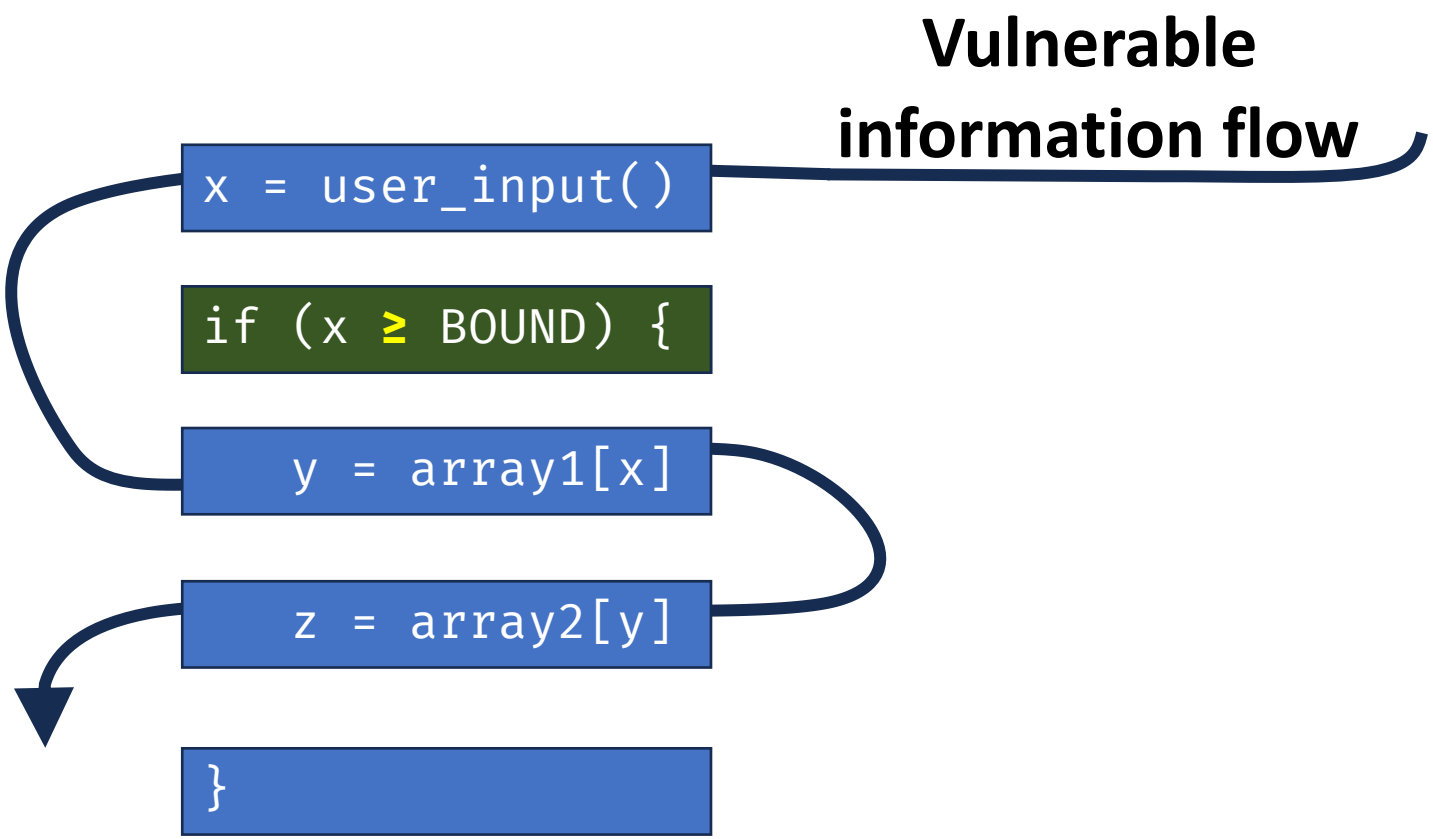
How do existing gadget scanners work?

```
x = user_input()

if (x < BOUND) {

    y = array1[x]

    z = array2[y]

}
```

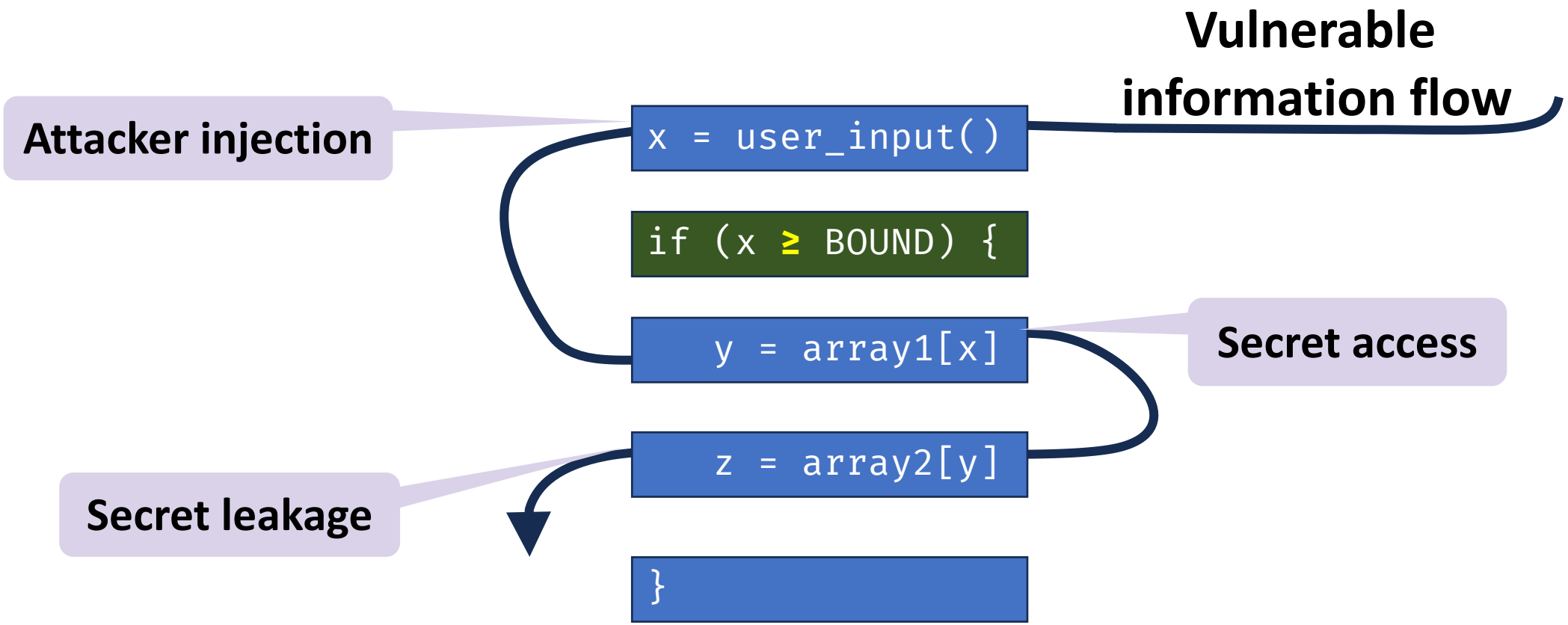
How do existing gadget scanners work?

```
x = user_input()

if (x ≥ BOUND) {

    y = array1[x]

    z = array2[y]

}
```

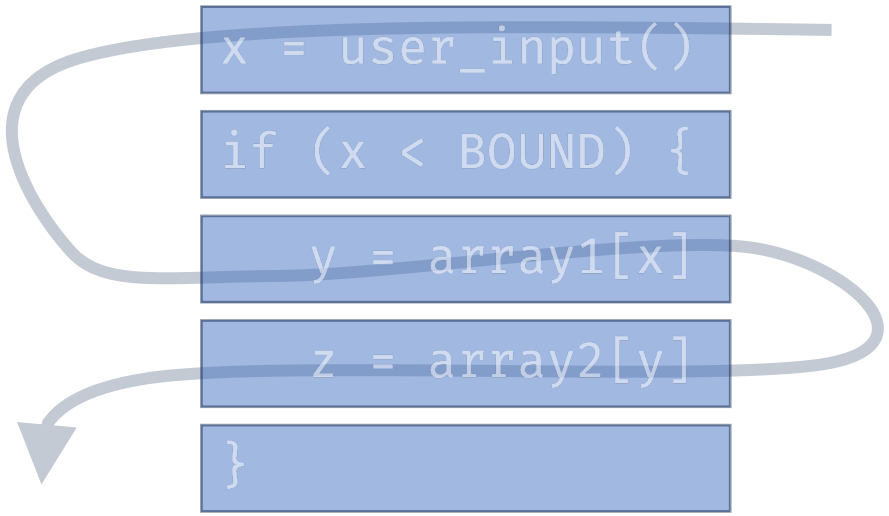
How do existing gadget scanners work?



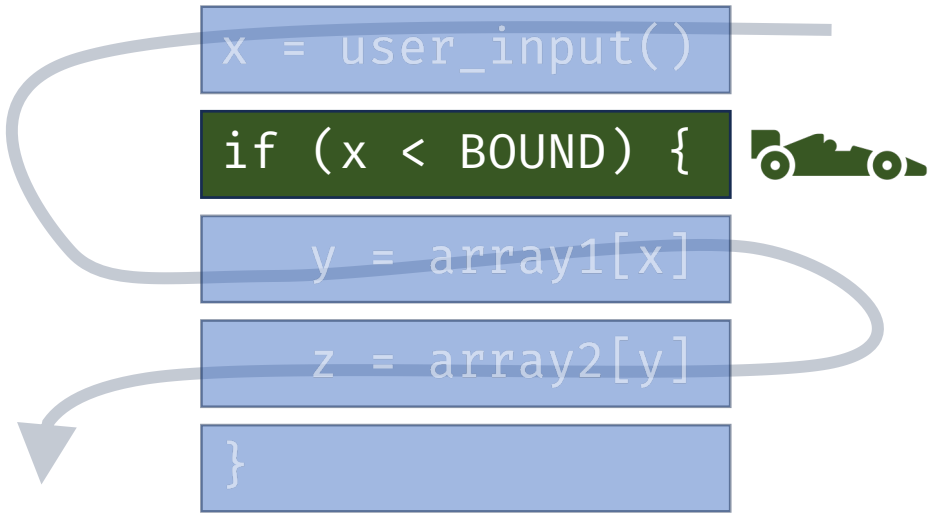
How do existing gadget scanners work?



