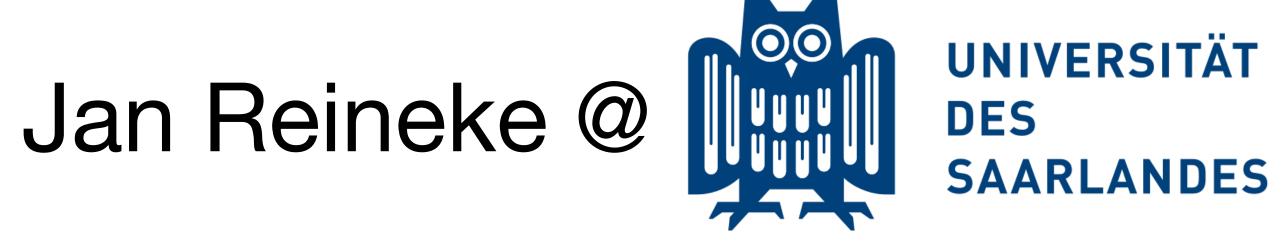
Hardware-Software Contracts for Safe and Secure Systems

Joint work with Marco Guarnieri, Pepe Vila @ IMDEA Software, Madrid Boris Köpf @ Microsoft Research, Cambridge, UK Andreas Abel, Sebastian Hahn, Valentin Touzeau @ Saarland University

Supported by the European Research Council and an Intel Strategic Research Alliance (ISRA)



The Need for HW/SW Contracts

"Stone-age" Computing

Applications implemented data transformations: e.g. payroll processing

"Stone-age" Computing

Applications implemented data transformations: e.g. payroll processing

Hardware:

- isolated, on-site
- Imited interaction with environment



Author: <u>ArnoldReinhold</u> License: <u>CC BY-SA 3.0</u>

"Stone-age" Computing

Applications implemented data transformations: e.g. payroll processing

Hardware:

- isolated, on-site
- Imited interaction with environment

HW/SW Contract: Instruction Set Architecture



Author: ArnoldReinhold License: CC BY-SA 3.0

ISA Abstraction

High-level languages

Compiler

Instruction set architecture (ISA)

Implementation Microarchitecture



ISA Abstraction: Benefits

Can program **independently** of microarchitecture

Instruction set architecture (ISA)

Can implement **arbitrary optimizations** as long as ISA semantics are obeyed

"Modern" (?) Computing

Applications are:

- *Data-driven*: e.g. deep neural networks • *Distributed*: e.g. locally + in the cloud • Open: e.g. untrusted code in the browser 🜔 • *Real-time*: interacting with the physical environment



"Modern" (?) Computing

Applications are:

- Data-driven: e.g. deep neural networks Real-time: interacting with the physical environment

- *Distributed*: e.g. locally + in the cloud • Open: e.g. untrusted code in the browser 🔘

What are the implications for HW/SW contracts?

Inadequacy of the ISA + current µArchitectures: Real-time Systems

Instruction set architecture (ISA)



Abstracts from time

Inadequacy of the ISA + current µArchitectures: Real-time Systems

Instruction set architecture (ISA) Abstracts from time

Can implement arbitrary **unpredictable** optimizations as long as ISA semantics are obeyed



Inadequacy of the ISA + current µArchitectures: Real-time Systems

Programs do not have a timed semantics Programs have no control over timing

Instruction set architecture (ISA)

as long as ISA semantics are obeyed



Abstracts from time

Can implement arbitrary unpredictable optimizations

State-of-the-art: Handcrafted Microarchitectural Timing Models

Instruction set architecture (ISA)



Microarchitectural timing model

Manual Modeling

Microarchitecture





models timing behavior + still no control over timing

unpredictable



State-of-the-art: Handcrafted Microarchitectural Timing Models

Instruction set architecture (ISA)



Microarchitectural timing model

Manual Modeling

Microarchitecture





- limited to particular microarchitectures
- + probably incorrect
- + yield expensive or imprecise analysis

models timing behavior + still no control over timing

unpredictable





Timed Instruction Set Architecture



Timed Instruction Set Architecture

Admit wide range of high-performance microarchitectural implementations





Programs have a **timed semantics** that is **efficiently predictable** Programs have **control** over timing

Timed Instruction Set Architecture

Admit wide range of high-performance microarchitectural implementations





Some answers:

D. Bui, E. Lee, I. Liu, H. Patel, and J. Reineke: Temporal Isolation on Multiprocessing Architectures DAC 2011

S. Hahn and J. Reineke: Design and Analysis of SIC:

A Provably Timing-Predictable Pipelined Processor Core



Inadequacy of the ISA + current µArchitectures: Side-channel security

Instruction set architecture (ISA)



No guarantees about side channels



Inadequacy of the ISA + current µArchitectures: Side-channel security

Instruction set architecture (ISA) No guarantees about side channels

Can implement arbitrary **insecure** optimizations as long as ISA semantics are obeyed **SPECTRE**



Inadequacy of the ISA + current µArchitectures: Side-channel security

Impossible to program securely on top of ISA cryptographic algorithms? sandboxing untrusted code?

Instruction set architecture (ISA)

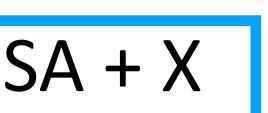


- No guarantees about side channels
- Can implement arbitrary insecure optimizations as long as ISA semantics are obeyed **SPECTRE**



