Pensieve: Microarchitectural Modeling for Formal Security Evaluation

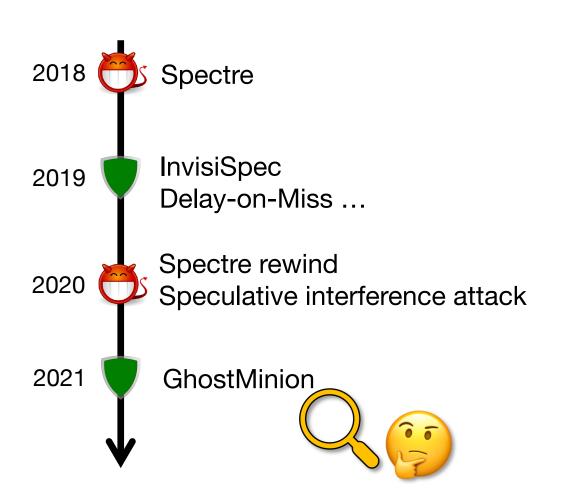
Yuheng Yang, Thomas Bourgeat, Stella Lau, Mengjia Yan

To appear at ISCA'23



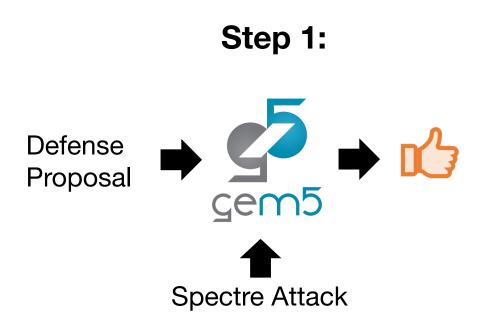


Problem: the Cat-and-Mouse Game





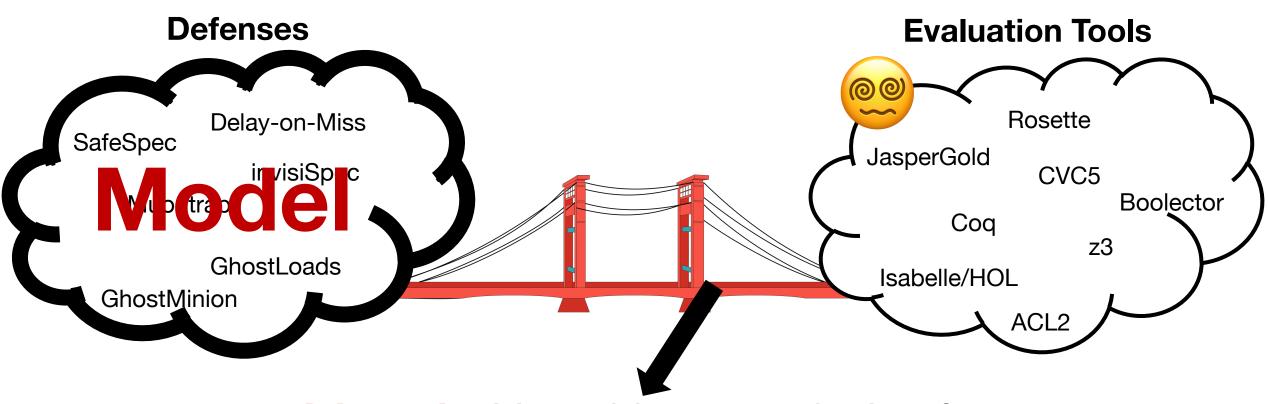
Problem: Weak Security Evaluation





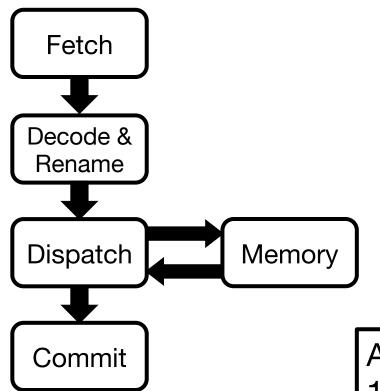
We need a principled, trustworthy security evaluation tool!

Challenge: Bridge the Gap



Aligned with architectural design flow.

Defense Design flow



Example: delay speculative requests

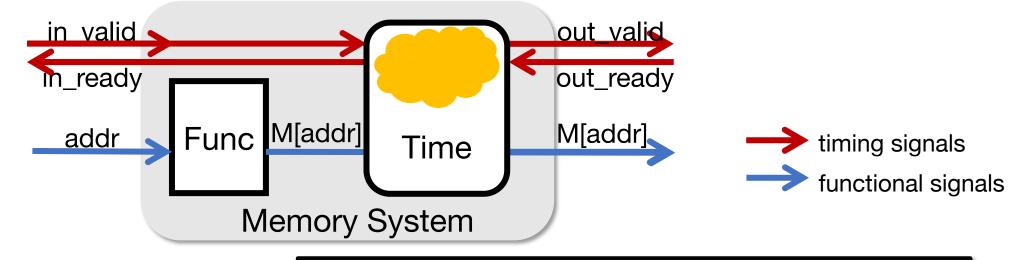


An architecture modeling method should be

- 1. Modular
- 2. Precise on describing timing behaviors
- 3. Represent a space of designs

Pensieve Modeling

- Decouple timing and functionality using the hand-shaking interface
- Represent a space of timing behavior