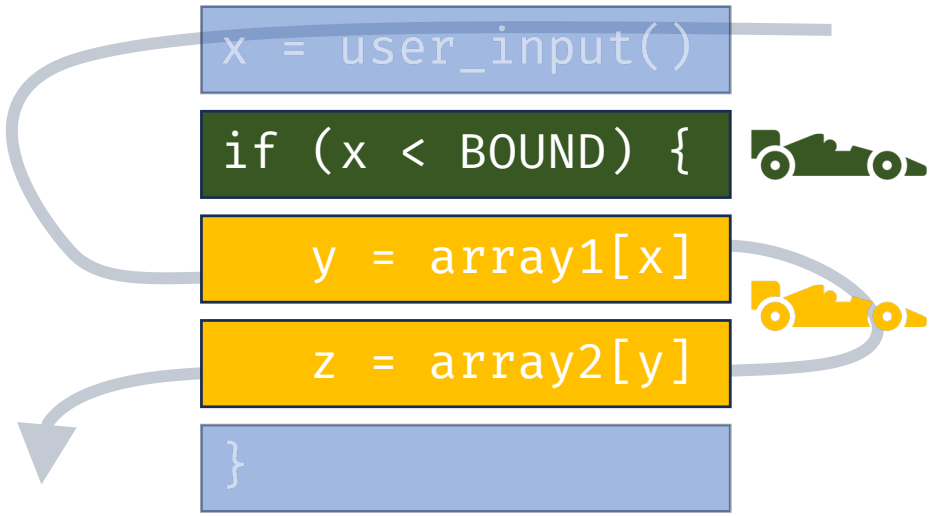
Missing piece: Timing condition



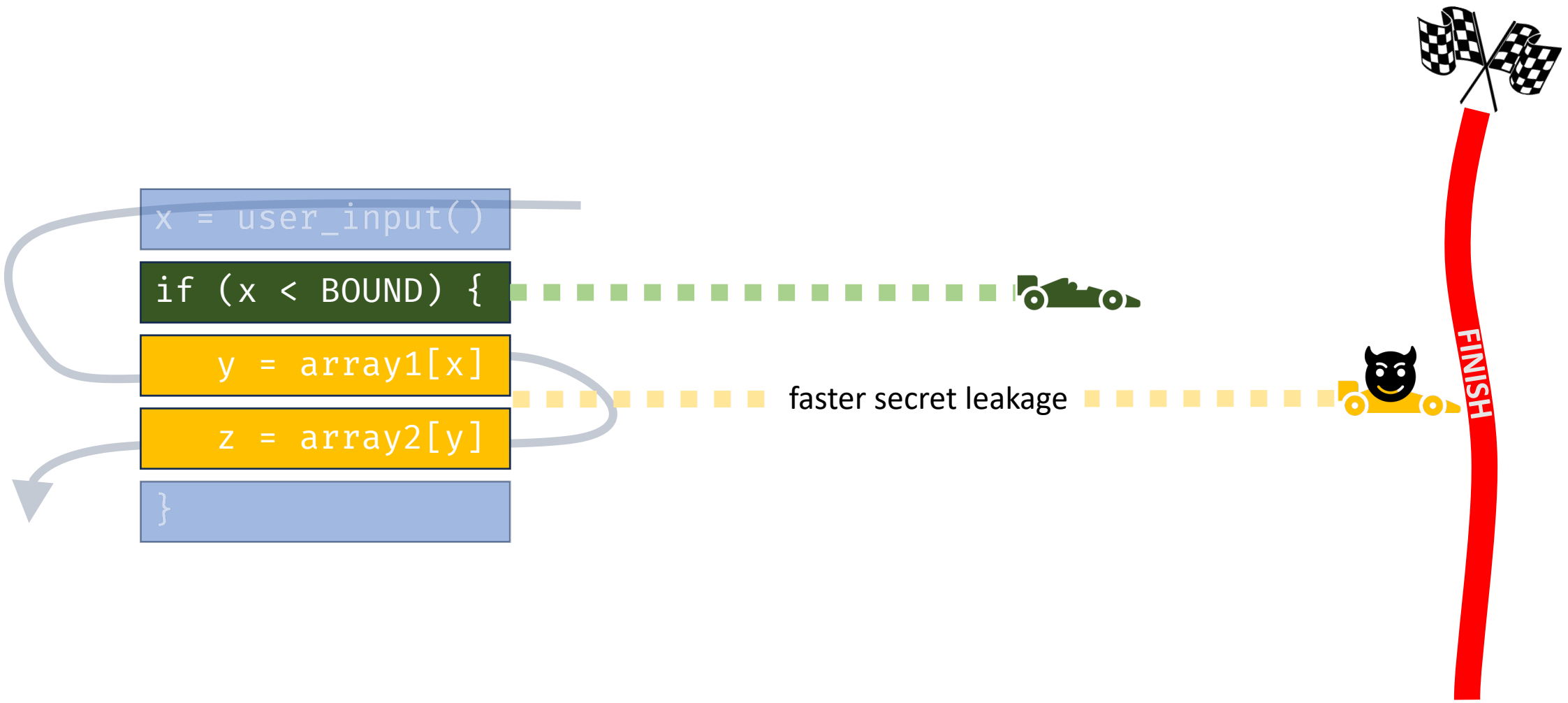
Missing piece: Timing condition



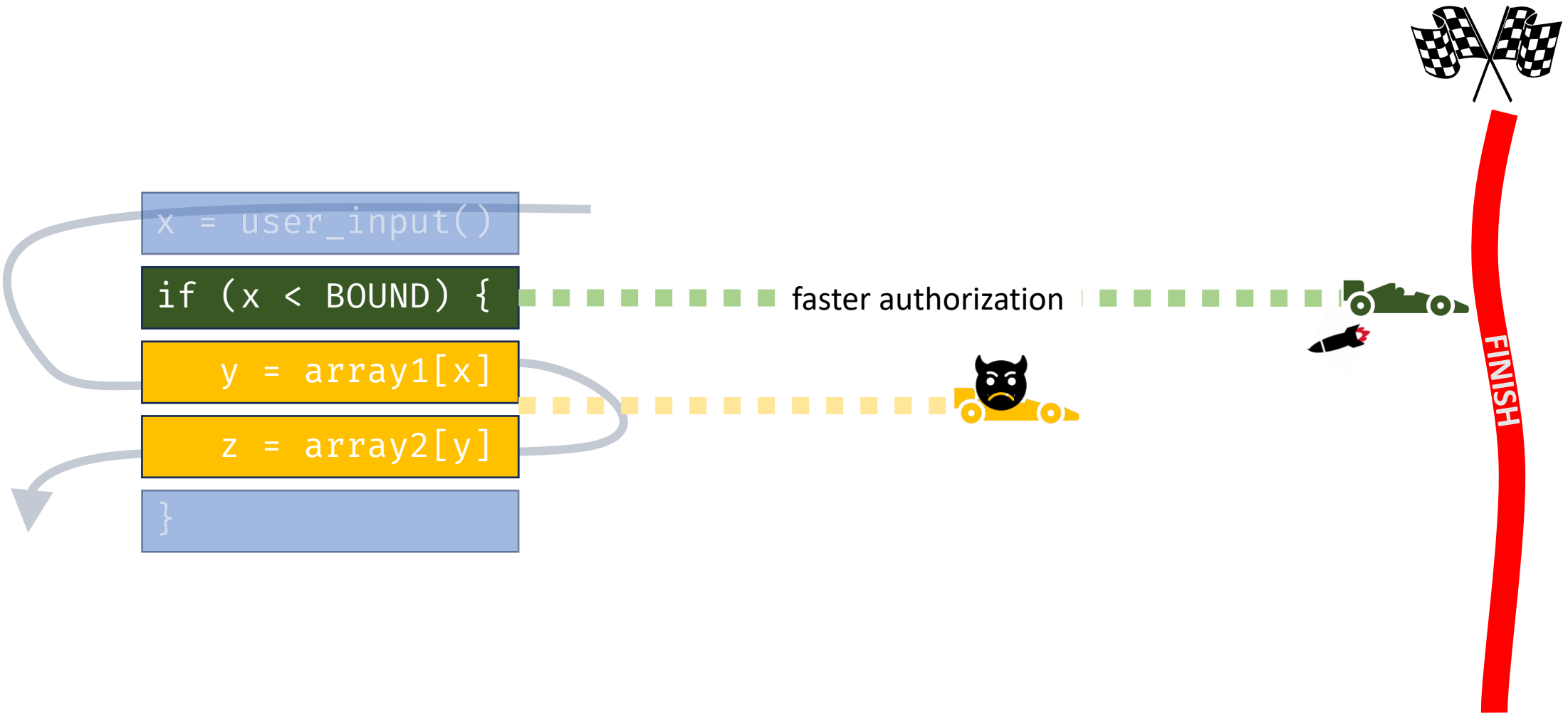
Missing piece: Timing condition



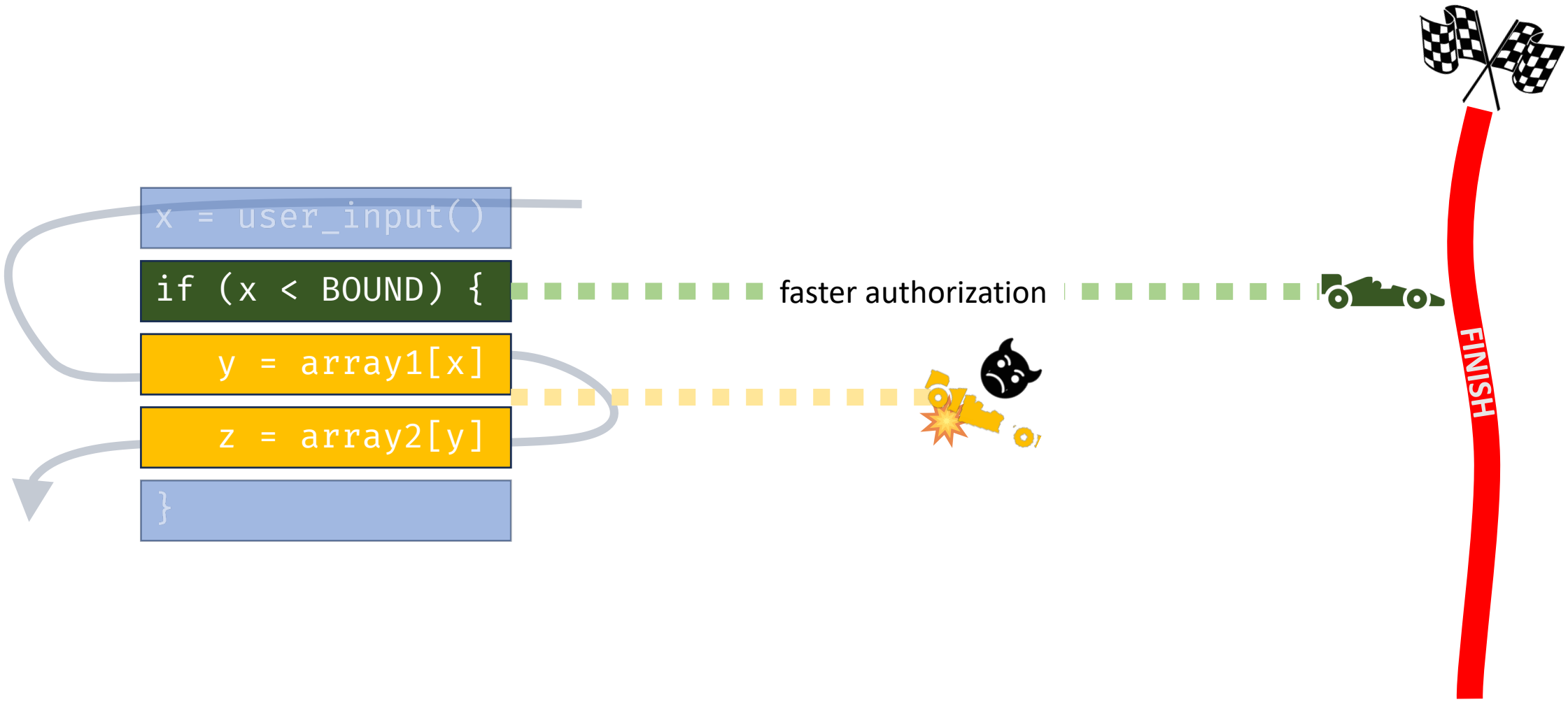
Missing piece: Timing condition



Missing piece: Timing condition



Missing piece: Timing condition



Satisfying the **timing condition** is necessary for a gadget to be **exploitable**.

Satisfying the **timing condition** is necessary for a gadget to be **exploitable**.

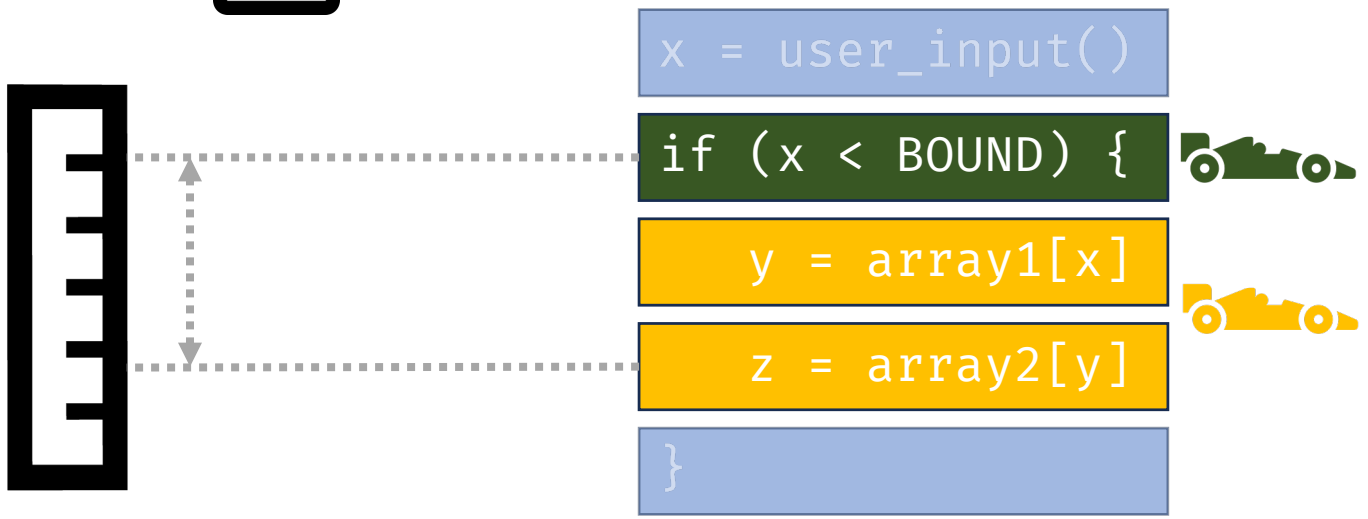
How do **existing works** model the timing condition?

Most scanners: Approximating with RoB size

```
x = user_input()
if (x < BOUND) {
    y = array1[x]
    z = array2[y]
}
```

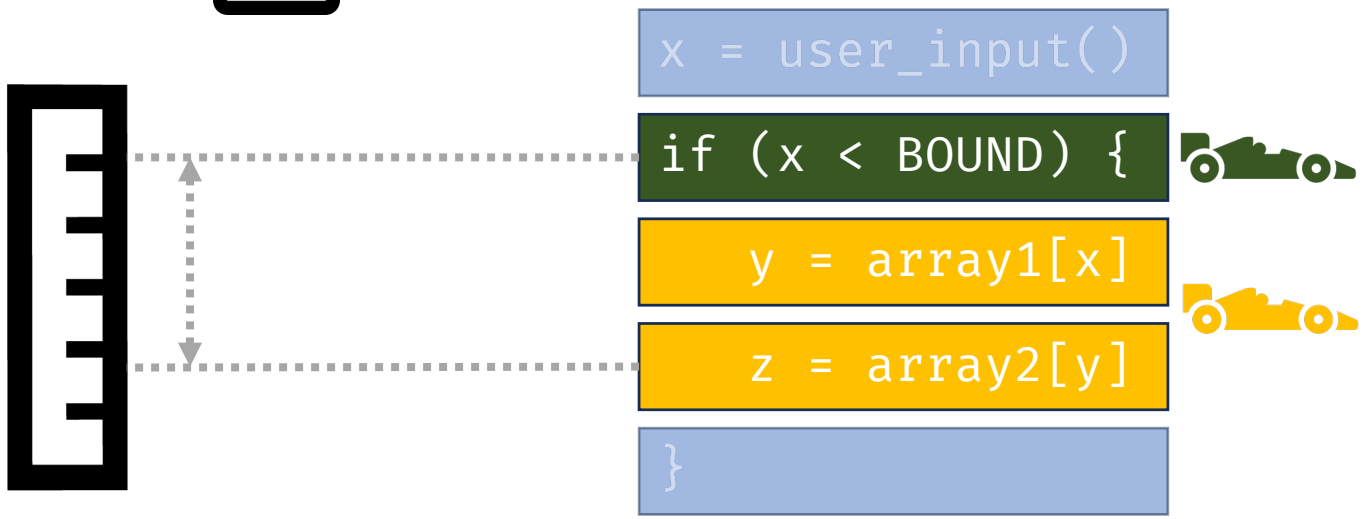
Most scanners: Approximating with RoB size

Fit in RoB ?

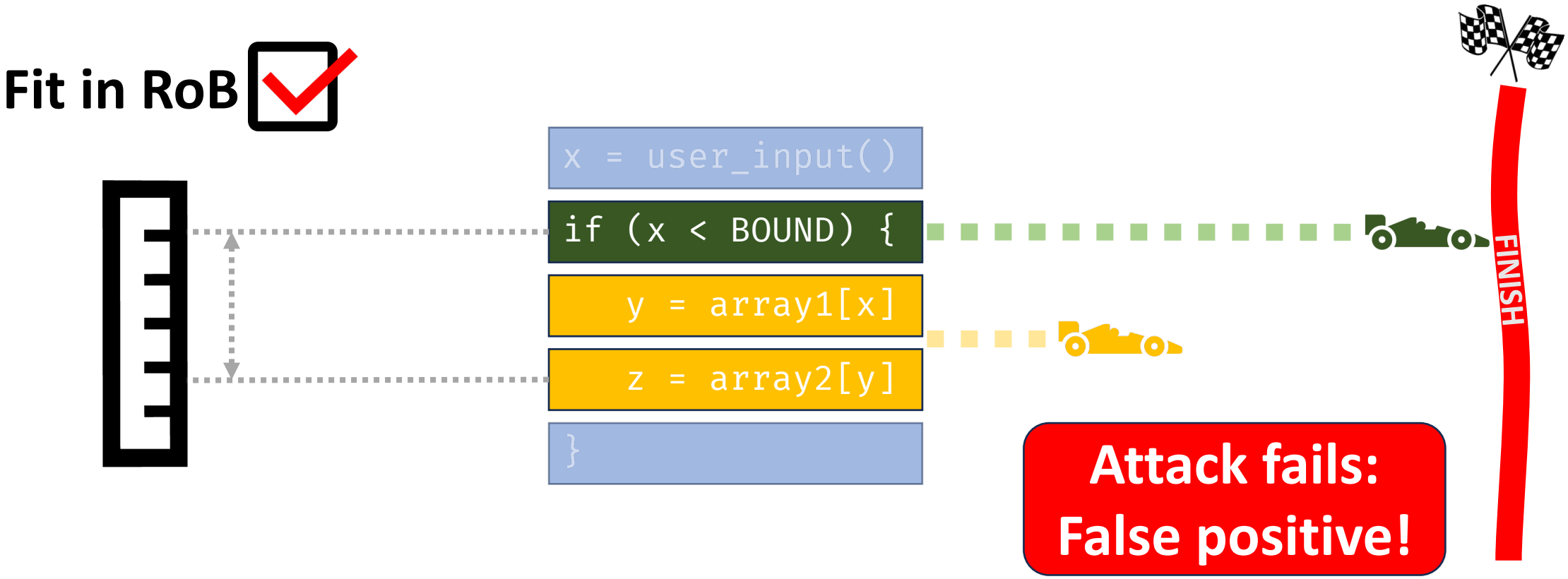


Most scanners: Approximating with RoB size

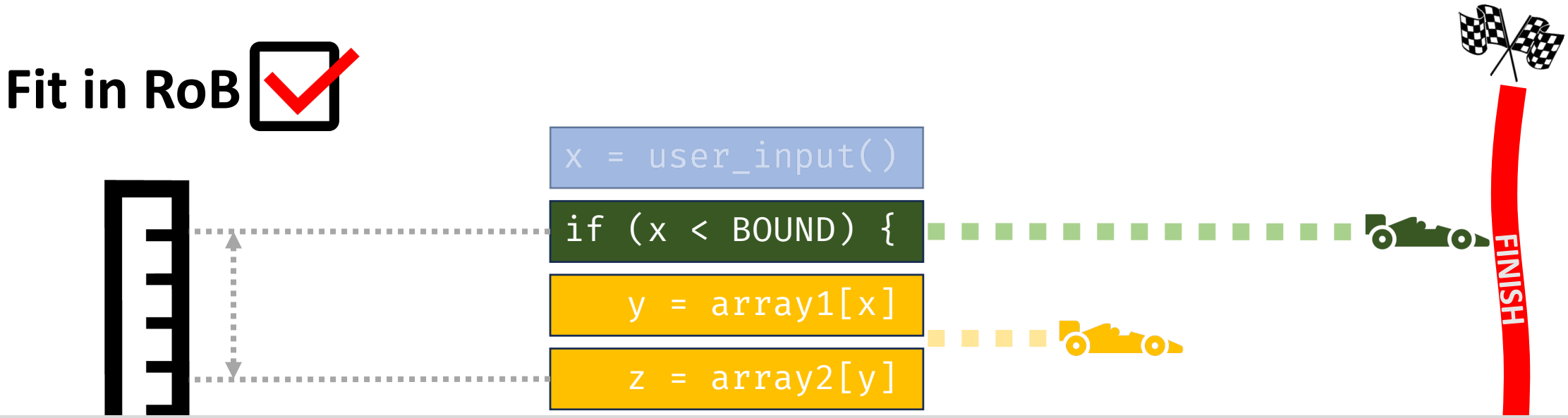
Fit in RoB ☑



Most scanners: Approximating with RoB size

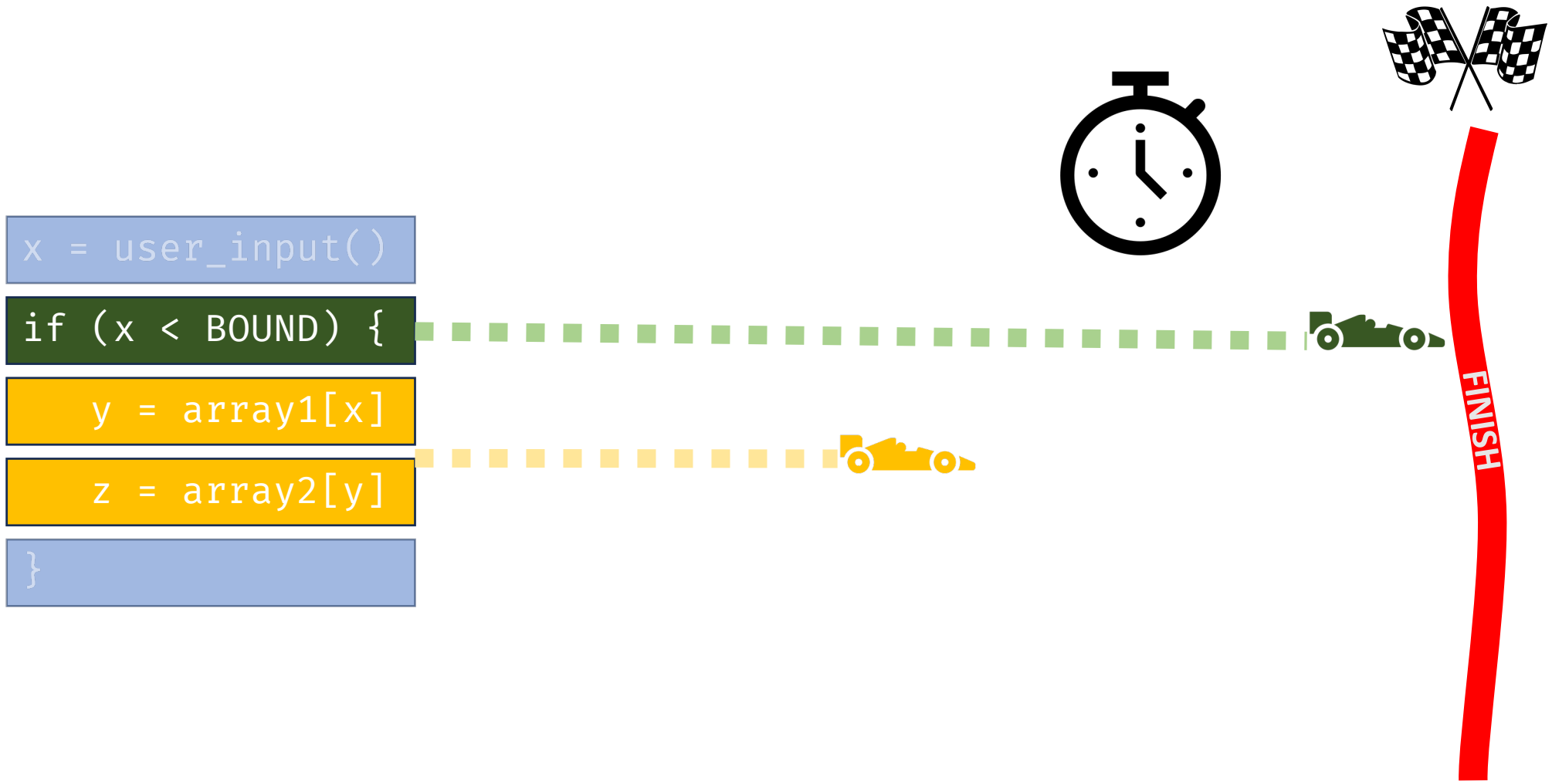


Most scanners: Approximating with RoB size

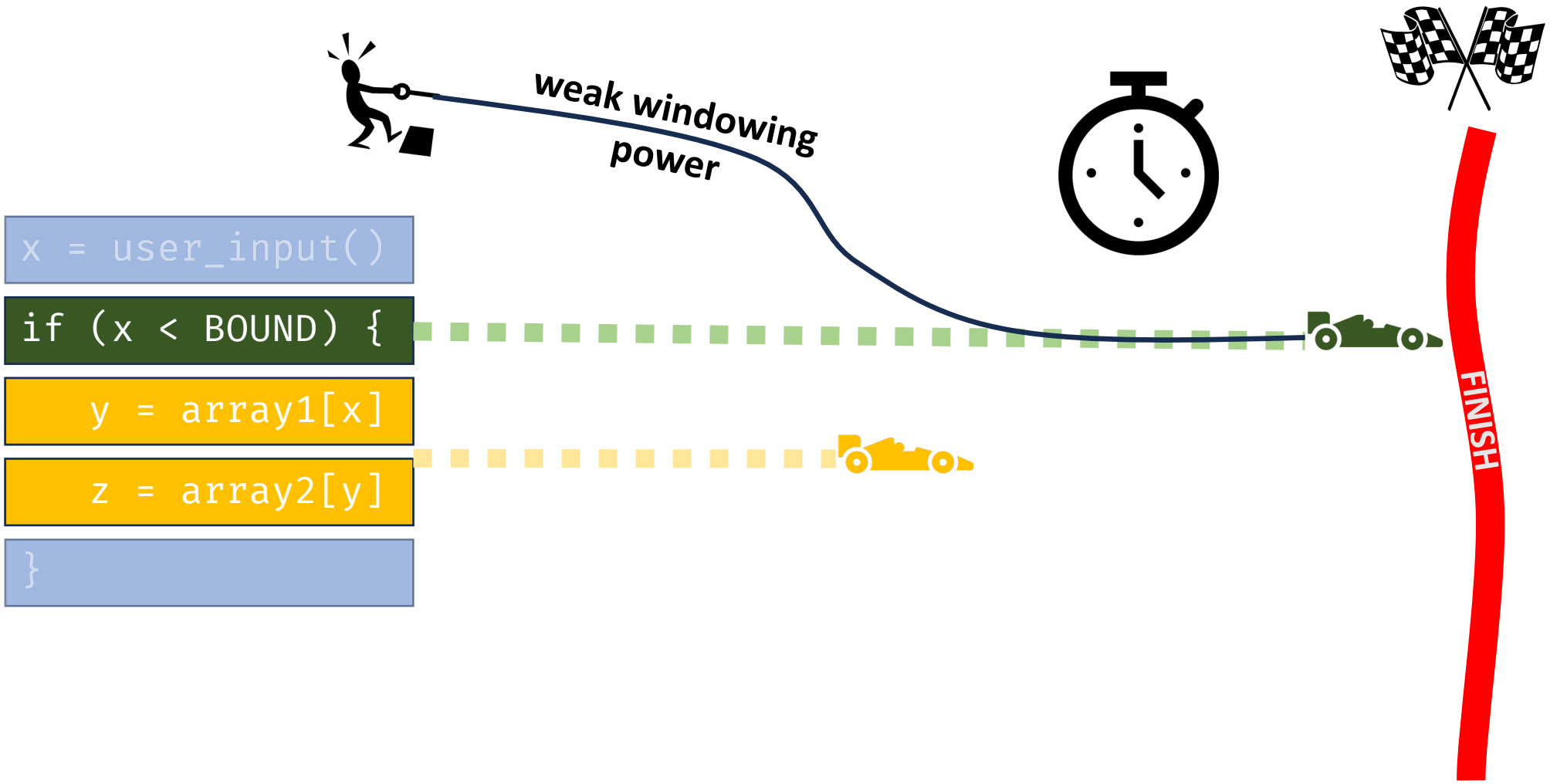


Measure the timing condition accurately!