A Way Forward: HW/SW Security Contracts

Hardware-Software Contract = ISA + X



Succinctly captures possible information leakage



A Way Forward: HW/SW Security Contracts

Hardware-Software Contract = ISA + X

Can implement **arbitrary insecure optimizations** as long as contract is obeyed

Succinctly captures possible information leakage



A Way Forward: HW/SW Security Contracts

Can program **securely** on top contract **independently** of microarchitecture

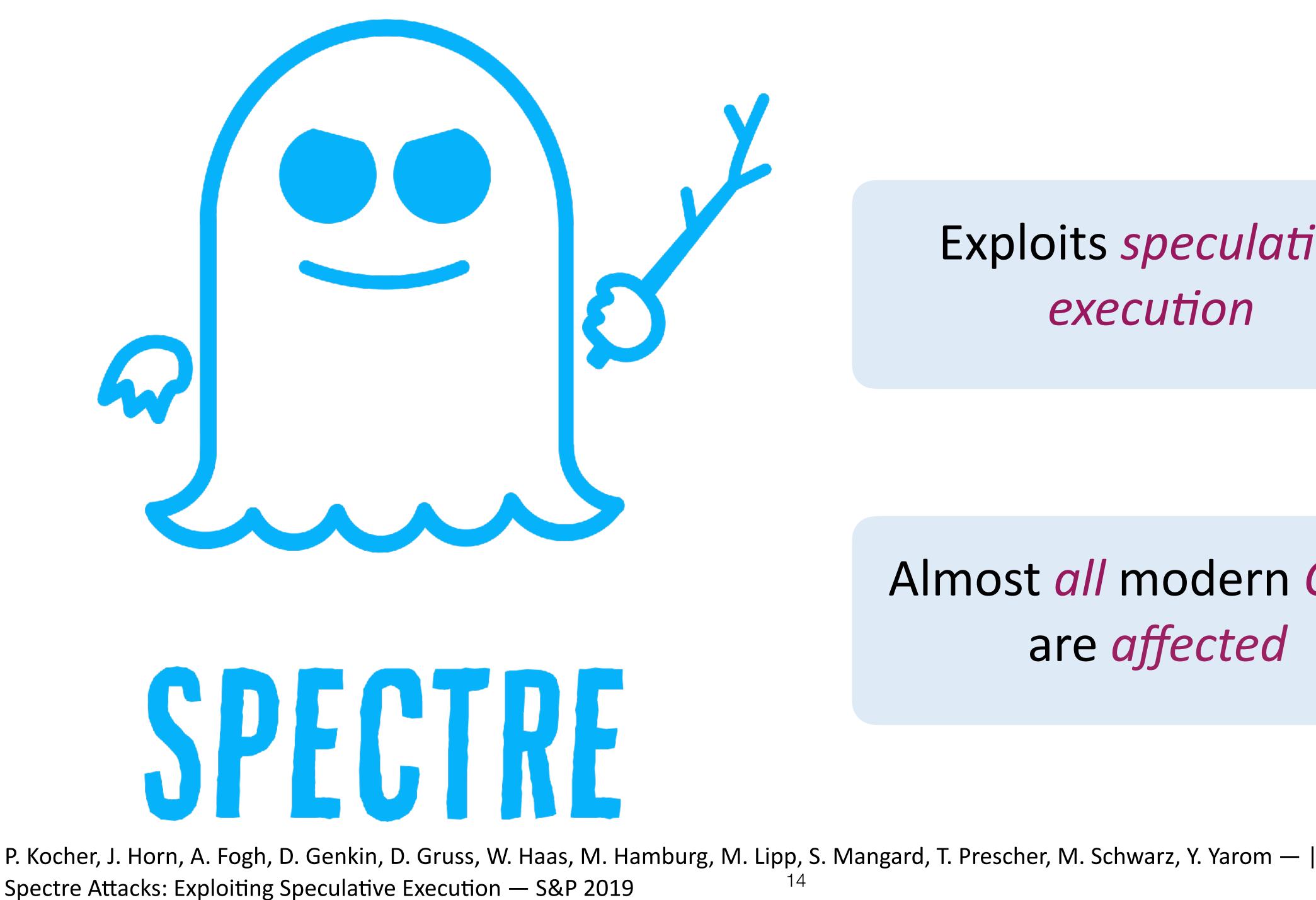
Hardware-Software Contract = ISA + X

Can implement **arbitrary insecure optimizations** as long as contract is obeyed

Succinctly captures possible information leakage



A Concrete Challenge: Spectre



Exploits *speculative* execution

Almost *all* modern *CPUs* are affected

14

if (x < A_size) y = A[x] z = B[y*512] end

1. x is out of bounds 1. if $(\mathbf{x} < \mathbf{A} \ size)$ 2. $\mathbf{y} = \mathbf{A}[\mathbf{x}]$ 3. z = B[y*512]end 4.





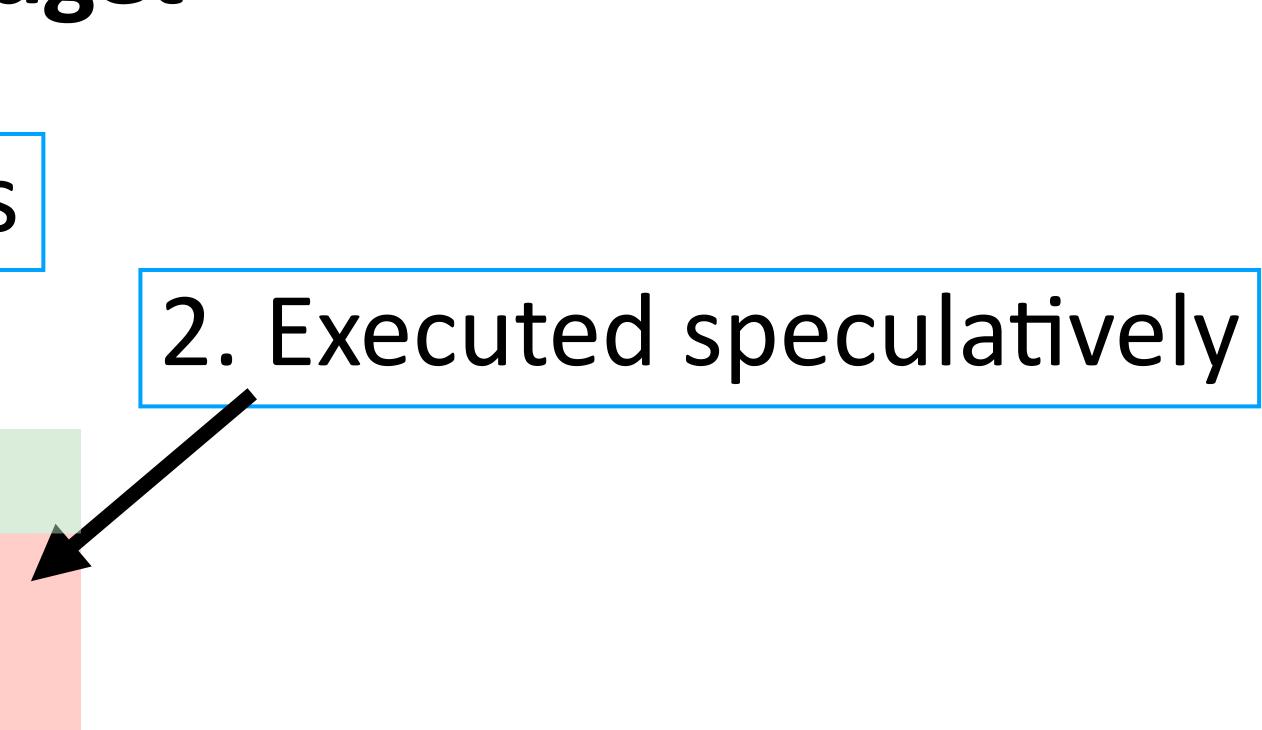
1. x is out of bounds 1. if $(\mathbf{x} < \mathbf{A} \text{ size})$ $\mathbf{y} = \mathbf{A}[\mathbf{x}]$ 2. z = B[y*512]3. end 4.