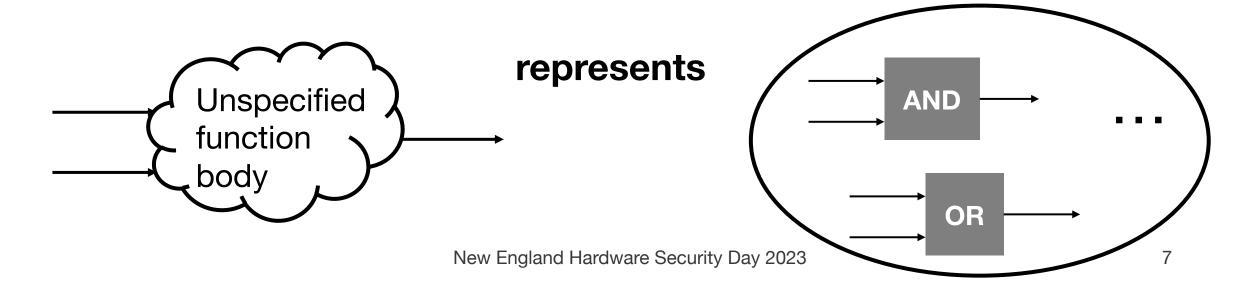


An architecture modeling method should be

- Modular
- 2. Precise on describing timing behaviors
 - 3. Represent a space of designs

Uninterpreted Function (UF)

- A UF represents space of functions with the same input/output types
 - Example: Bool UF (Bool, Bool)
- UF helps us
 - state "what" affects the output,
 - abstract away the details on "how" the input affects the output



Pensieve Modeling: Using UFs

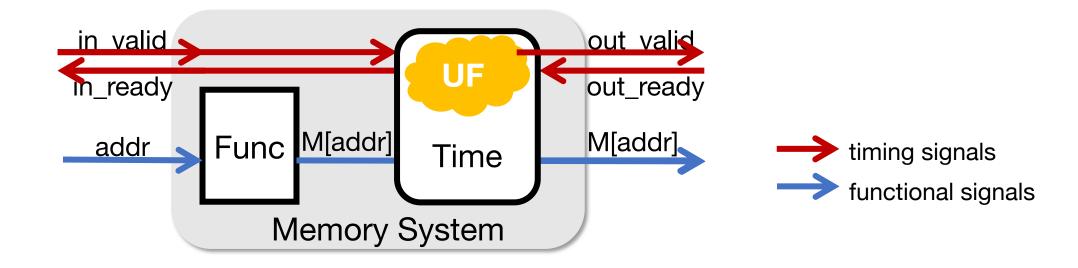
Examples:

```
Multiply_req_latency = UF(historyOf(in_valid))
Multiply_req_latency = UF(historyOf(in_valid, in_operands))
Memory_req_latency = UF(historyOf(in_valid, in_addr))
```

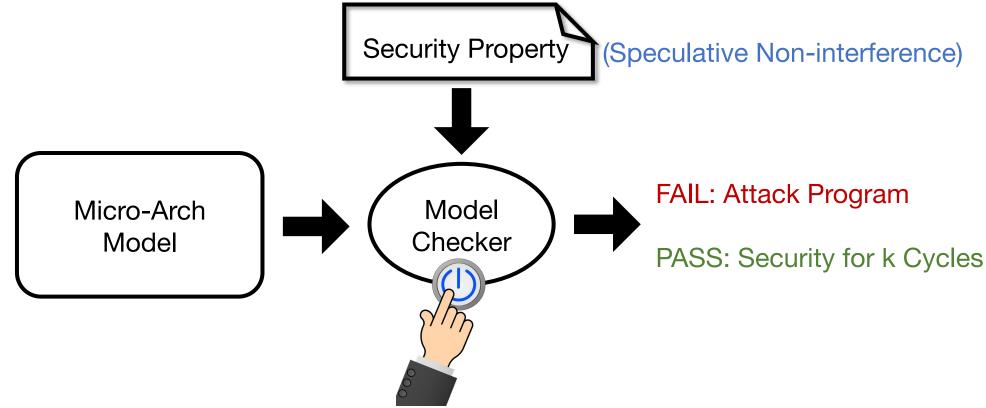
Pensieve can use **simple** models to cover **space** of microarchitectures with **complex** timing behaviors

Pensieve Modeling

- Decouple timing and functionality using the hand-shaking interface
- Represent a space of timing behavior with uninterpreted functions



Pensieve Security Evaluation Framework



Pensieve finds **unknown** security vulnerabilities in the latest speculative execution defense, i.e., GhostMinion [2021]

New Attack on GhostMinion Summary

speculative interference attack

```
y = .....

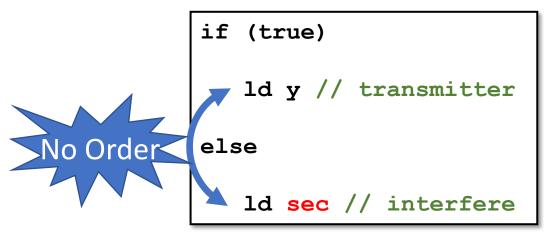
Older

ld y // transmitter

if (false)

Younger ld sec // interfere
```

new attack variant



Takeaway: Manual evaluation can easily be unsound, we need Pensieve, a trustworthy evaluation tool