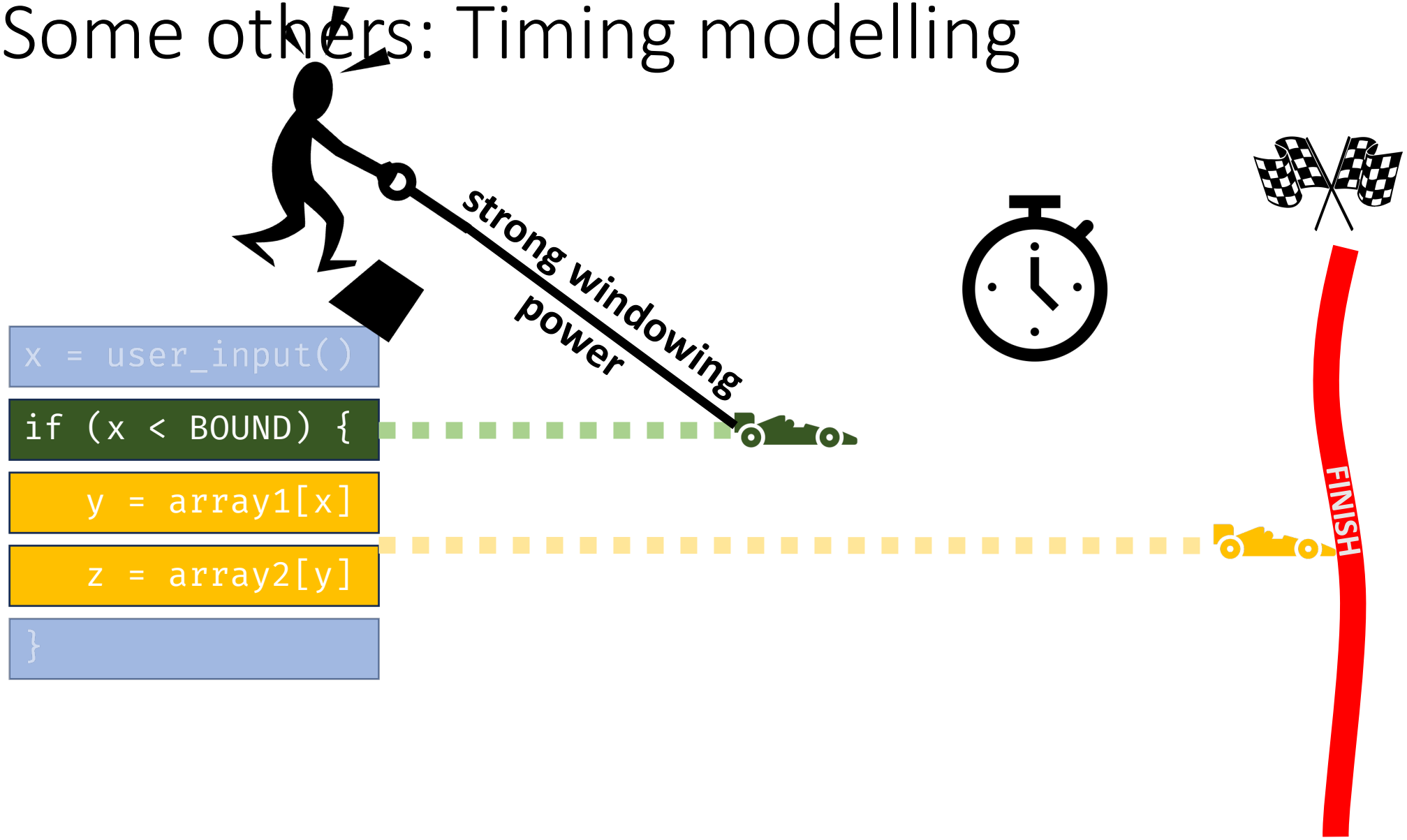
Some others: Timing modelling



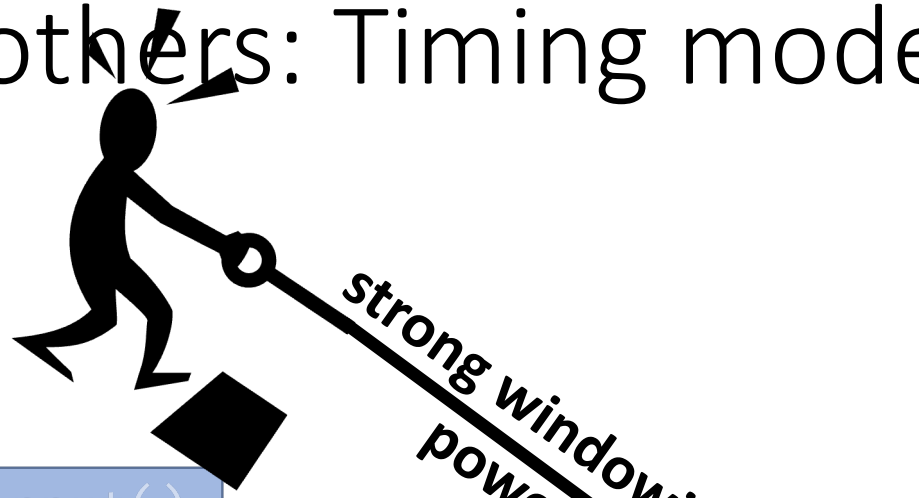
Some others: Timing modelling

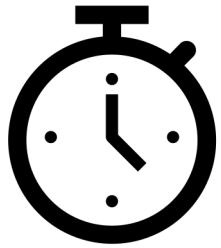


Some others: Timing modelling



Some others: Timing modelling

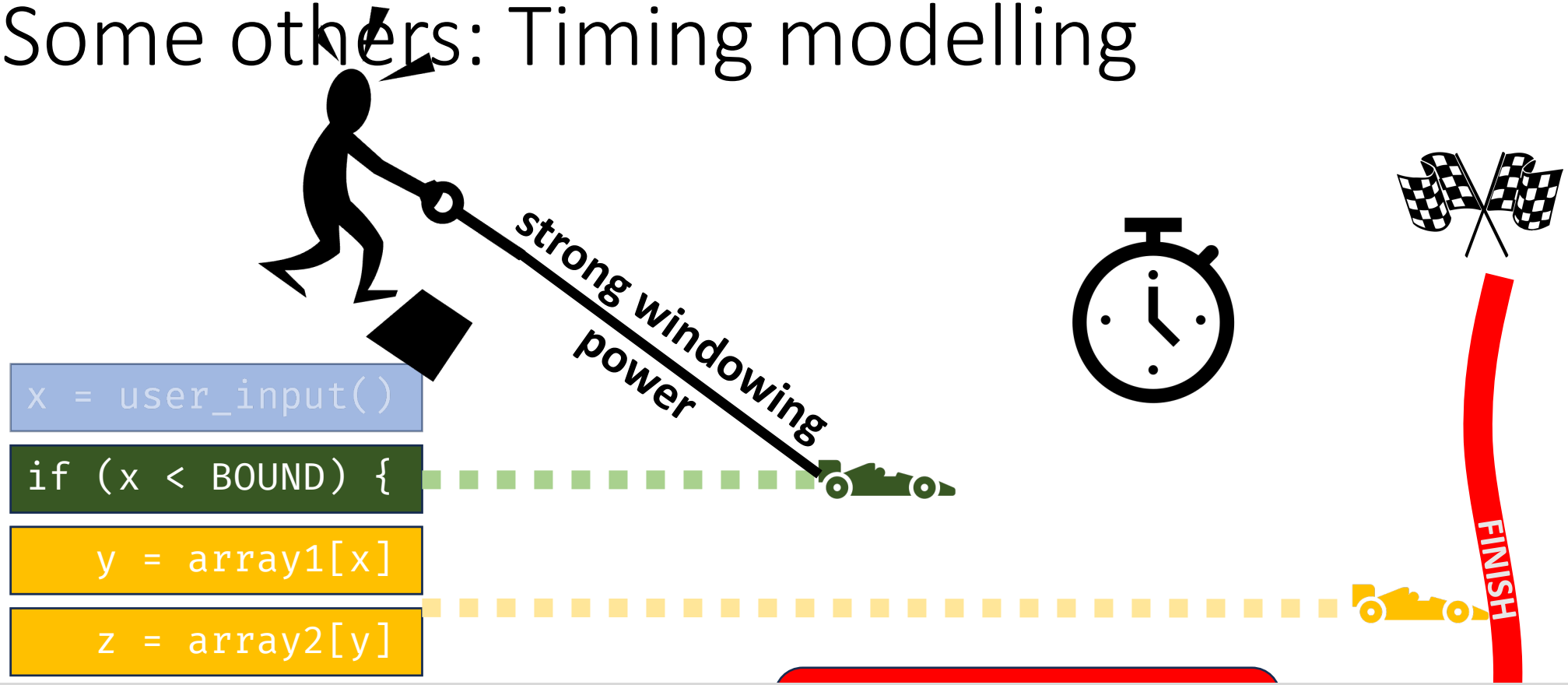


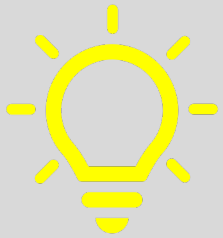



```
x = user_input()
if (x < BOUND) {
    y = array1[x]
    z = array2[y]
}
```

Attack succeeds:
False negative!

Some others: Timing modelling

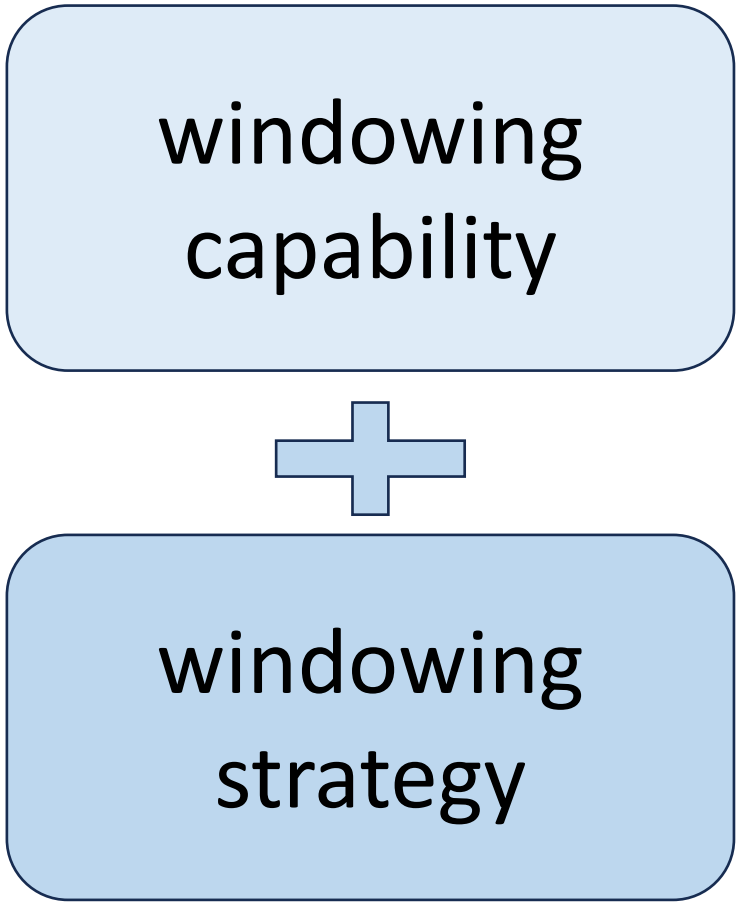




Measure the timing condition accurately,

under strong windowing power!