2. Executed speculatively



1. x is out of bounds 1. if (x < A size) $y = \mathbf{A}[\mathbf{x}]$ 2. z = B[y*512]3. end 4. 3. Leaks **A**[**x**] via data cache



InvisiSpec: Making Speculative Execution Invisible in the Cache Hierarchy

Mengjia Yan[†], Jiho Choi[†], Dimitrios Skarlatos, Adam Morrison^{*}, Christopher W. Fletcher, and Josep Torrellas University of Illinois at Urbana-Champaign *Tel Aviv University {myan8, jchoi42, skarlat2}@illinois.edu, mad@cs.tau.ac.il, {cwfletch, torrella}@illinois.edu

InvisiSpec: Making Speculative Execution Invisible in the Cache Hierarchy

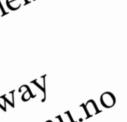
Mengjia Yan[†], Jiho Choi[†], Dimitrios Skarlatos, Adam Morrison^{*}, Christopher W. Fletcher, and Josep Torrellas University of Illinois at Urbana-Champaign *Tel Aviv University {myan8, jchoi42, skarlat2}@illinois.edu, mad@cs.tau.ac.il, {cwfletch, torrella}@illinois.edu

CleanupSpec: An "Undo" Approach to Safe Speculation Moinuddin K. Qureshi moin@gatech.edu Georgia Institute of Teck

Efficient Invisible Speculative Execution through Efficient Invisible Delay and Value Prediction InvisiSpec: Making Speculative Execution Invisible in the Cache Hierarchy Mengjia Yan[†], Jiho Choi[†], Dimitrios Skarlatos, Adam Morrison^{*}, Christopher W. Fletcher, and Josep Torrellas University of Illinois at Urbana-Champaign *Tel Aviv University {myan8, jchoi42, skarlat2}@illinois.edu, mad@cs.tau.ac.il, {cwfletch, torrella}@illinois.edu Norwegian University of Science and CleanupSpec: An "Undo" Approach to Safe Speculation Trondheim, Norway magnus.sjalander@ntnu.no Alexandra Jimborean Uppsala University istos.sakalis@it.uu.se Uppsala, Sweden lexandra.jimborean@it.uu.se Moinuddin K. Qureshi moin@gatech.edu Georgia Institute of Tech





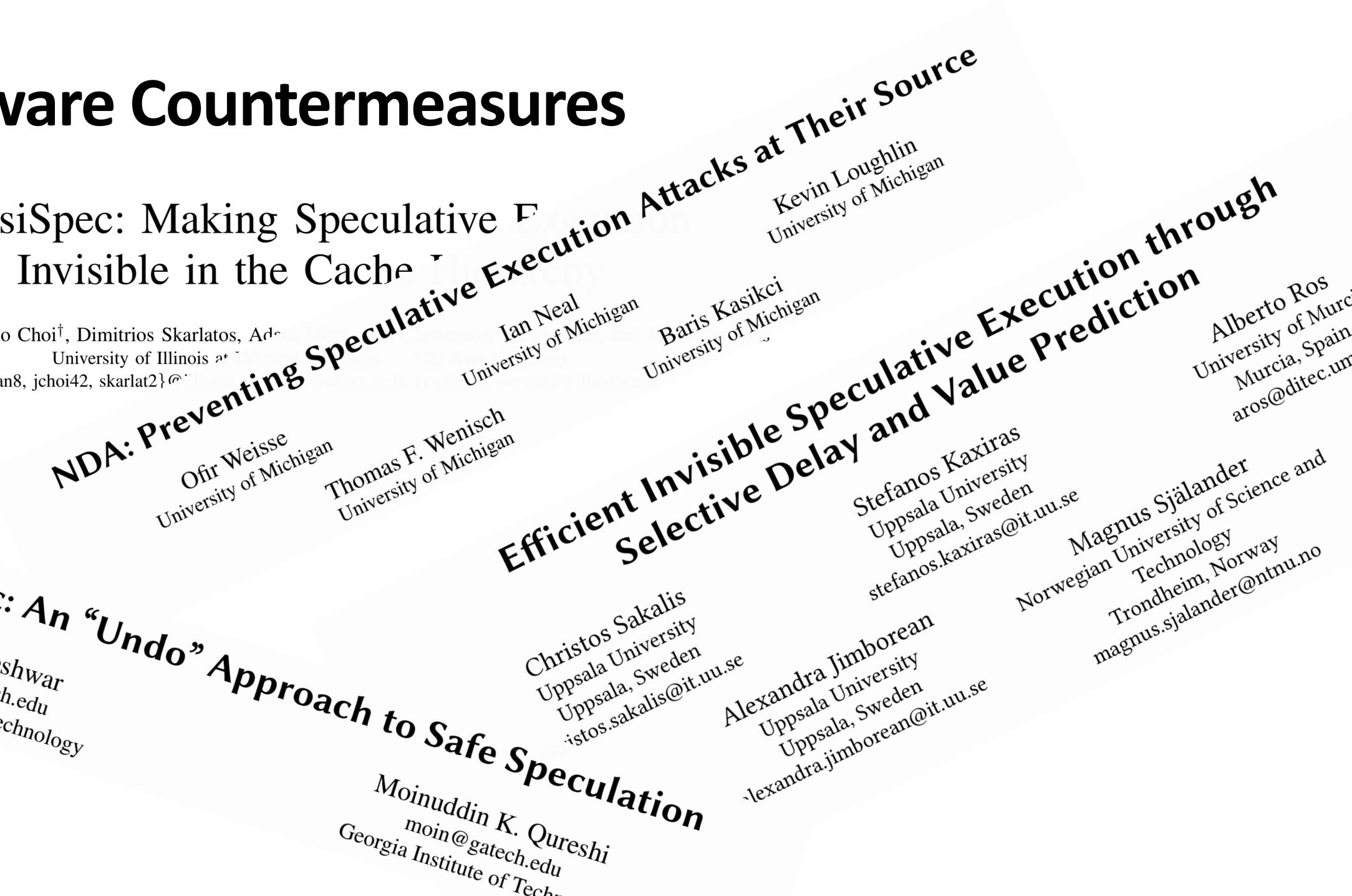


InvisiSpec: Making Speculative F

Mengjia Yan[†], Jiho Choi[†], Dimitrios Skarlatos, Adr {myan8, jchoi42, skarlat2}@"

University of Michigan CleanupSpec: An "Undo" Approach to Safe Speculation

Moinuddin K. Qureshi moin@gatech.edu Georgia Institute of Tech



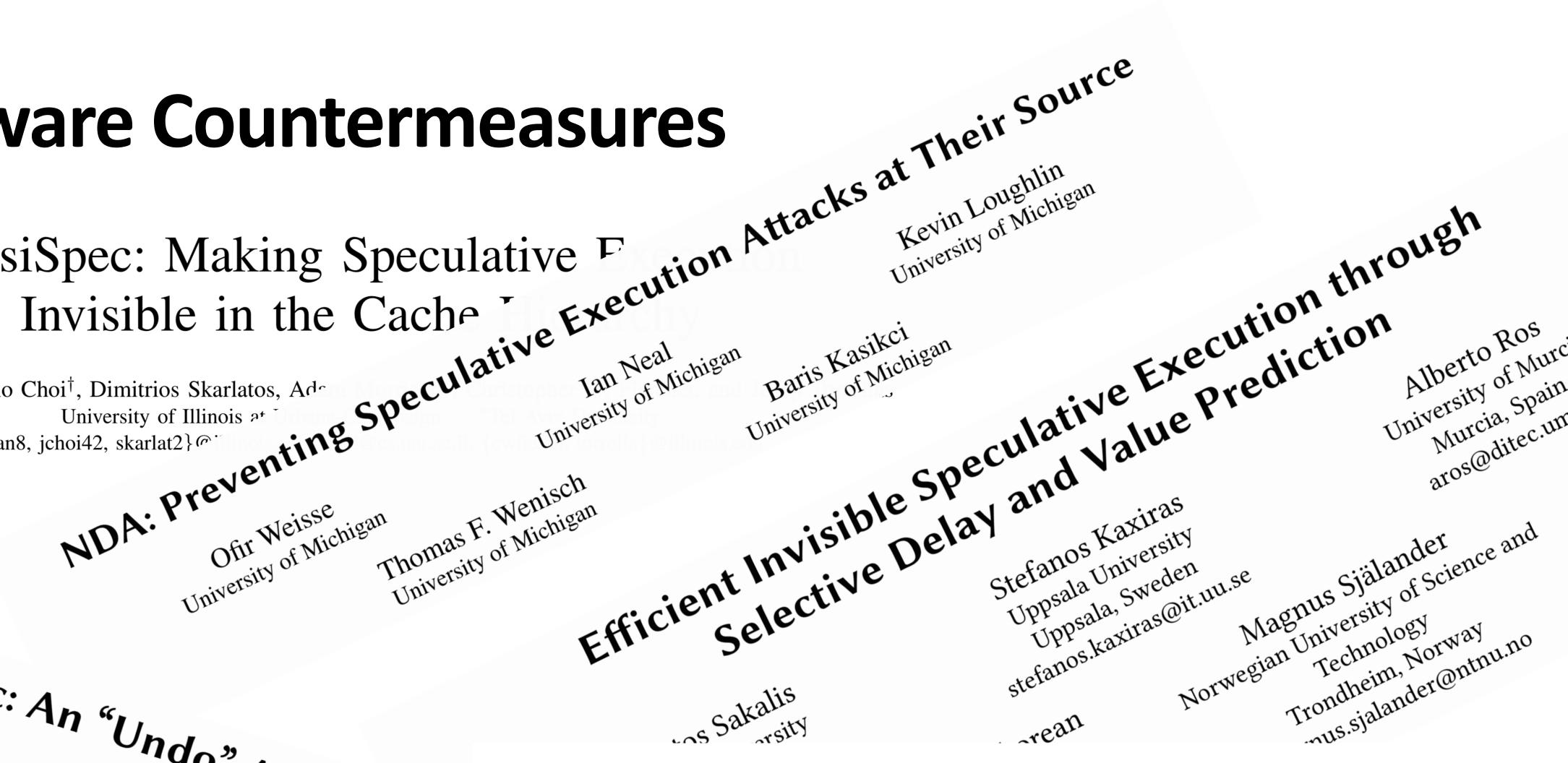


InvisiSpec: Making Speculative F

Mengjia Yan[†], Jiho Choi[†], Dimitrios Skarlatos, Adr {myan8, jchoi42, skarlat2}@`

University of Michigan CleanupSpec: An "Undo" Approach to Safe Speculation Gururaj Saileshwar Moinuddin K. Qureshi Moinuddin K. Qureshi

Georgia Institute of Teck



Speculative Taint Tracking (STT): A Comprehensive Protection for Speculatively Accessed Data

Mengjia Yan University of Illinois at Urbana-Champaign myan8@illinois.edu

Josep Torrellas

Artem Khyzha Tel Aviv University artkhyzha@mail.tau.ac.il

Christopher W. Fletcher

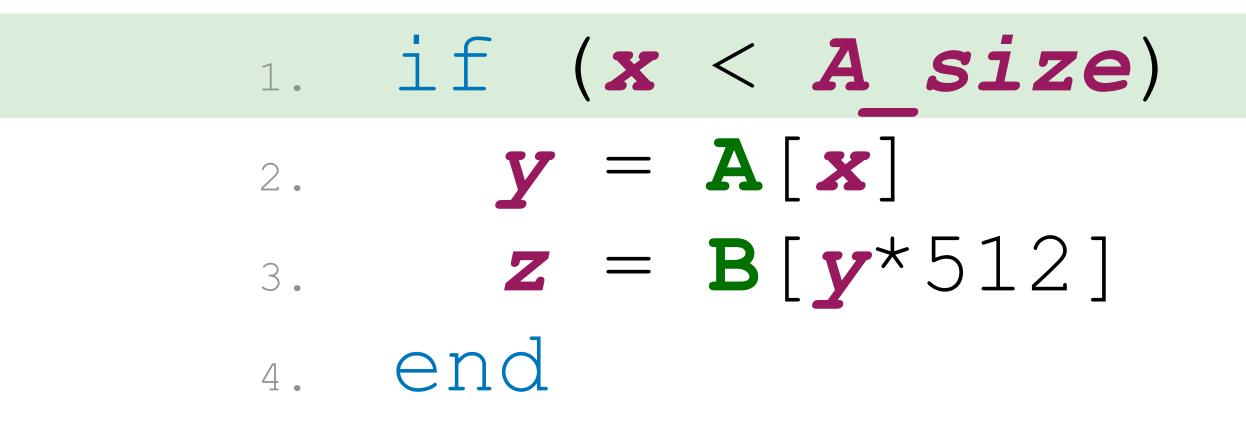








if (x < A_size) y = A[x] z = B[y*512] end



17

if (x < A_size) y = A[x] z = B[y*512] end

if (x < A size) y = A[x] z = B[y*512] end

Delay loads until they can be retired [Sakalis et al., ISCA'19]

Delay loads until they cannot be squashed [Sakalis et al., ISCA'19]





if (x < A size) y = A[x] z = B[y*512] end

Delay loads until they can be retired [Sakalis et al., ISCA'19]

Delay loads until they cannot be squashed [Sakalis et al., ISCA'19]

Taint speculatively loaded data + delay tainted loads [STT and NDA, MICRO'19]







y = A[x] if (x < A_size) z = B[y*512] end

y = A[x] if (x < A_size) z = B[y*512] end

Delay loads until they can be retired [Sakalis et al., ISCA'19]

Delay loads until they cannot be squashed [Sakalis et al., ISCA'19]





y = A[x] if (x < A size) z = B[y*512] end

Delay loads until they can be retired [Sakalis et al., ISCA'19]

Delay loads until they cannot be squashed [Sakalis et al., ISCA'19]

Taint speculatively loaded data + delay tainted loads [STT and NDA, MICRO'19]









What security properties do HW countermeasures enforce?