Our approach: Modelling windowing power

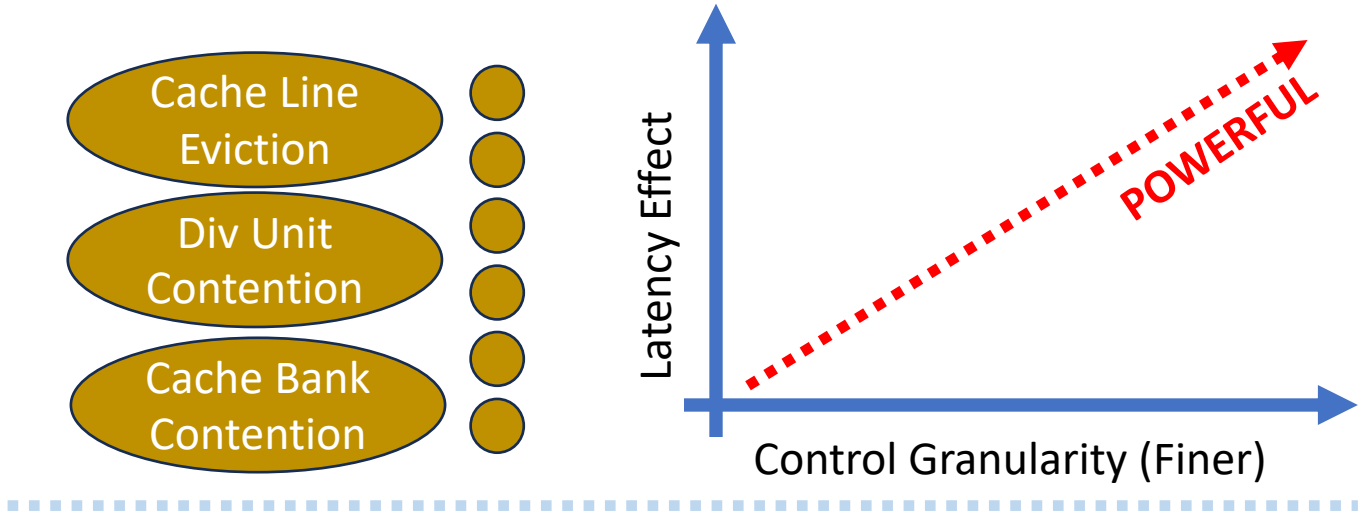


Our approach: Modelling windowing power

windowing capability

+

windowing strategy

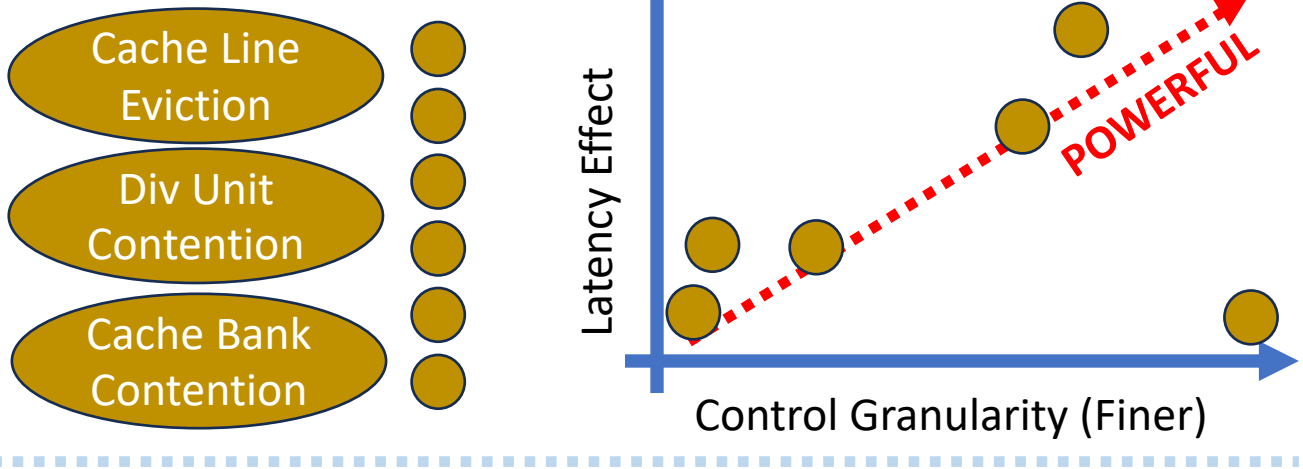


Our approach: Modelling windowing power

windowing capability

+

windowing strategy

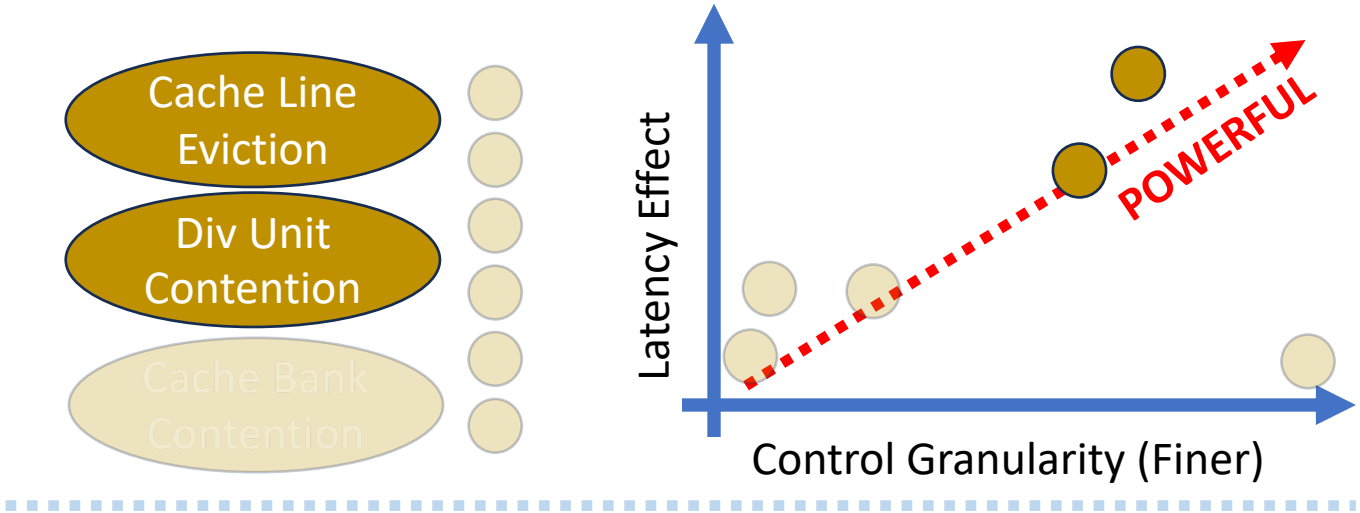


Our approach: Modelling windowing power

windowing capability

+

windowing strategy



Our approach: Modelling windowing power

windowing capability

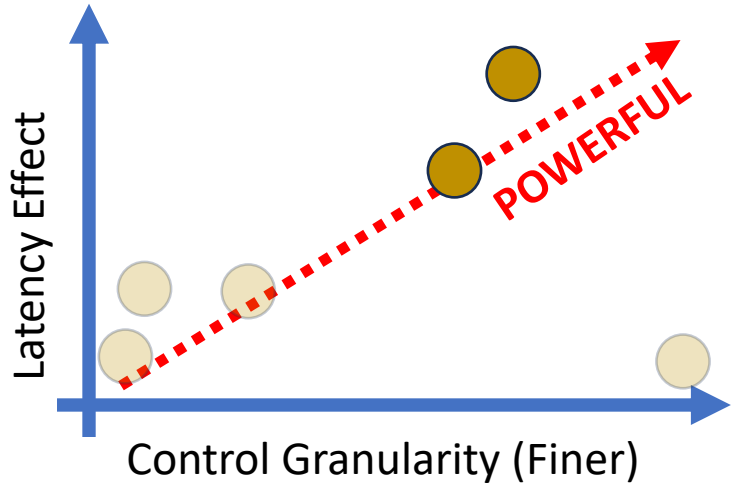
+

windowing strategy

Cache Line Eviction

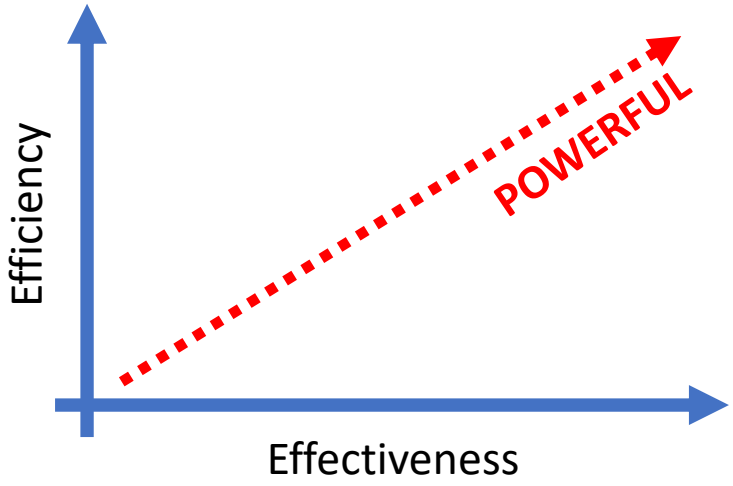
Div Unit Contention

Cache Bank Contention



Naïve Strategy

Brute force Strategy

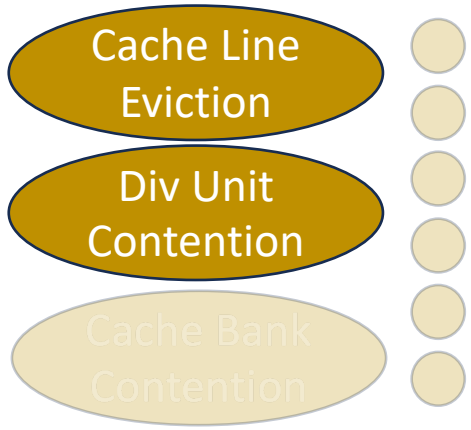


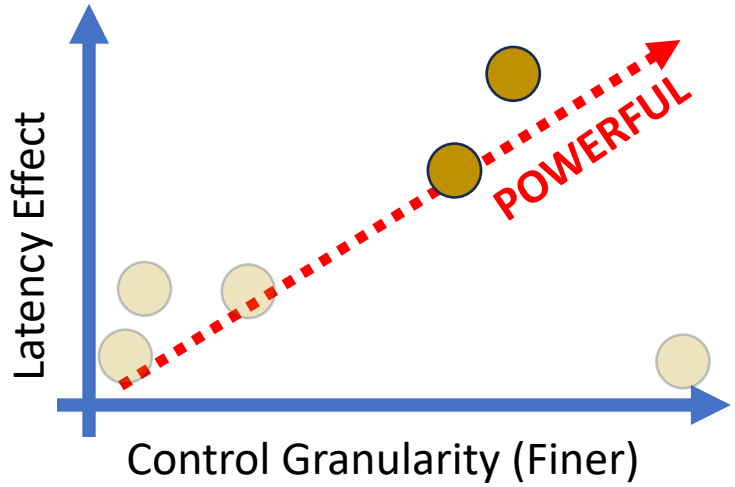
Our approach: Modelling windowing power

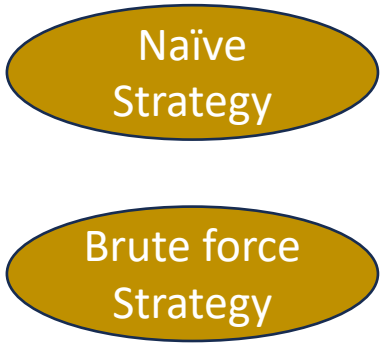
windowing capability

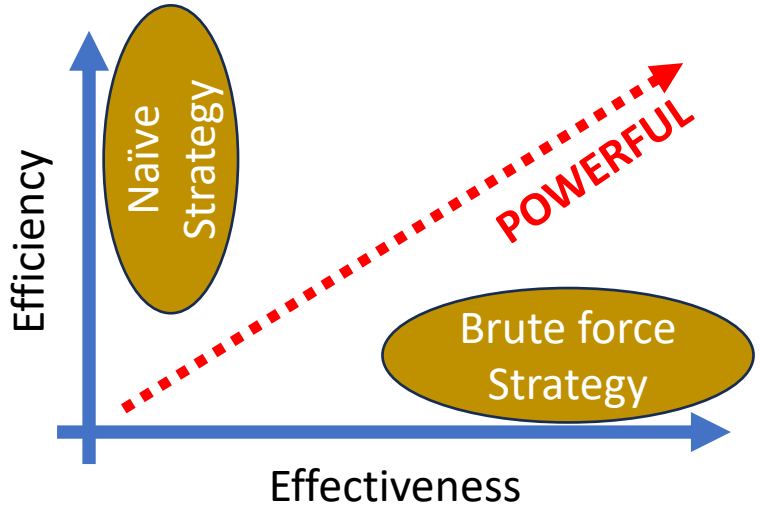
+

windowing strategy







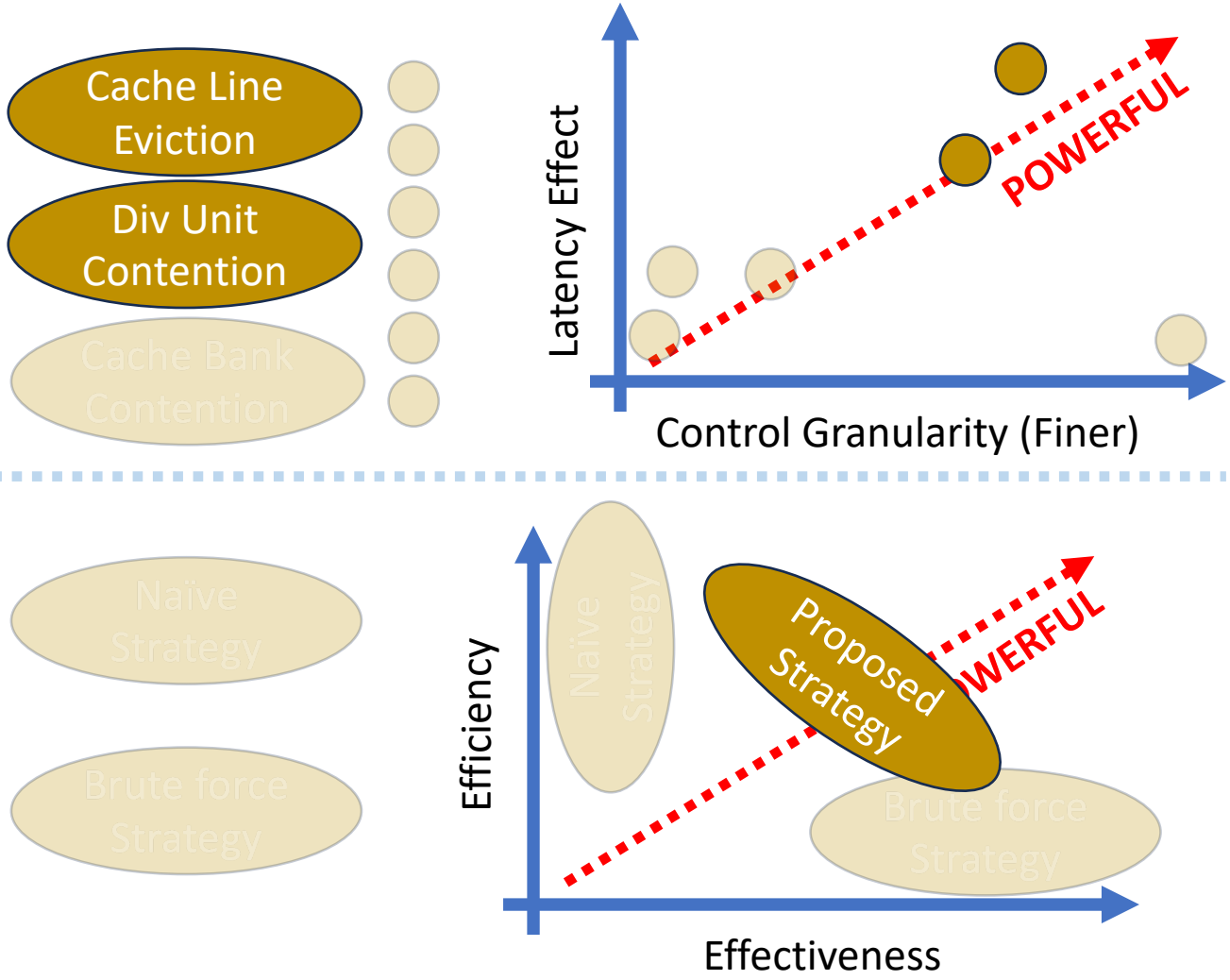


Our approach: Modelling windowing power

windowing capability

+

windowing strategy



Well, how do we evaluate gadgets?

Well, how do we evaluate gadgets?

Step A: **Modelling** timing condition

Well, how do we evaluate gadgets?

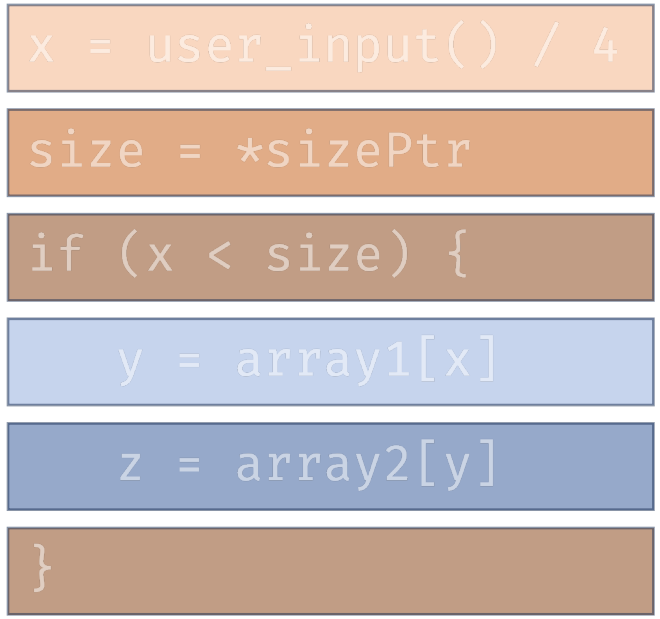
Step A: **Modelling** timing condition

Step B: **Simulating** windowing power

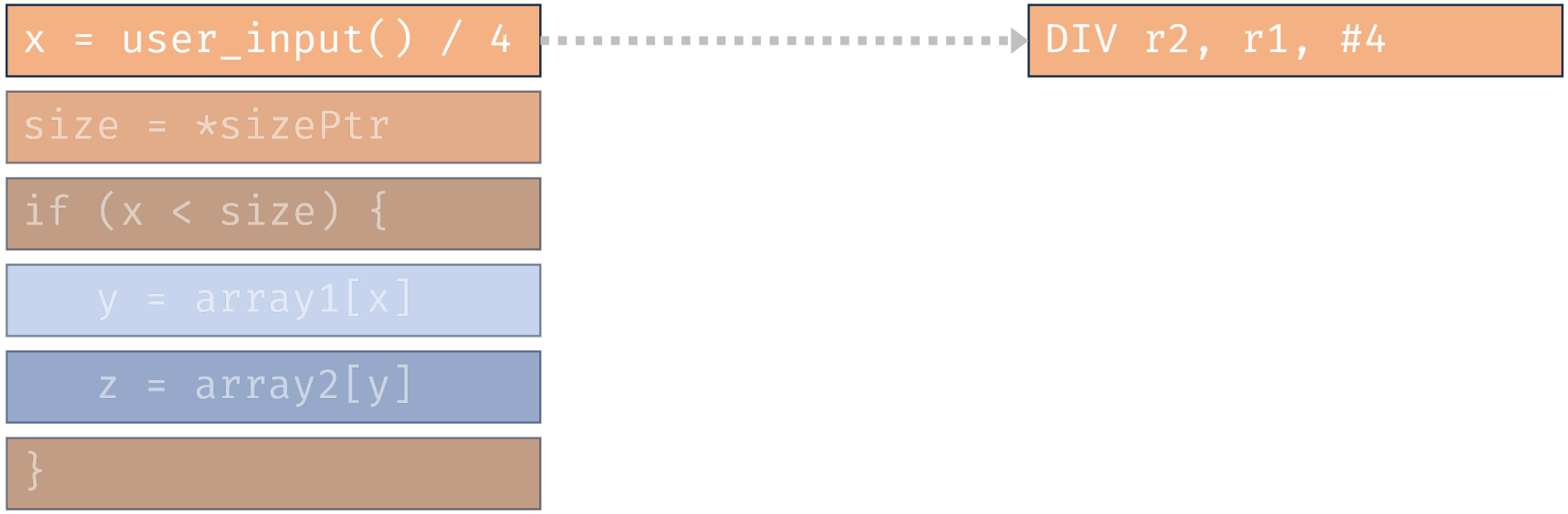
Well, how do we evaluate gadgets?

Step A: **Modelling** timing condition
Step B: **Simulating** windowing power
Step C: **Quantifying** exploitability