How can we program securely?



A Proof of Concept

M. Guarnieri, B. Köpf, J. Reineke, and P. Vila Hardware–Software Contracts for Secure Speculation S&P (Oakland) 2021



Hardware-Software Contracts

HW/SW Contracts for Secure Speculation

HW/SW Contracts for Secure Speculation

Hardware Countermeasures

Load Delay Taint Tracking

No speculation

No countermeasures

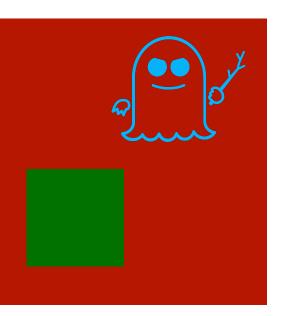
22

Secure Programming

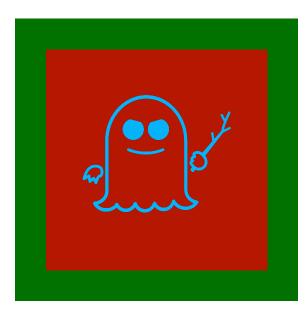
HW/SW Contracts for Secure Speculation

Hardware Countermeasures

Load Delay



Constant-time



Sandboxing

No speculation

Taint Tracking

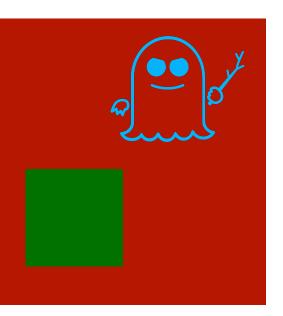
No countermeasures

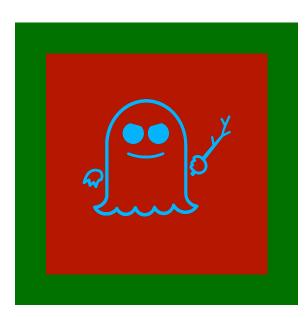
Secure Programming

HW/SW Contracts for Secure Speculation

Hardware Countermeasures

Load Delay





Constant-time

Sandboxing

mechanism-independent simple Desiderata: precise

No speculation

Taint Tracking

No countermeasures



Instruction Set Architecture Arch. states: σ Arch. semantics: $\sigma \rightsquigarrow \sigma'$

Instruction Set Architecture Arch. states: σ Arch. semantics: $\sigma \rightsquigarrow \sigma'$

> Microarchitecture Hardware states: $\langle \sigma, \mu \rangle$

Hardware semantics: $\langle \sigma, \mu \rangle \Rightarrow \langle \sigma', \mu' \rangle$

Instruction Set Architecture Arch. states: σ Arch. semantics: $\sigma \rightsquigarrow \sigma'$

> Microarchitecture Hardware states: $\langle \sigma, \mu \rangle$

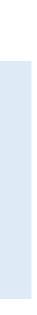
Hardware semantics: $\langle \sigma, \mu \rangle \Rightarrow \langle \sigma', \mu' \rangle$

Adversary model μ Arch traces: $\{p\}(\sigma) = \mu_0 \mu_1 \dots \mu_n$





A deterministic, labelled semantics $\xrightarrow{\tau}$ for the ISA





Observations expose security-relevant *µArch events* A deterministic, labelled semantics $\xrightarrow{\tau}$ for the ISA





Observations expose security-relevant *µArch events* A deterministic, labelled semantics $\stackrel{\tau}{\rightarrow}$ for the ISA

Contract traces: $[p](\sigma) = \tau_1 \tau_2 \dots \tau_n$







A deterministic, labelled semantics $\stackrel{\tau}{\rightarrow}$ for the ISA

Contract satisfaction Hardware $\{\cdot\}$ satisfies contract $[\cdot]$ if for all programs p and

Observations expose security-relevant *µArch events*

Contract traces: $[p](\sigma) = \tau_1 \tau_2 \dots \tau_n$

arch. states σ, σ' : if $[p](\sigma) = [p](\sigma')$ then $\{p\}(\sigma) = \{p\}(\sigma')$





Contract = Execution Mode · Observer Mode

Contract = Execution Mode · Observer Mode

How are programs executed?

Contract = Execution Mode · Observer Mode

How are programs executed?

What is visible about the execution?

Contract = Execution Mode · Observer Mode

Contract = Execution Mode · Observer Mode

seq — sequential execution **spec** — mispredict branch instructions

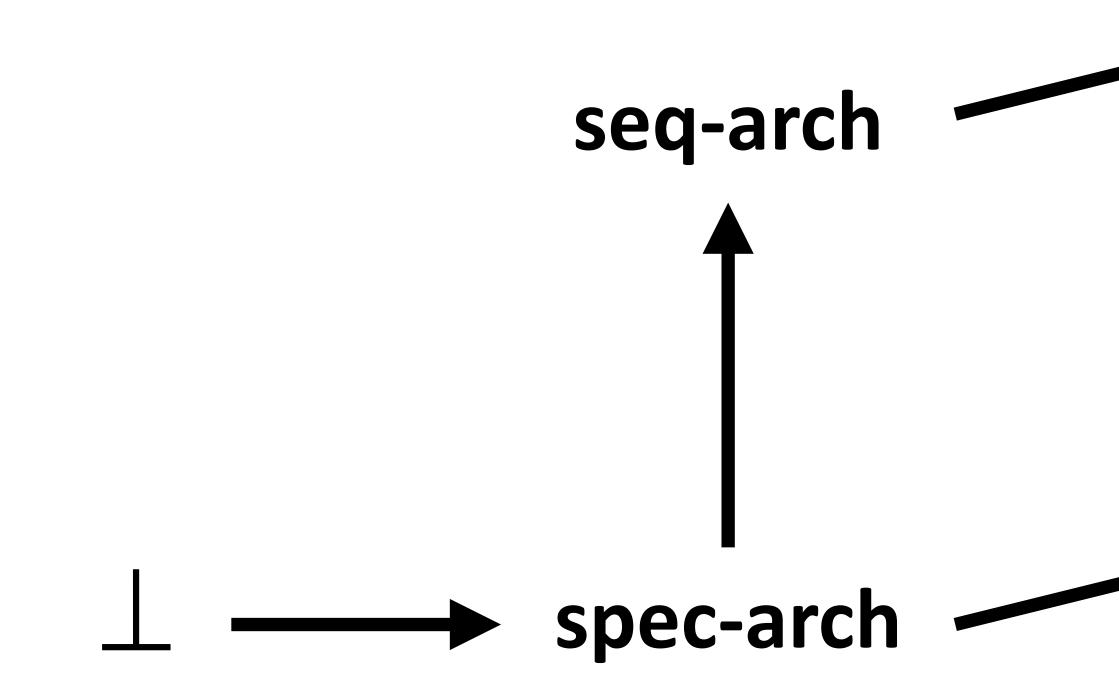
Contract = Execution Mode · Observer Mode

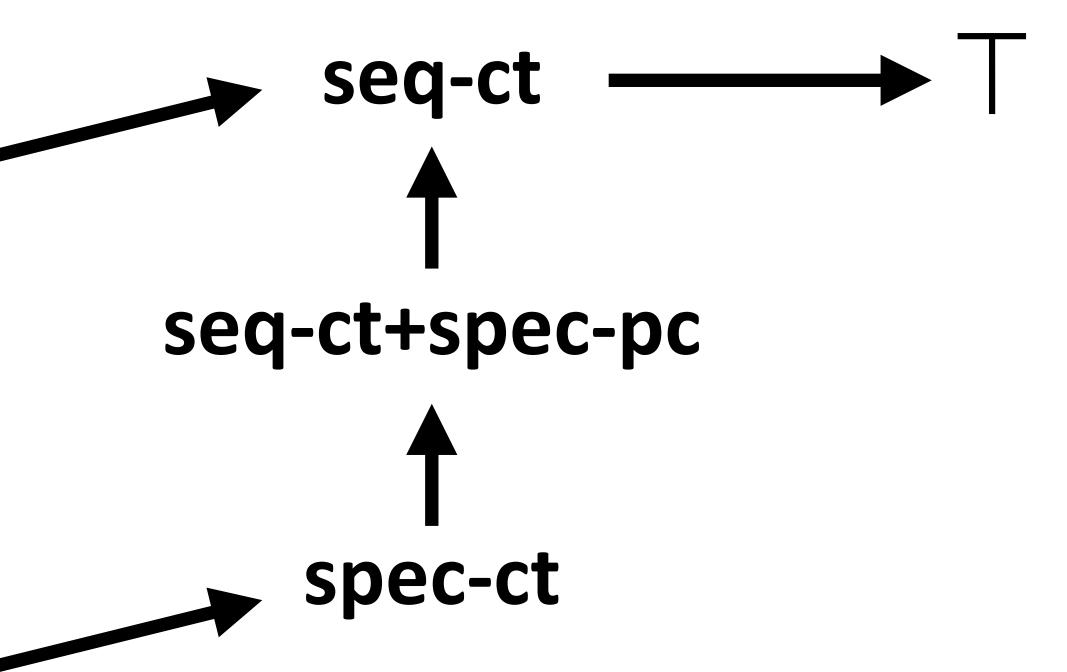
Contract = Execution Mode · Observer Mode