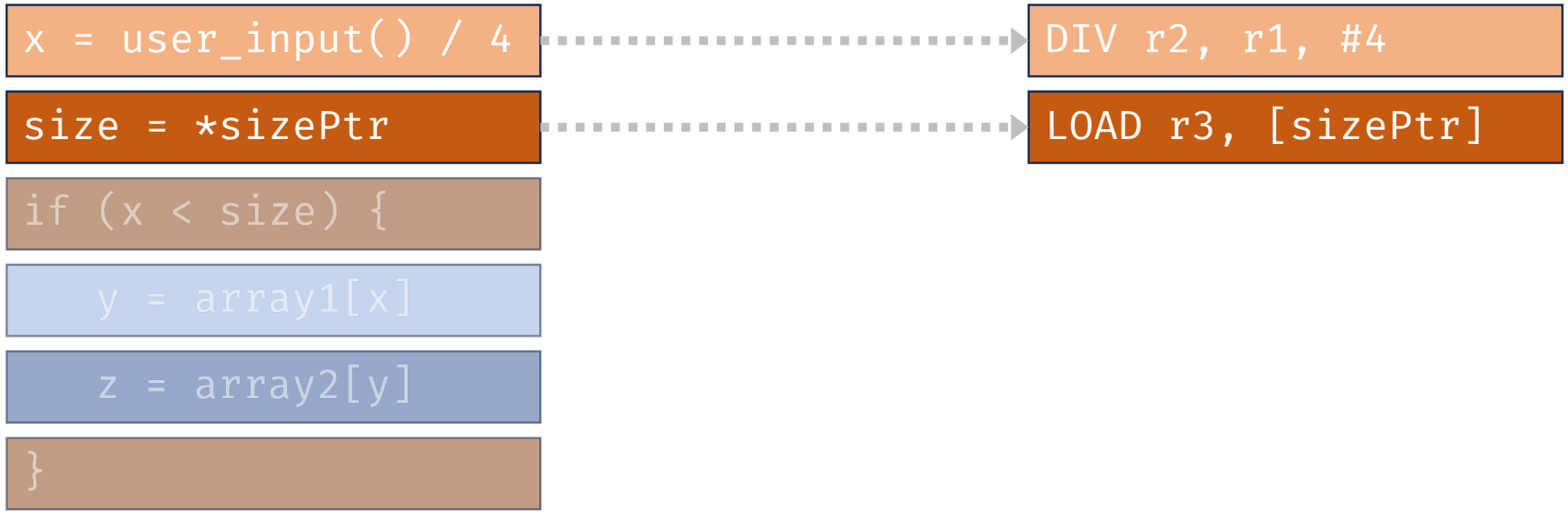
Our approach: An example



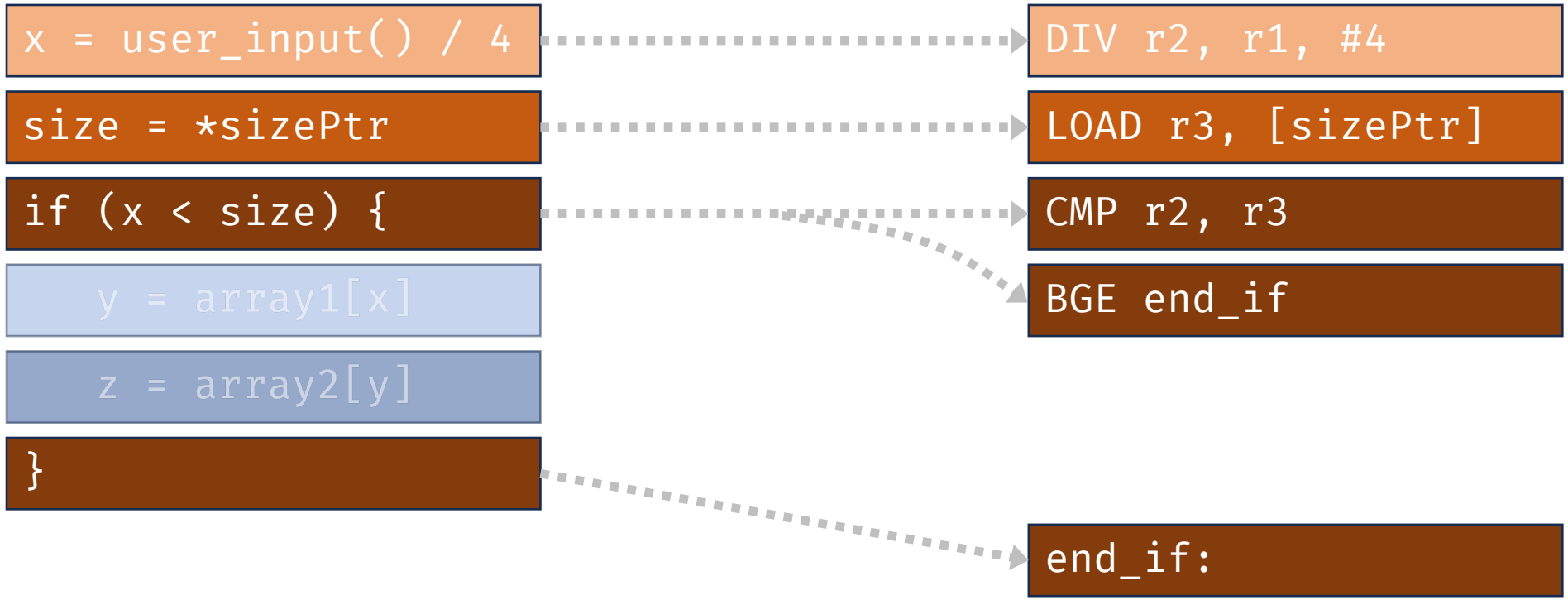
Our approach: An example



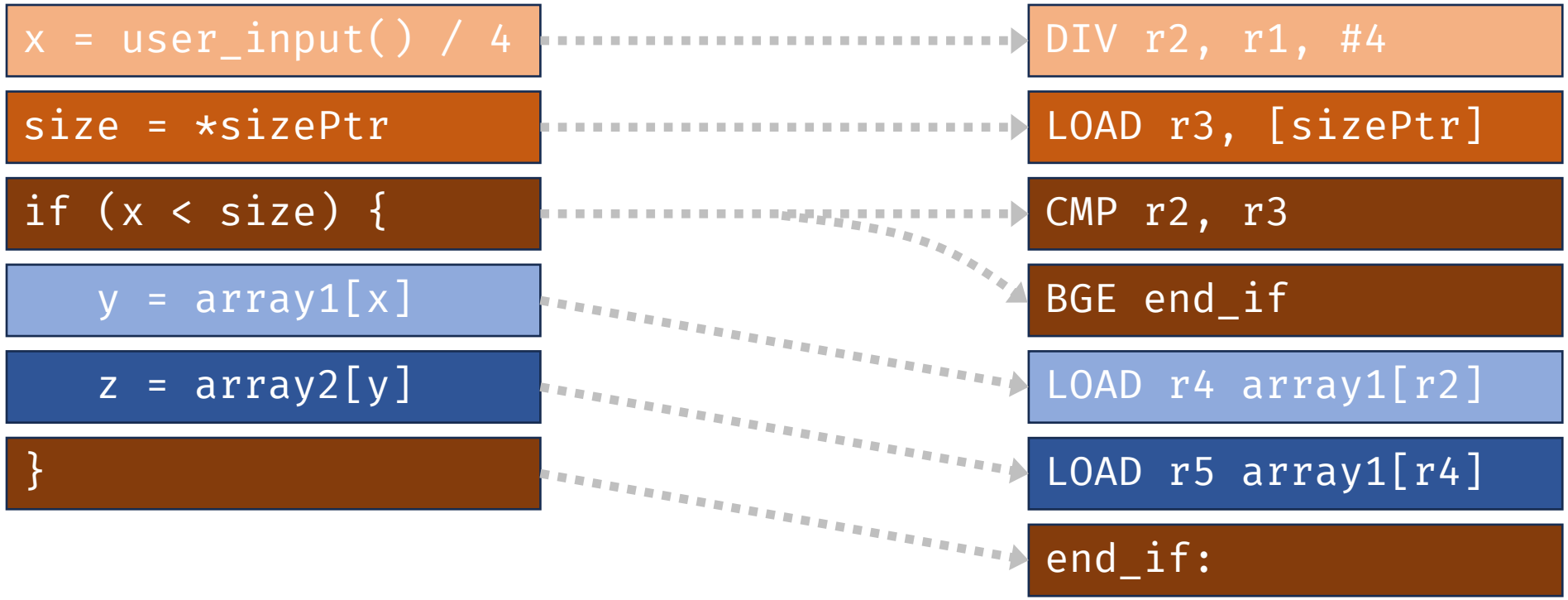
Our approach: An example



Our approach: An example

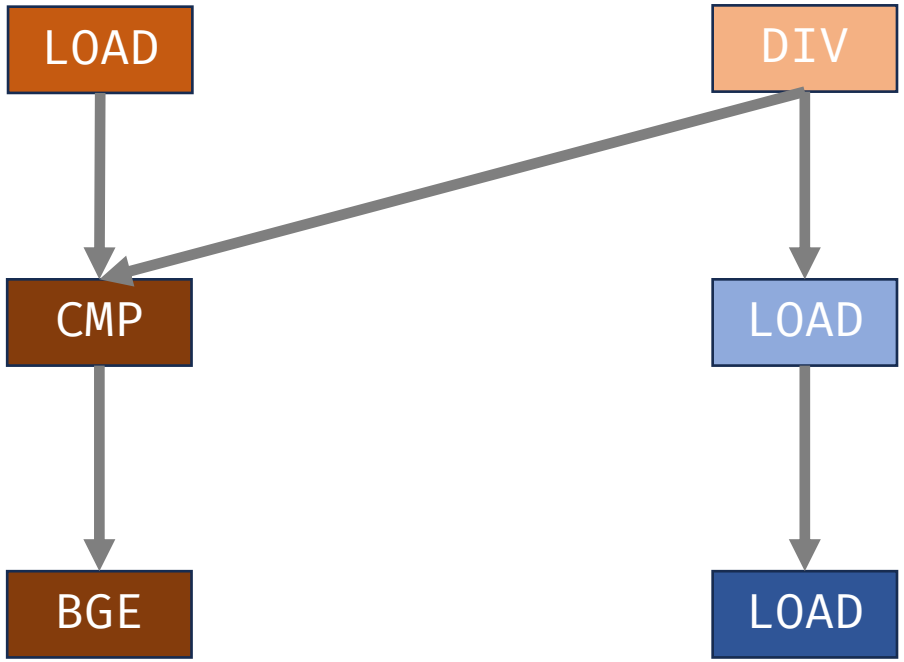


Our approach: An example



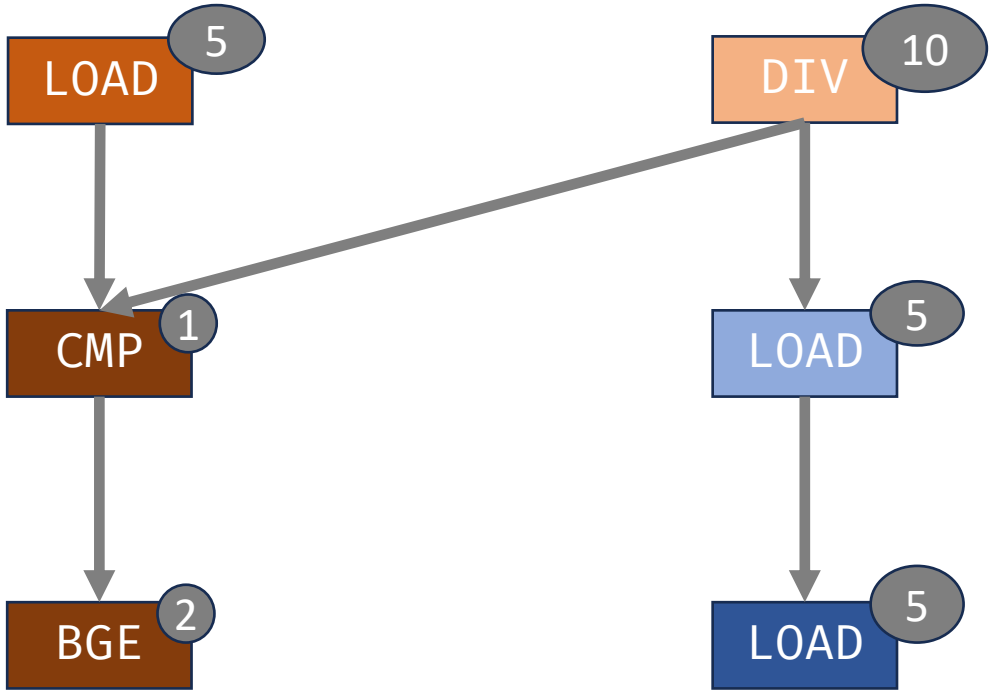
Step A: Modelling timing condition

```
DIV r2, r1, #4
LOAD r3, [sizePtr]
CMP r2, r3
BGE end_if
LOAD r4 array1[r2]
LOAD r5 array1[r4]
end_if:
```



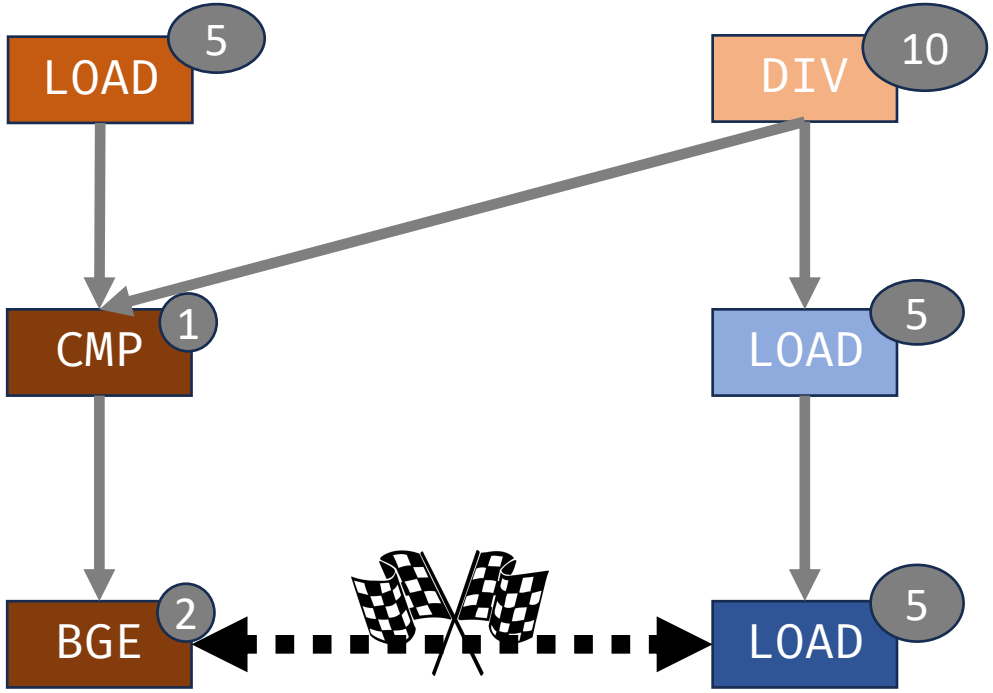
Step A: Modelling timing condition

```
DIV r2, r1, #4
LOAD r3, [sizePtr]
CMP r2, r3
BGE end_if
LOAD r4 array1[r2]
LOAD r5 array1[r4]
end_if:
```



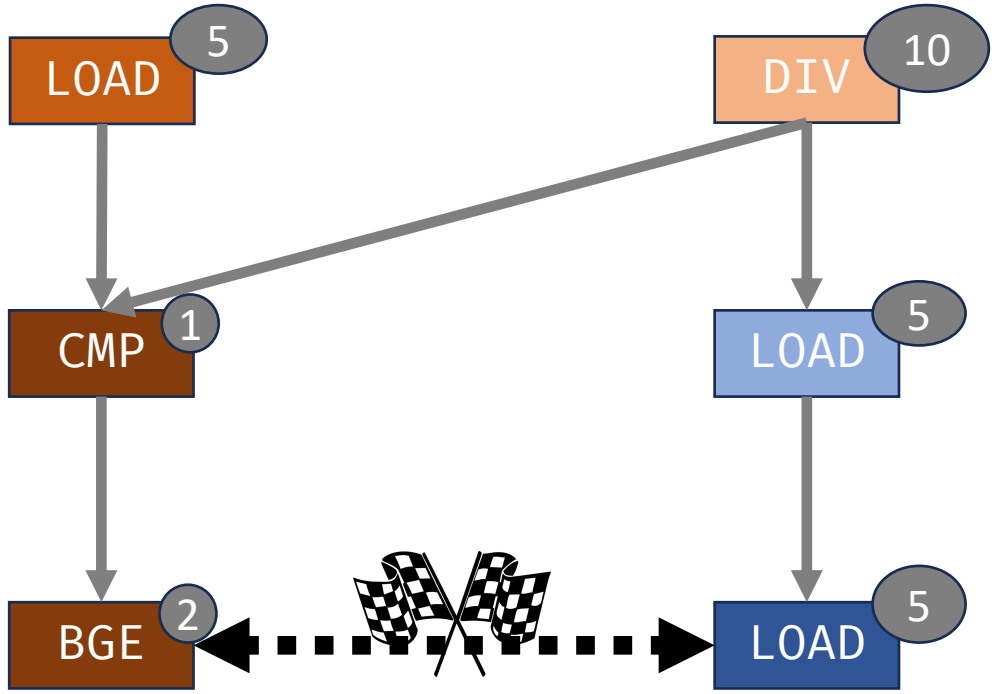
Step A: Modelling timing condition

```
DIV r2, r1, #4
LOAD r3, [sizePtr]
CMP r2, r3
BGE end_if
LOAD r4 array1[r2]
LOAD r5 array1[r4]
end_if:
```



Step A: Modelling timing condition

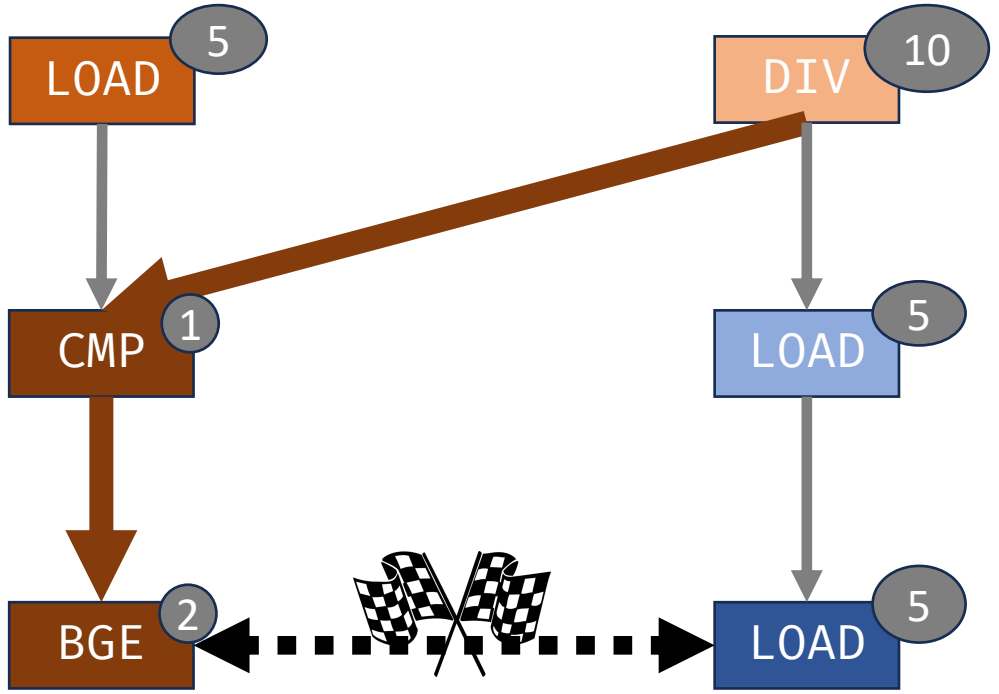
```
DIV r2, r1, #4
LOAD r3, [sizePtr]
CMP r2, r3
BGE end_if
LOAD r4 array1[r2]
LOAD r5 array1[r4]
end_if:
```



MaxPathWeight(BGE) — ***MaxPathWeight***(LOAD)

Step A: Modelling timing condition

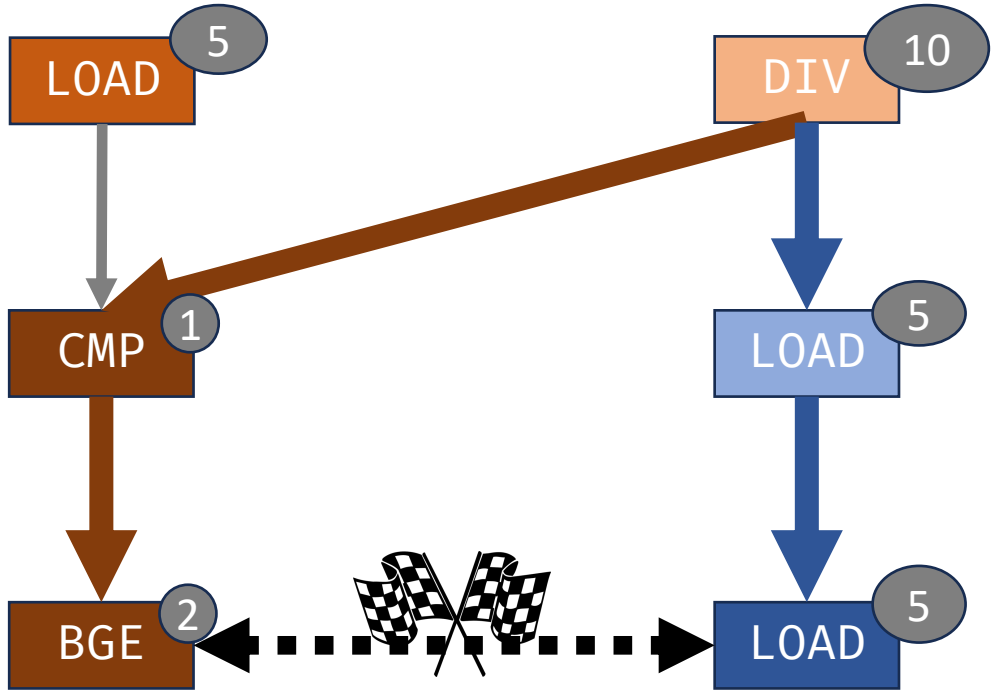
```
DIV r2, r1, #4
LOAD r3, [sizePtr]
CMP r2, r3
BGE end_if
LOAD r4 array1[r2]
LOAD r5 array1[r4]
end_if:
```



MaxPathWeight(BGE) — ***MaxPathWeight***(LOAD)

Step A: Modelling timing condition

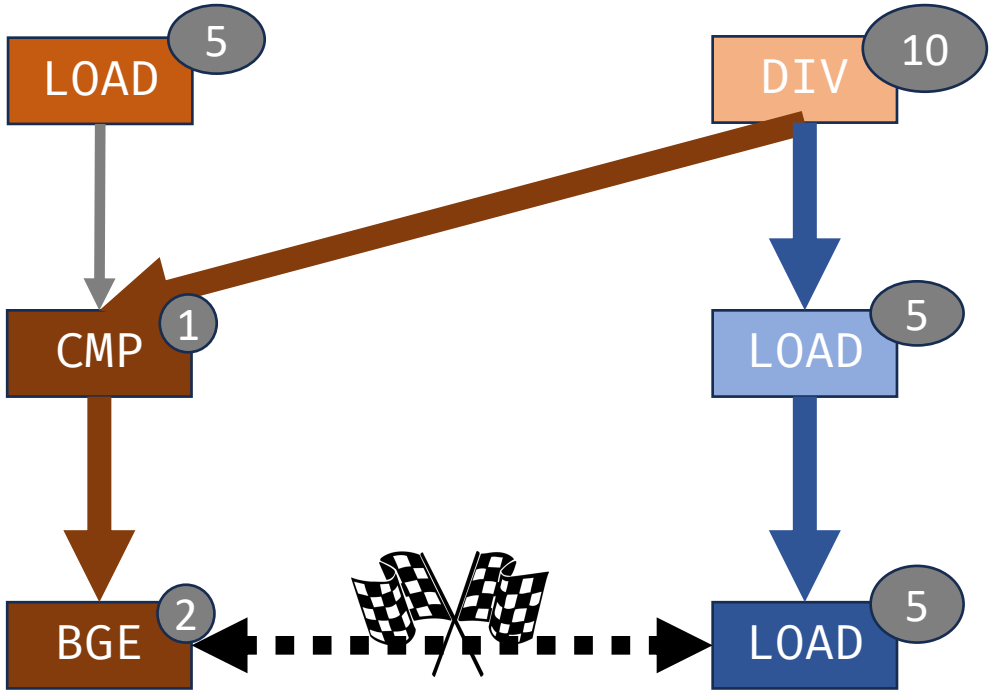
```
DIV r2, r1, #4
LOAD r3, [sizePtr]
CMP r2, r3
BGE end_if
LOAD r4 array1[r2]
LOAD r5 array1[r4]
end_if:
```



MaxPathWeight(BGE) — ***MaxPathWeight***(LOAD)

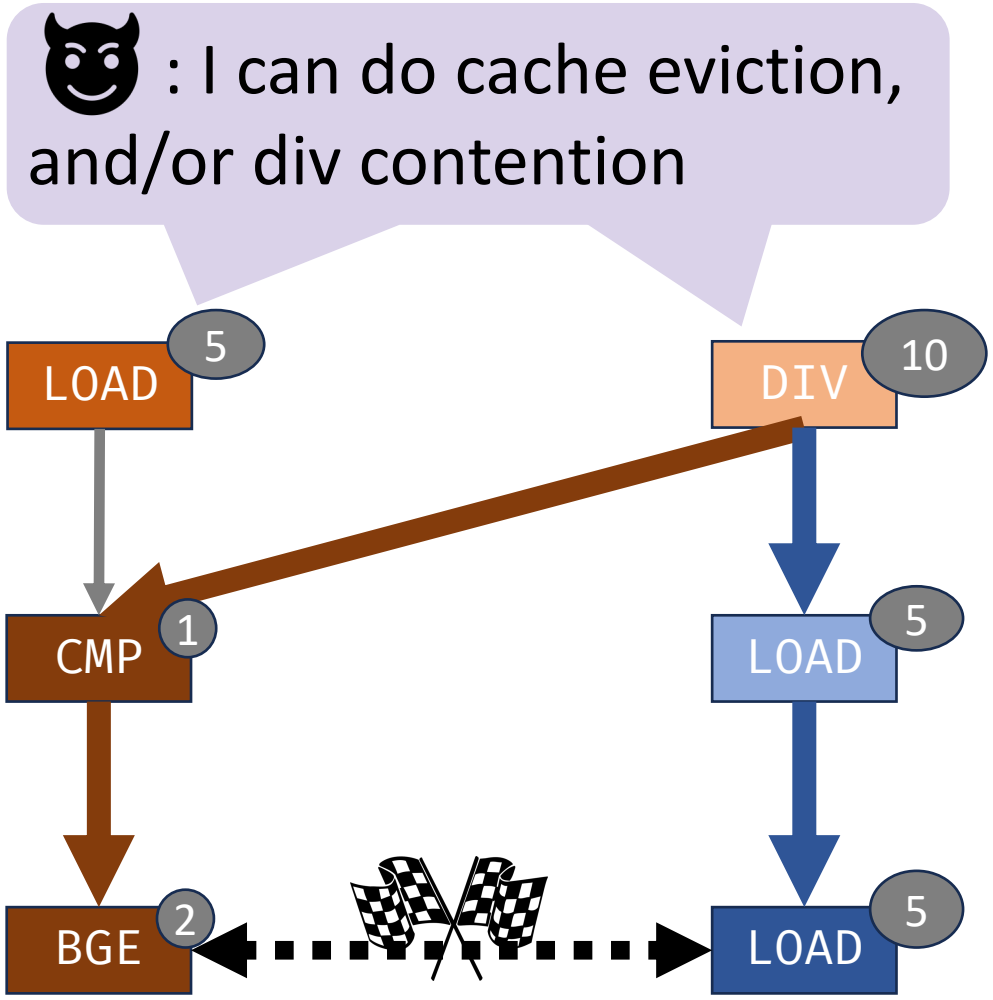
Step A: Modelling timing condition

```
DIV r2, r1, #4
LOAD r3, [sizePtr]
CMP r2, r3
BGE end_if
LOAD r4 array1[r2]
LOAD r5 array1[r4]
end_if:
```

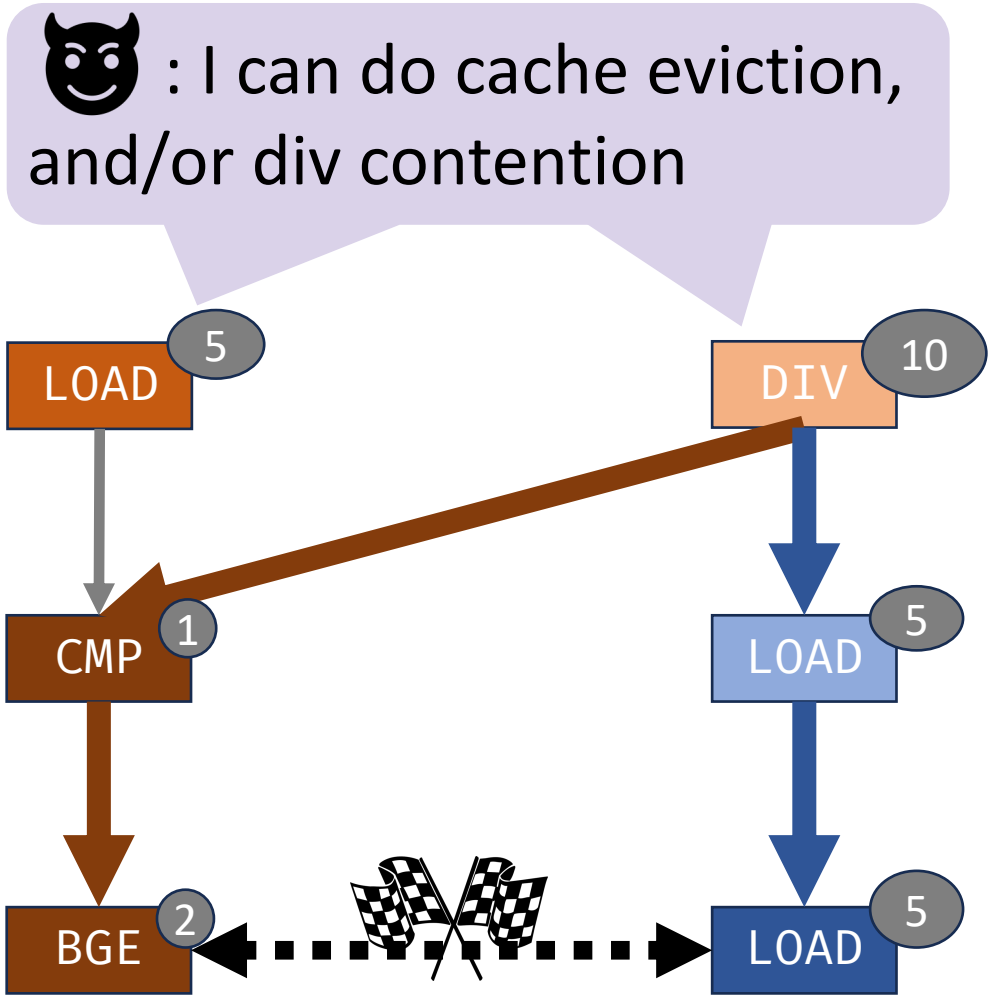


Timing Condition Index = ***MaxPathWeight***(BGE) − ***MaxPathWeight***(LOAD)

Step B: Simulating windowing power

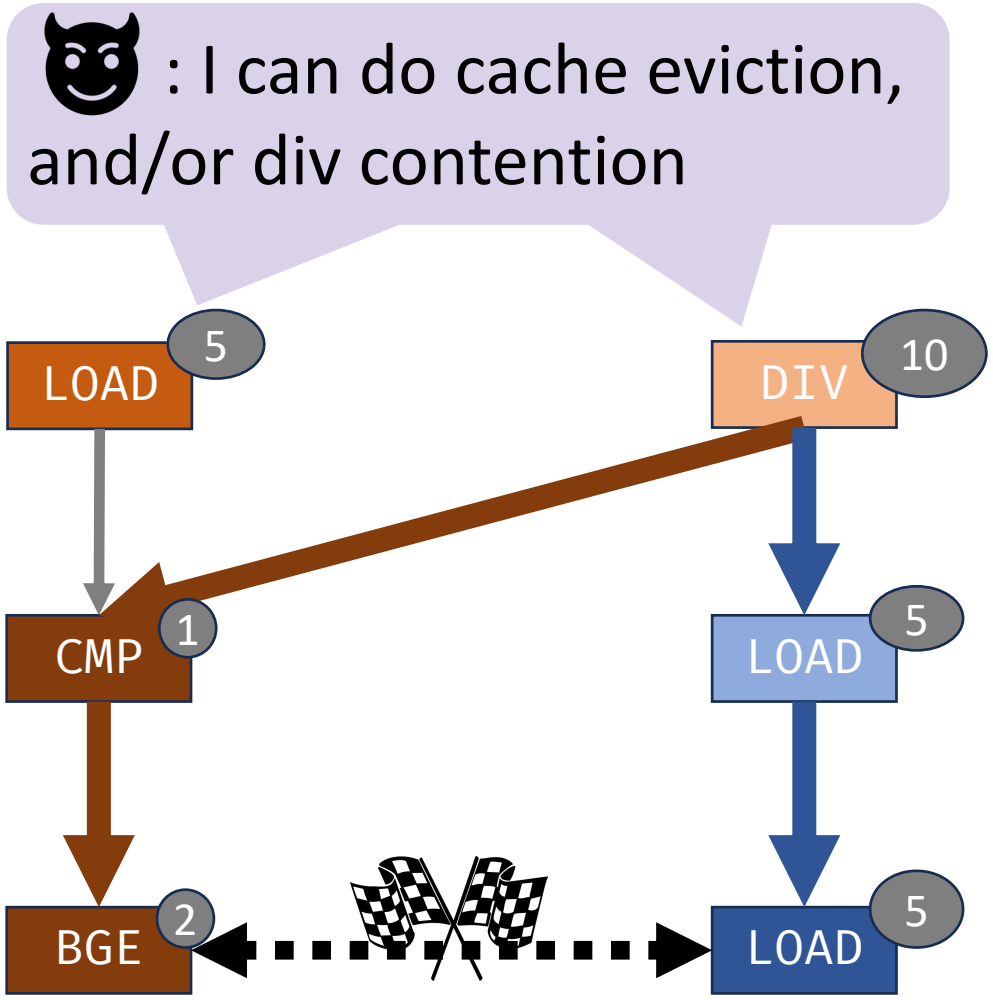


Step B: Simulating windowing power



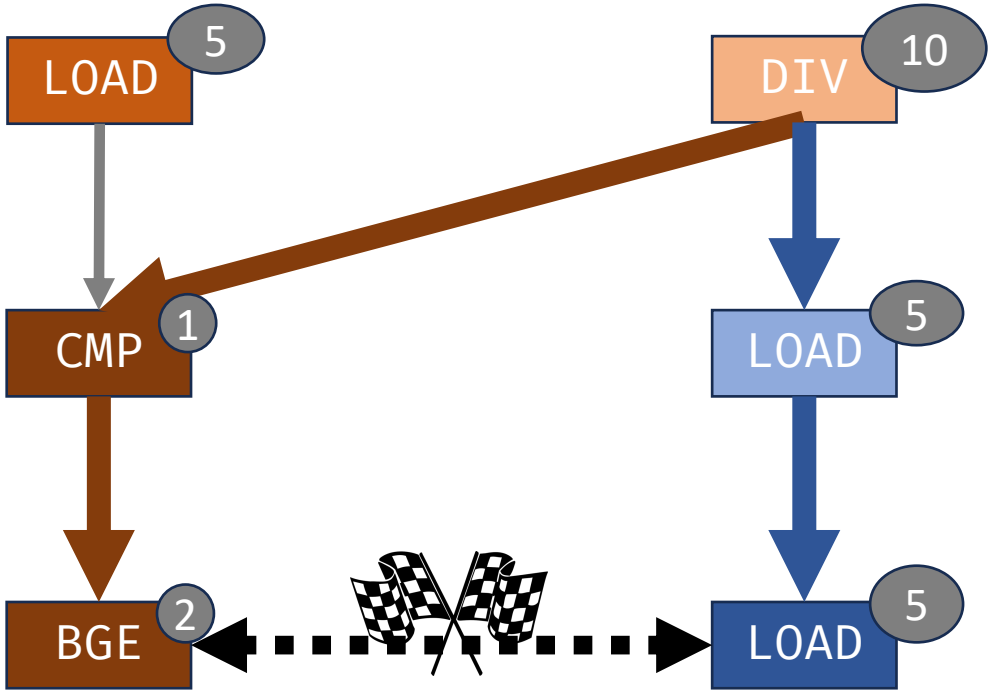
Attack Pattern
Do nothing
Cache eviction
Div contention
Cache eviction + Div contention

Step B: Simulating windowing power

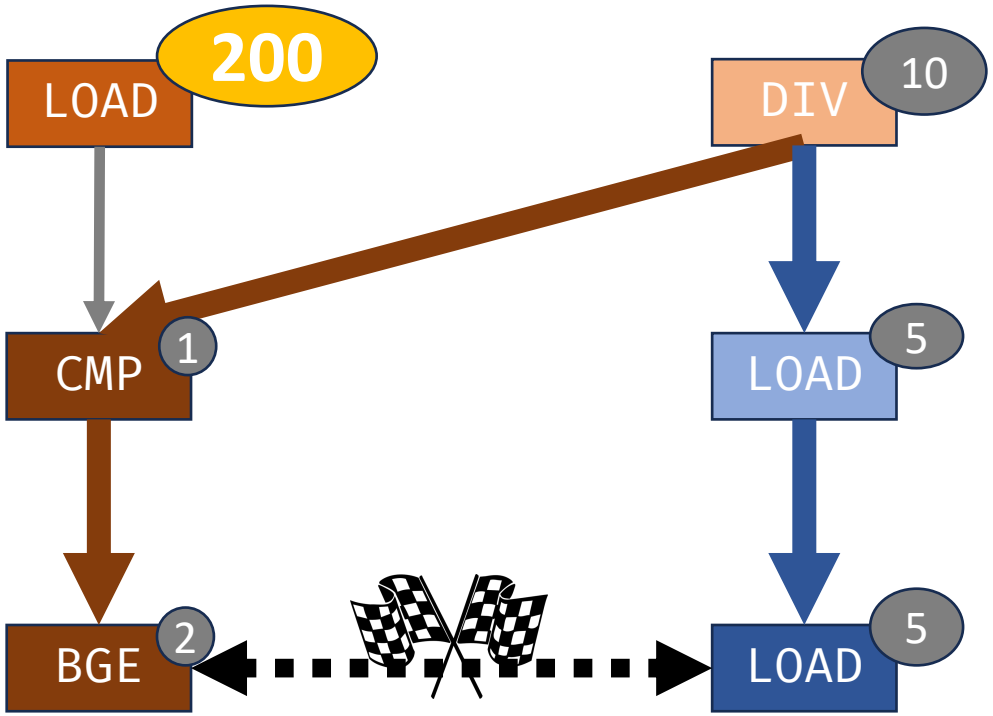


Attack Pattern	Increase in Timing Condition Index
Do nothing	?
Cache eviction	?
Div contention	?
Cache eviction + Div contention	?

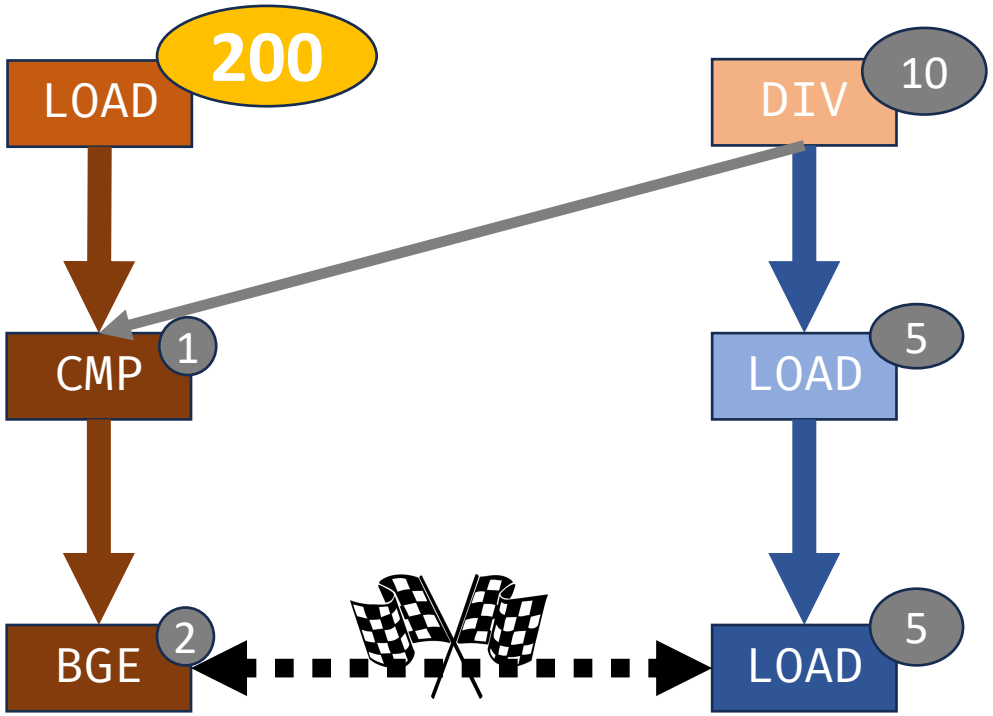
What if I do cache eviction?



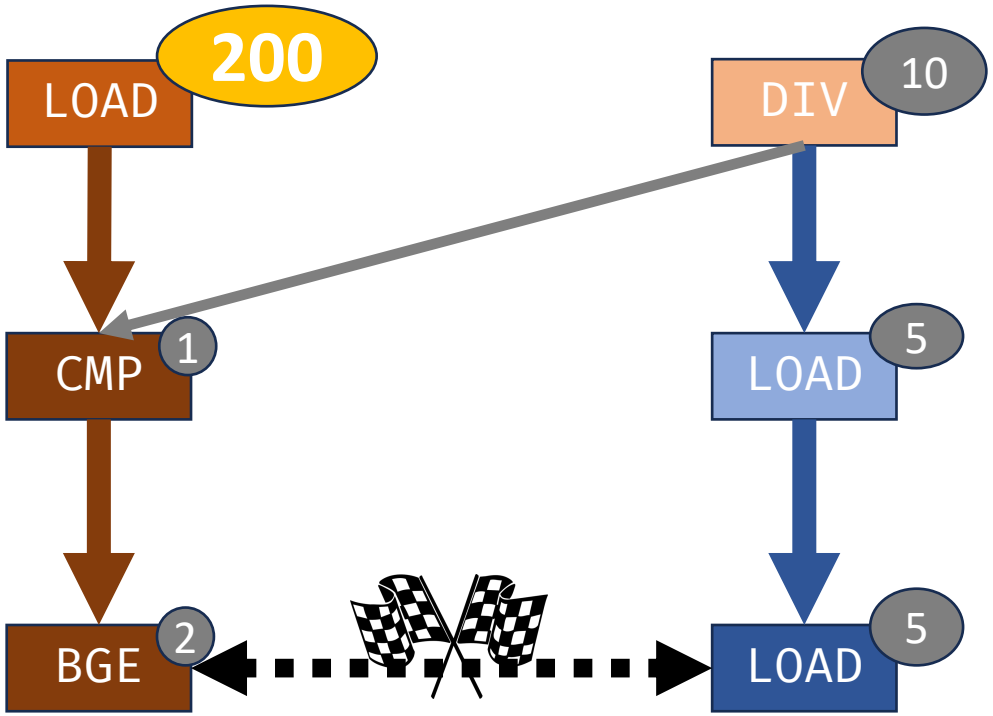
What if I do cache eviction?



What if I do cache eviction?



What if I do cache eviction?