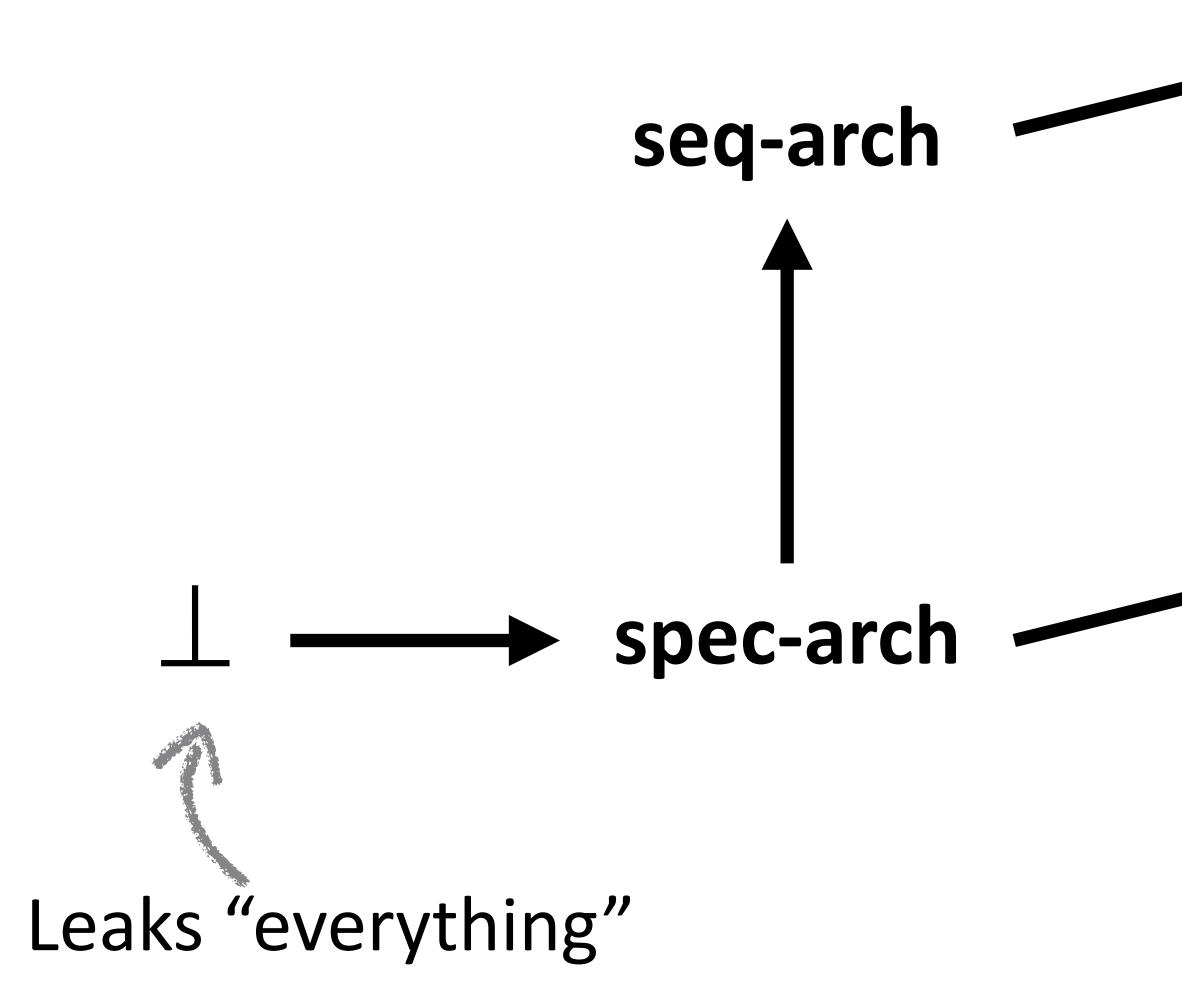
pc — only program counter

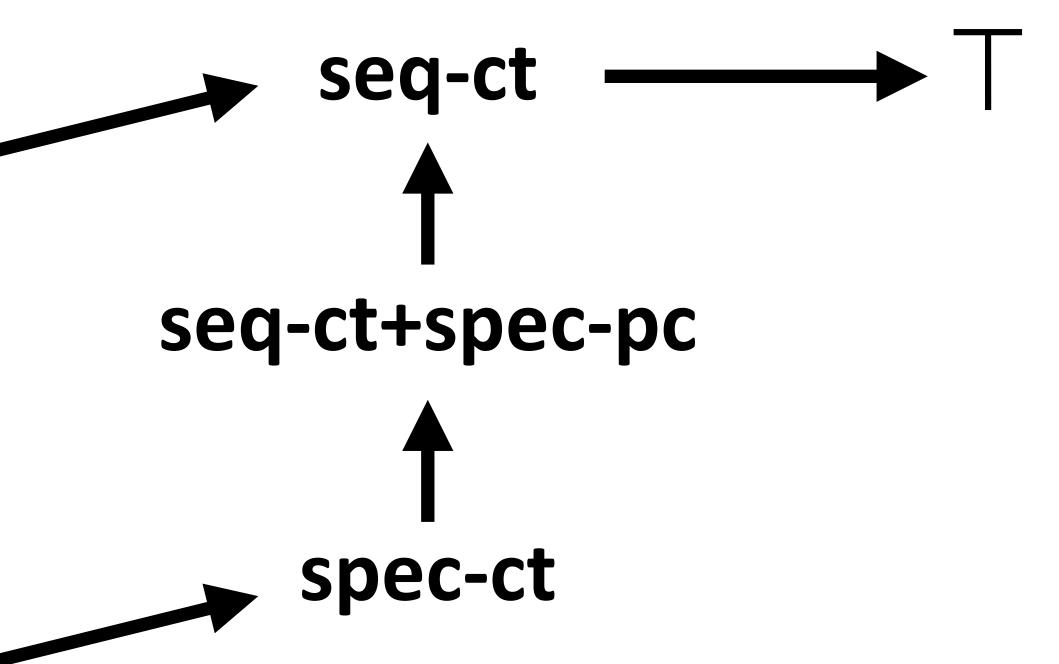
ct — pc + addr. of loads and stores