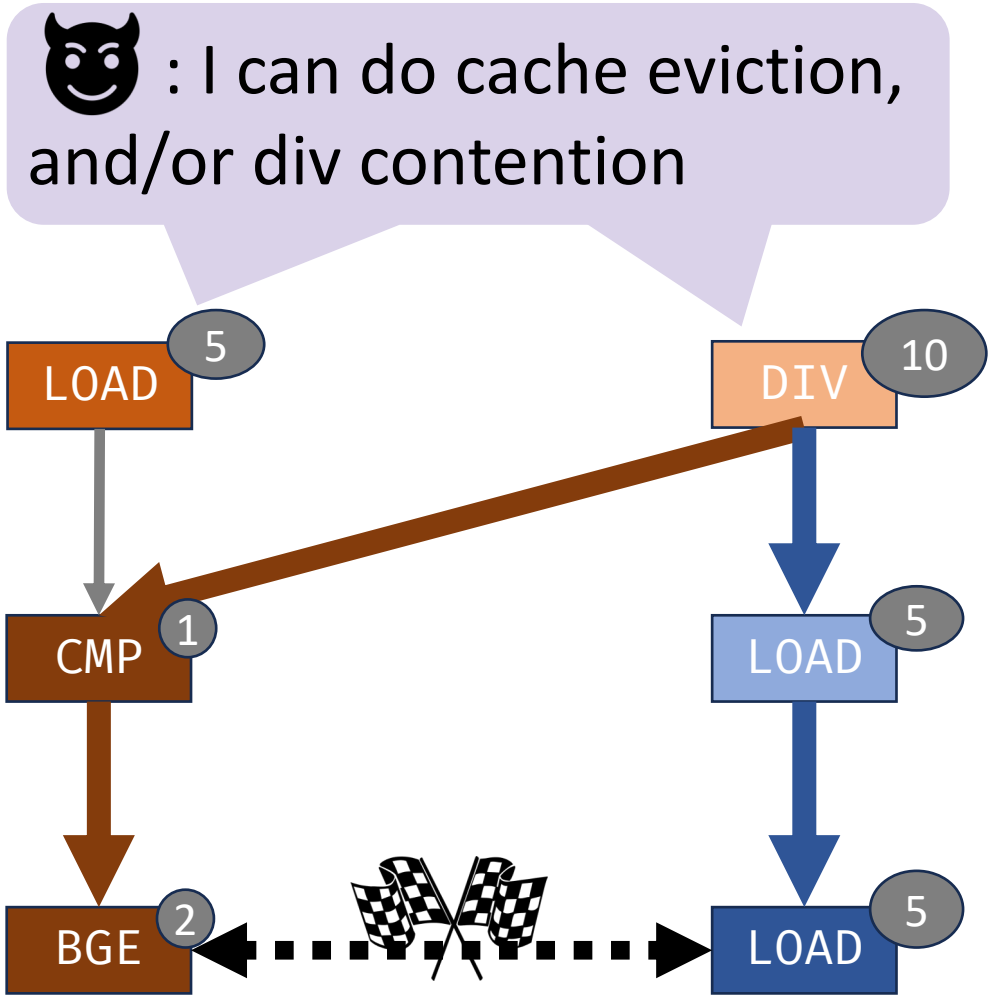


Timing Condition Index was -7.

Now, it increases to **183**!

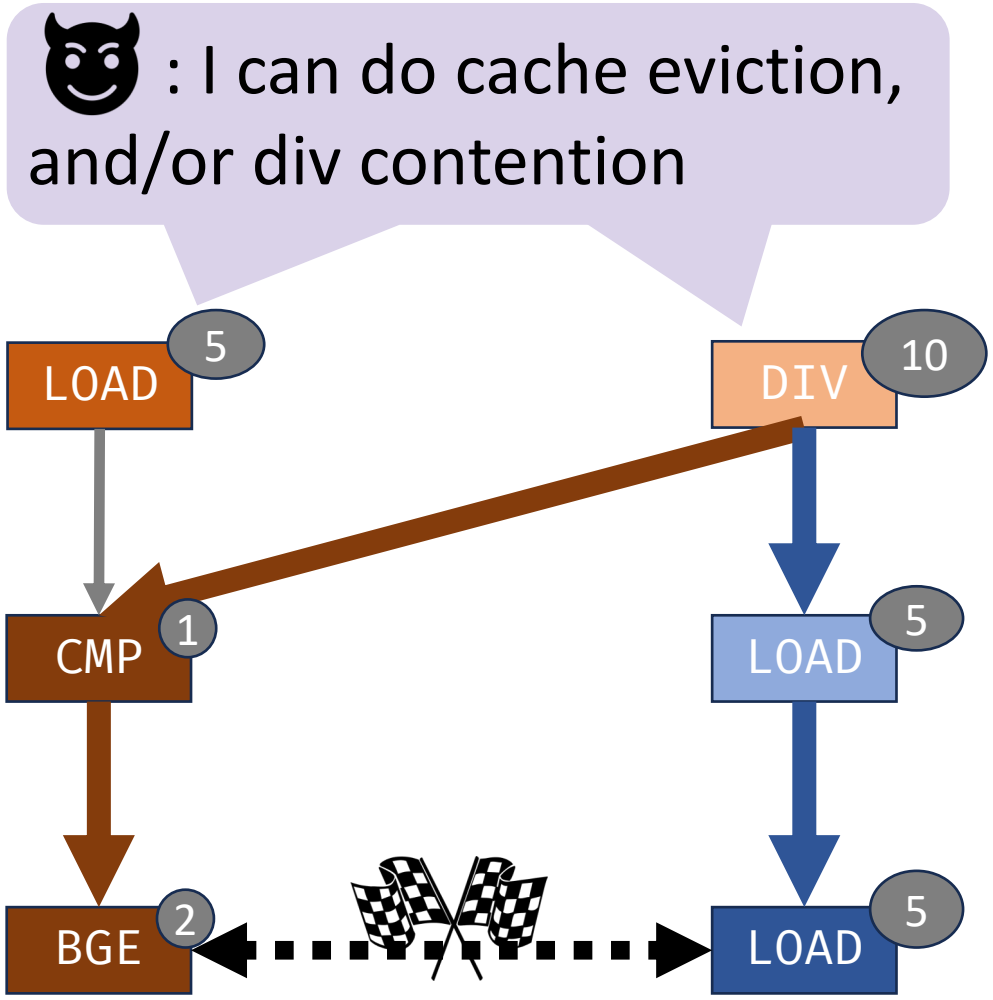
The attack pattern is effective :-)

Step B: Simulating windowing power



Attack Pattern	Increase in Timing Condition Index
Do nothing	0
Cache eviction	+190
Div contention	-140
Cache eviction + Div contention	+50

Step B: Simulating windowing power

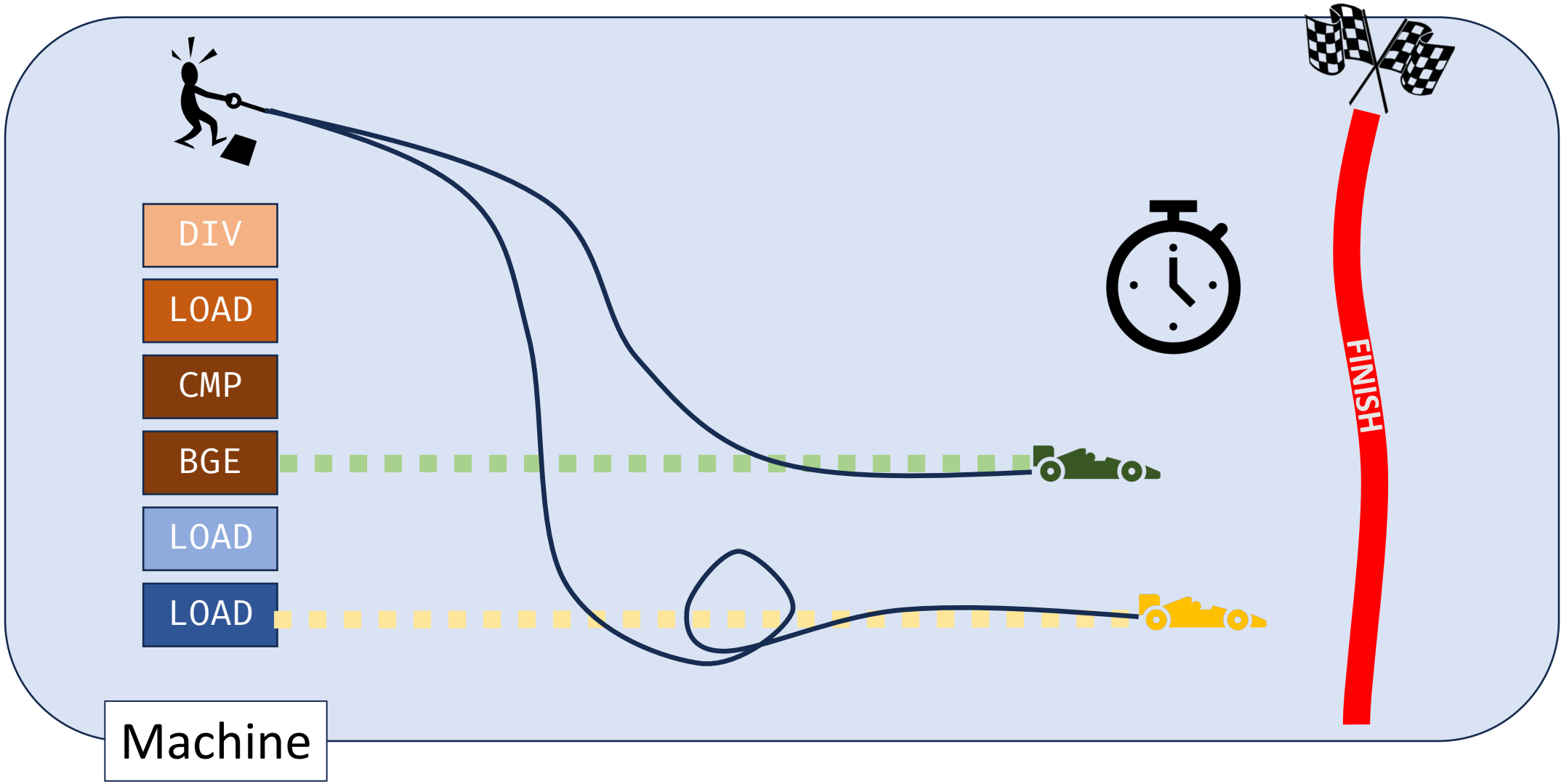


Attack Pattern	Increase in Timing Condition Index
Do nothing	0
Cache eviction	+190
Div contention	-140
Cache eviction + Div contention	+50

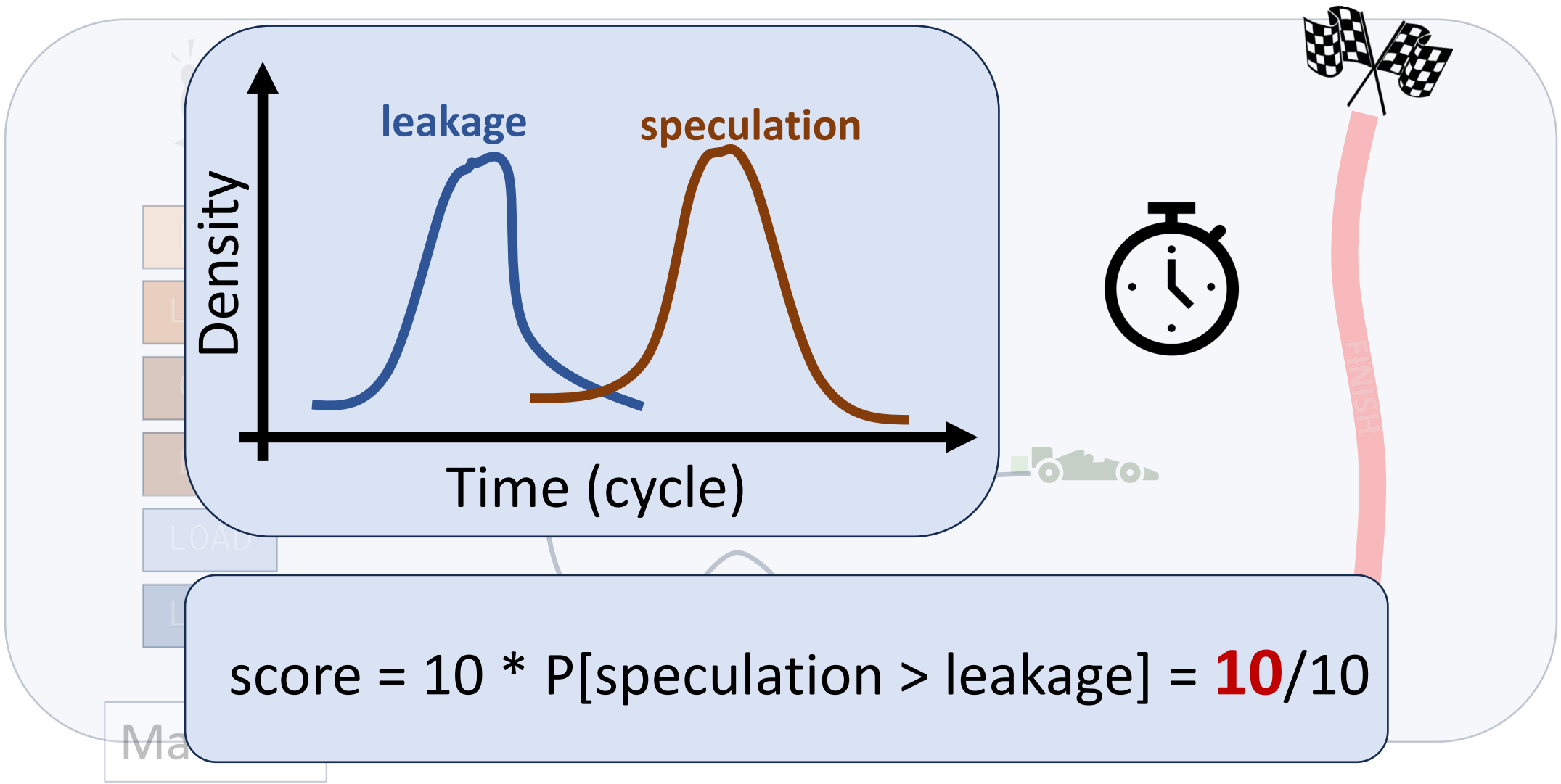
Step C: Measuring exploitability

Step C: Measuring exploitability

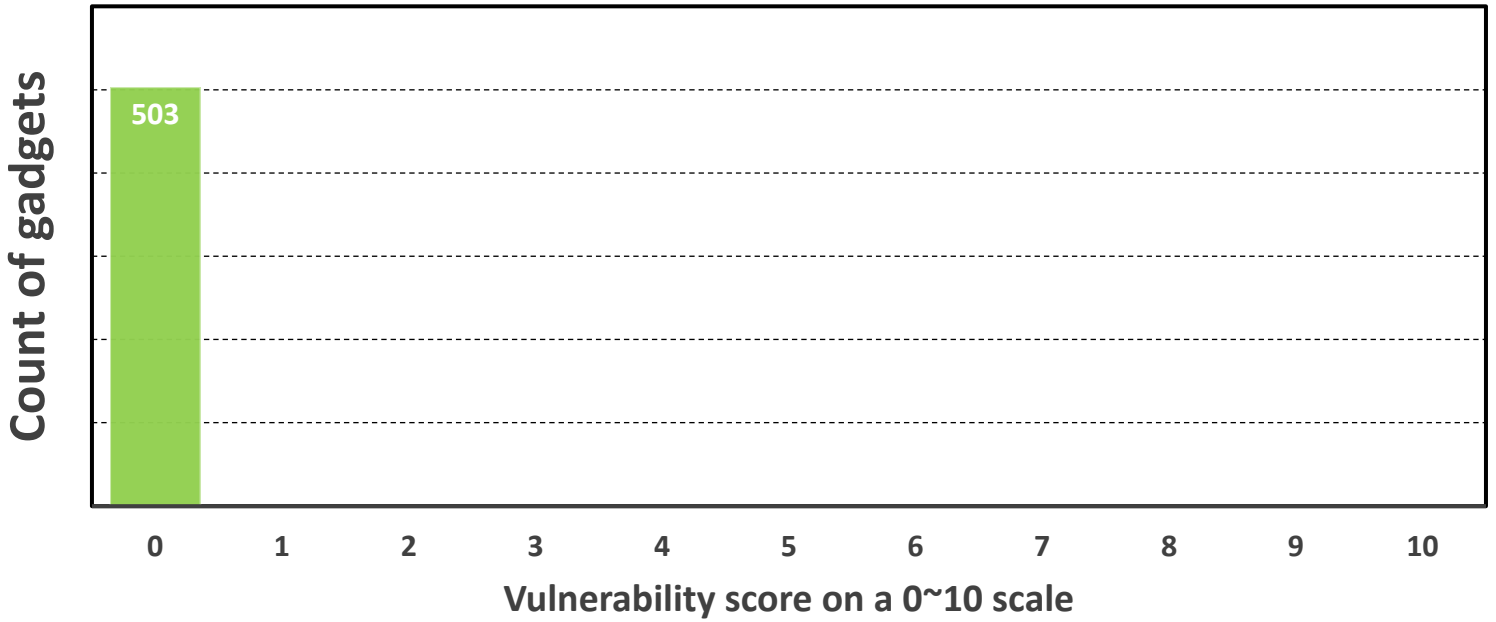
Step C: Measuring exploitability



Step C: Measuring exploitability

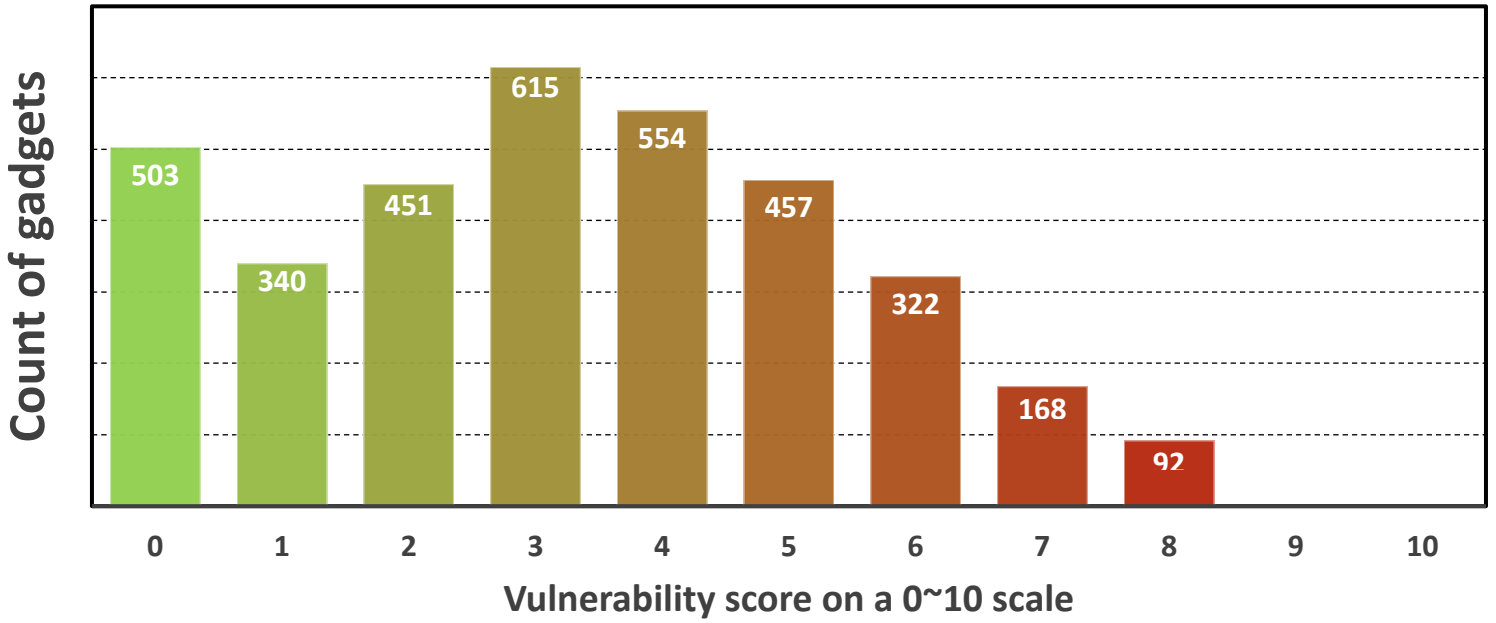


Evaluation: Gauging exploitability



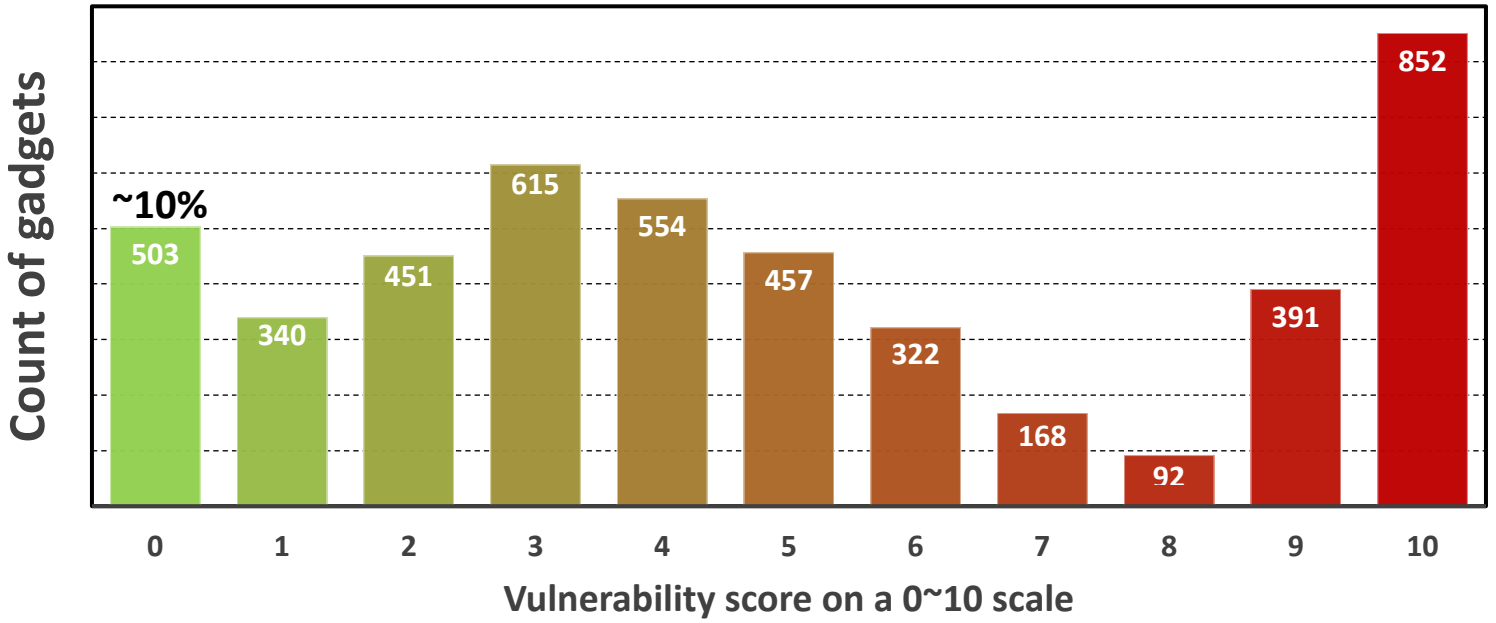
- Target: gadgets with vulnerable information flow, identified by SOTA scanners.
- Applications: 6 security-centric applications and Linux kernel

Evaluation: Gauging exploitability



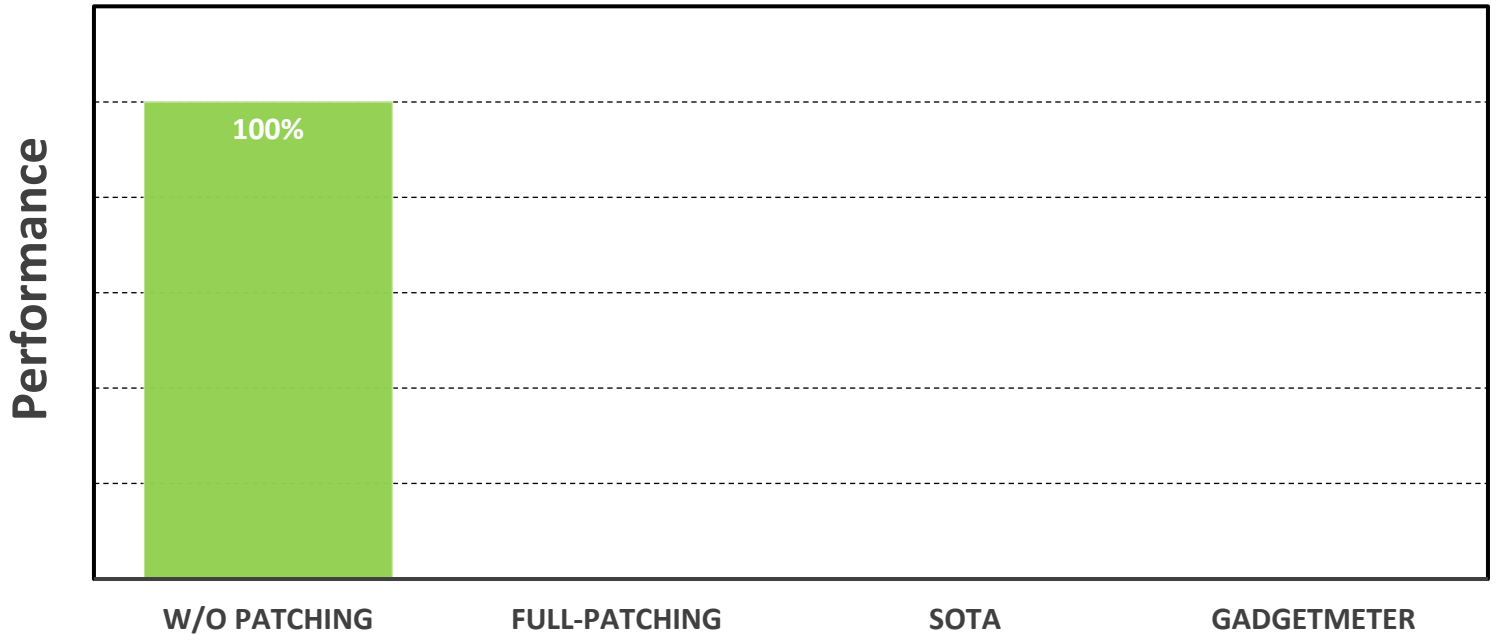
- Target: gadgets with vulnerable information flow, identified by SOTA scanners.
- Applications: 6 security-centric applications and Linux kernel

Evaluation: Gauging exploitability



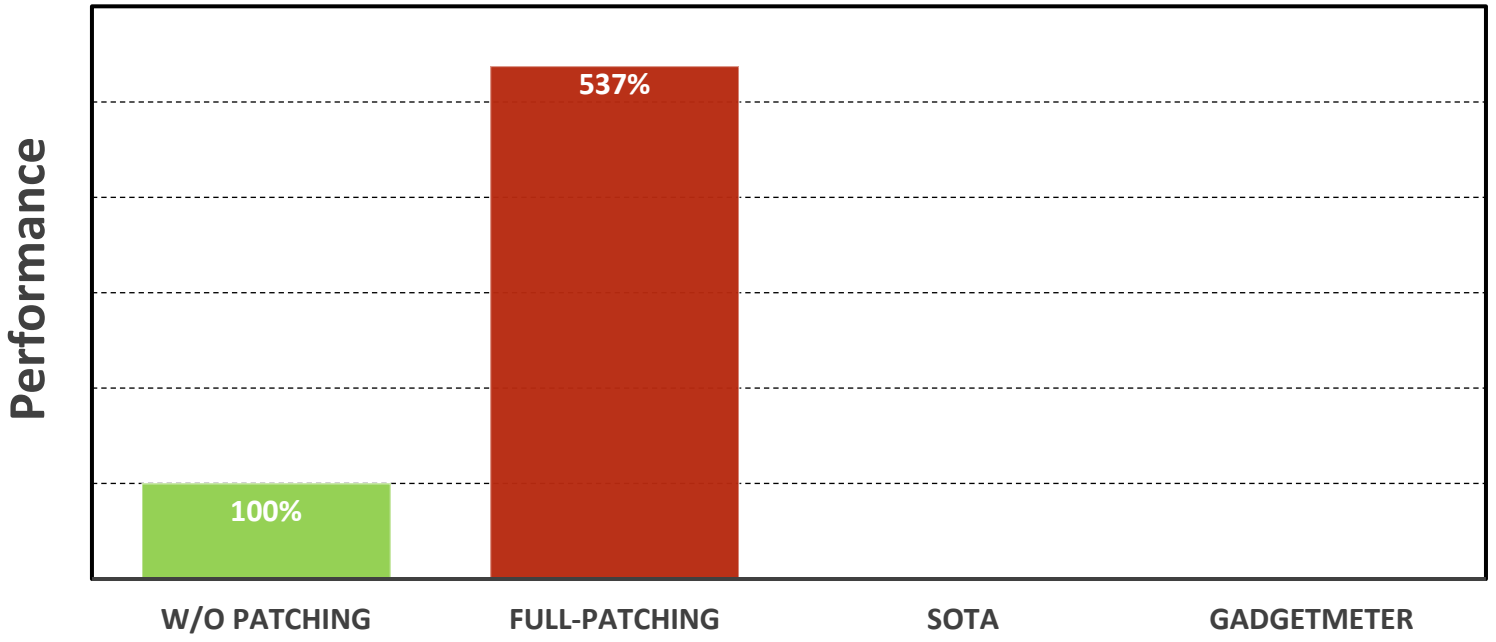
- Target: gadgets with vulnerable information flow, identified by SOTA scanners.
- Applications: 6 security-centric applications and Linux kernel

Evaluation: Performance improvement



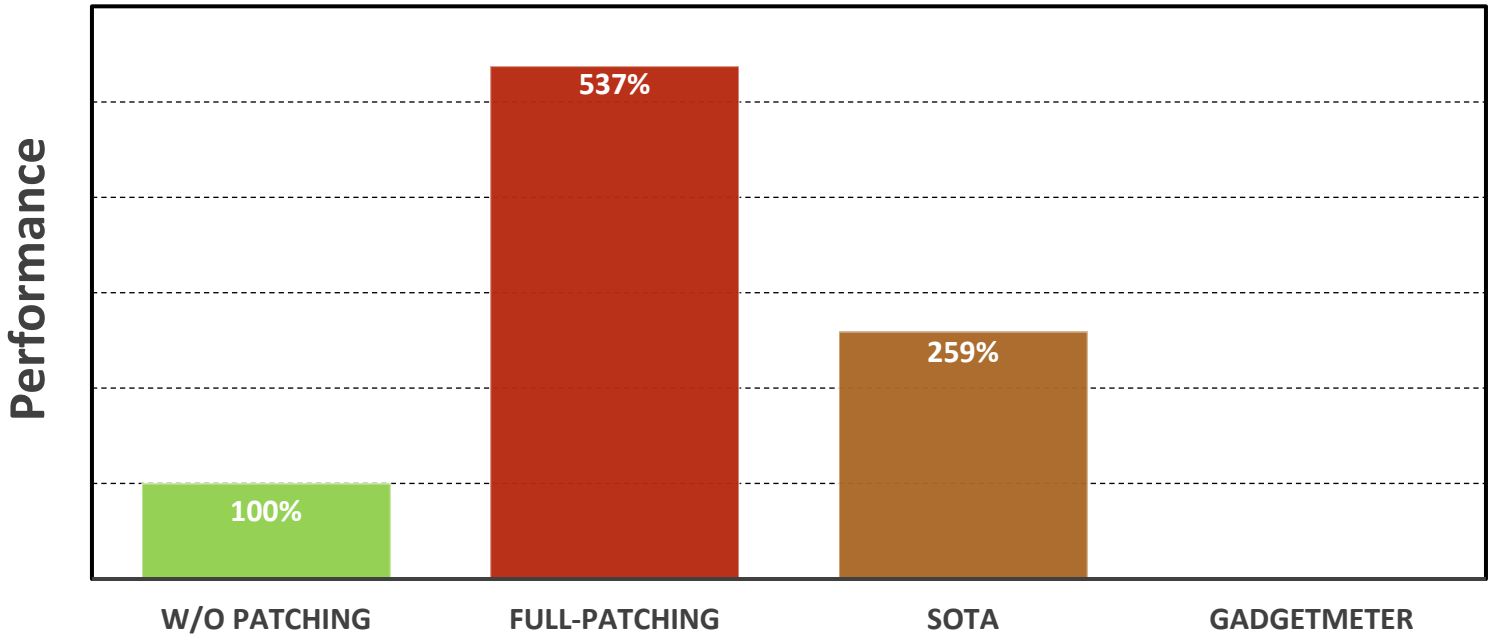
- Patching method: LFENCE serialization.

Evaluation: Performance improvement



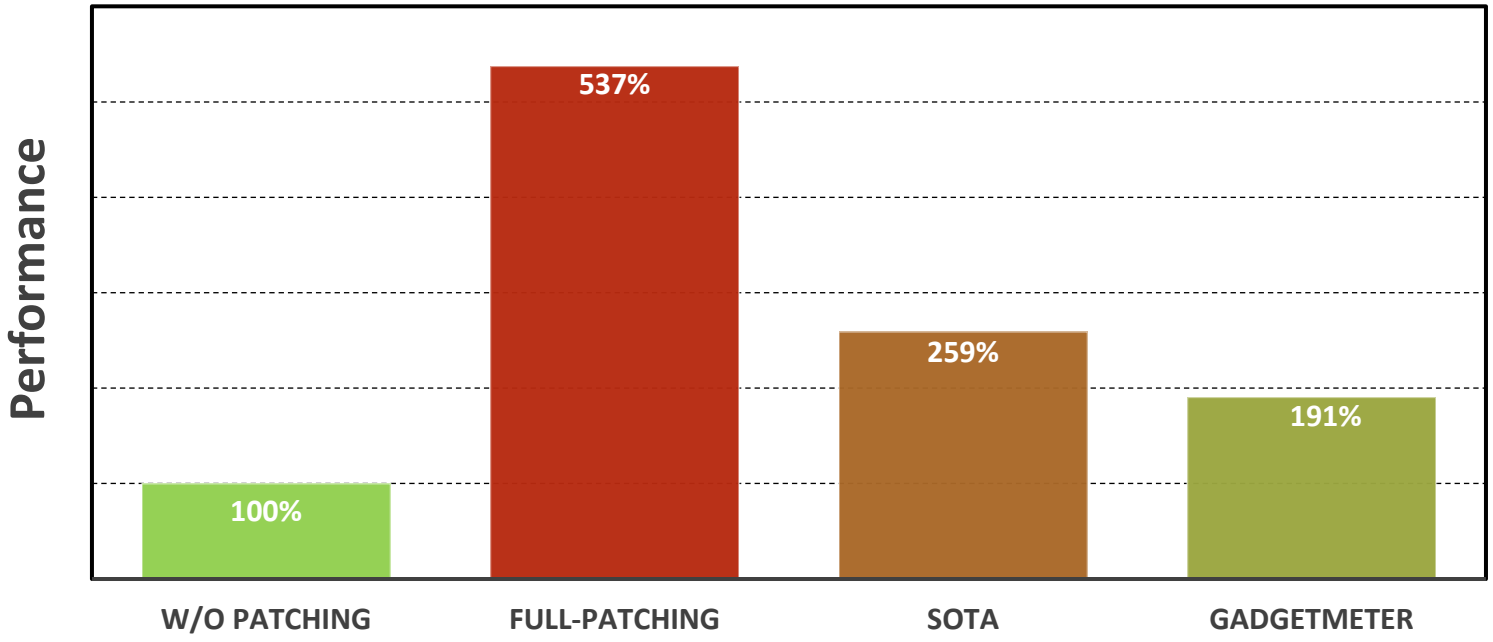
- Patching method: LFENCE serialization.

Evaluation: Performance improvement



- Patching method: LFENCE serialization.

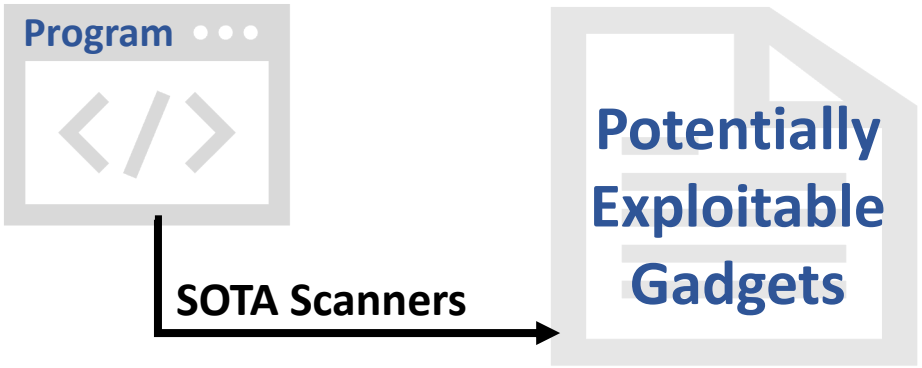
Evaluation: Performance improvement



- Patching method: LFENCE serialization.
- Reduce overhead by 20.66%, compared with SOTA.

Conclusion

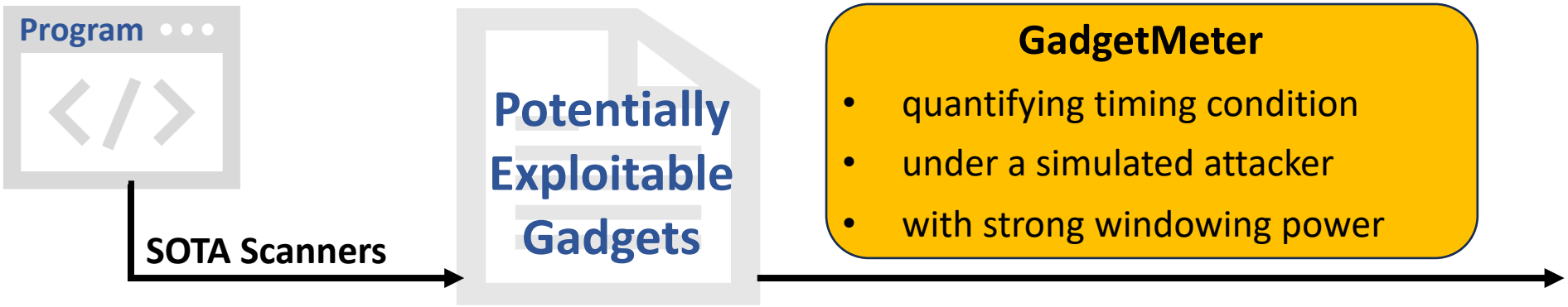
qiling07.github.io



- ✓ **Few** false negatives
- ✗ **Many** false positives
- ✗ **Binary** detection results

Conclusion

qiling07.github.io



✓ **Few** false negatives

✗ **Many** false positives

✗ **Binary** detection results

Conclusion

qiling07.github.io