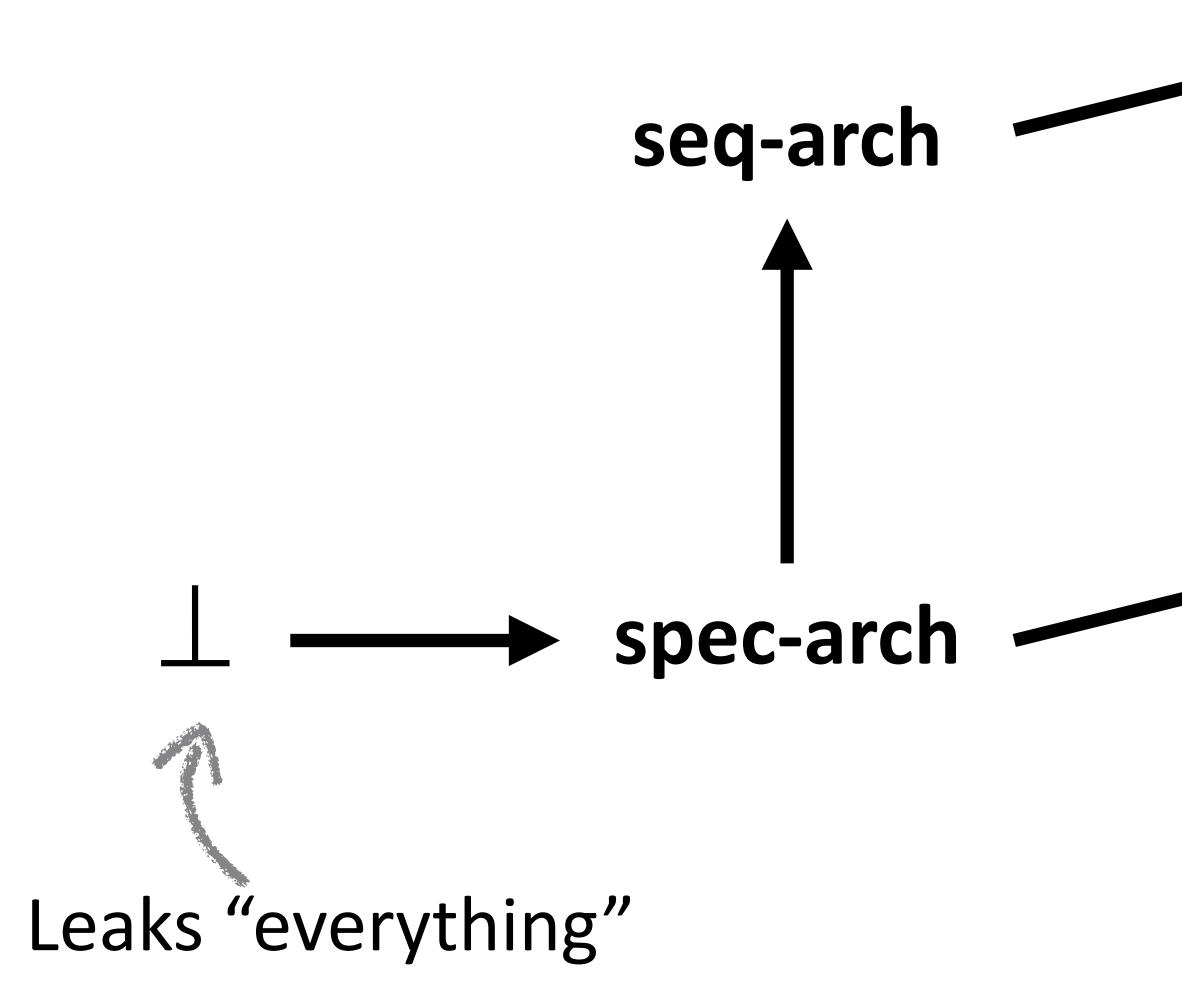
arch — ct + loaded values

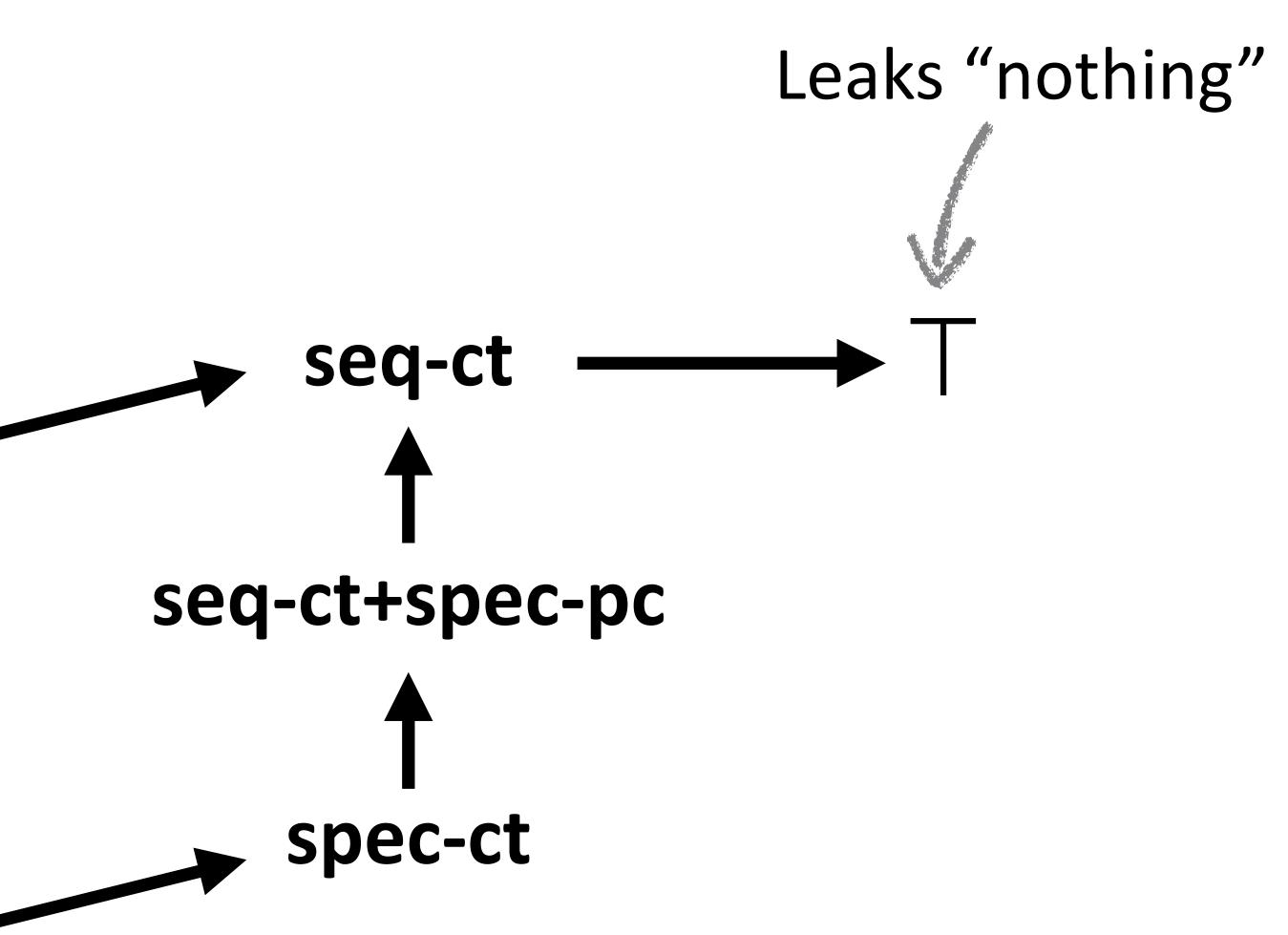


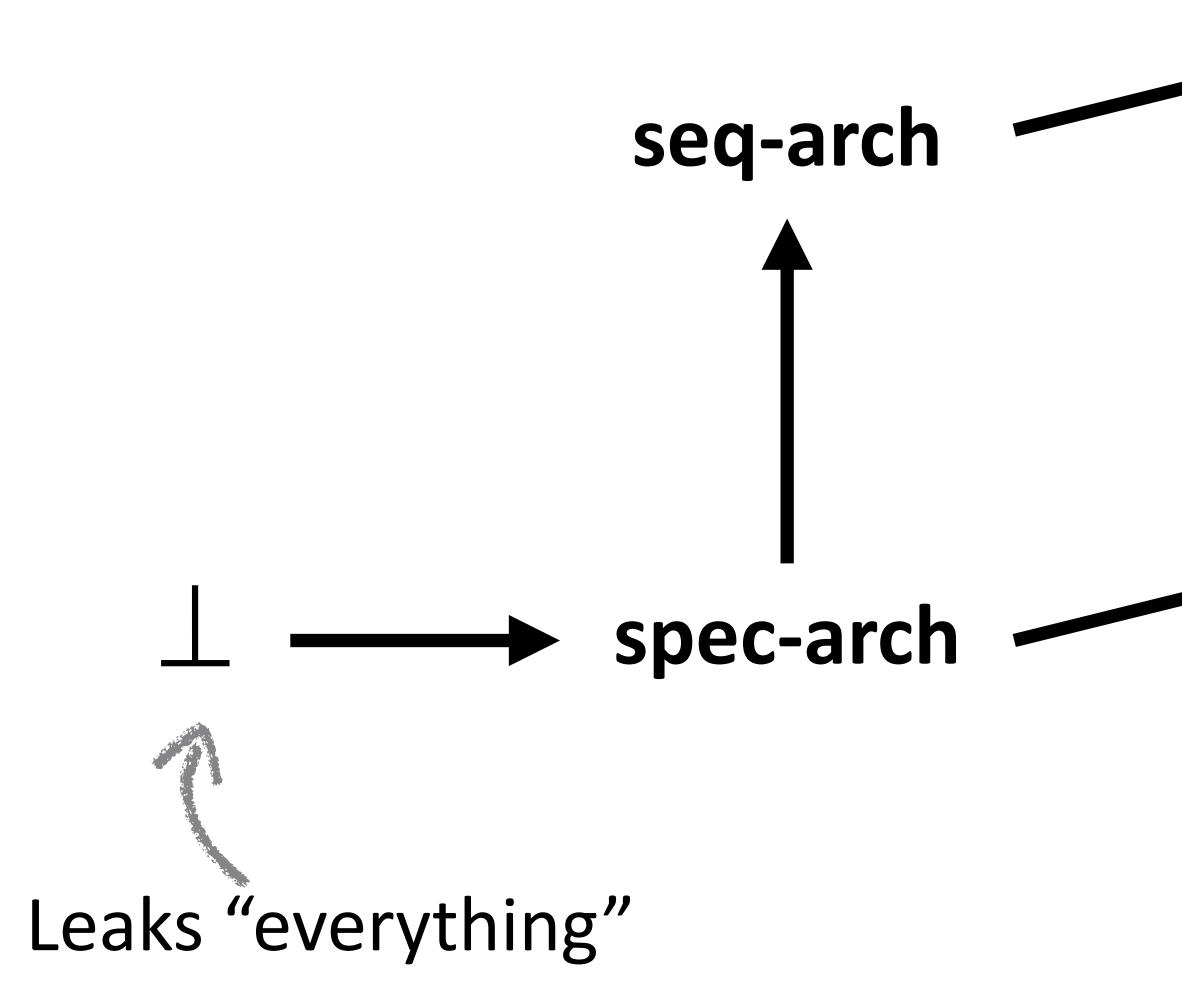


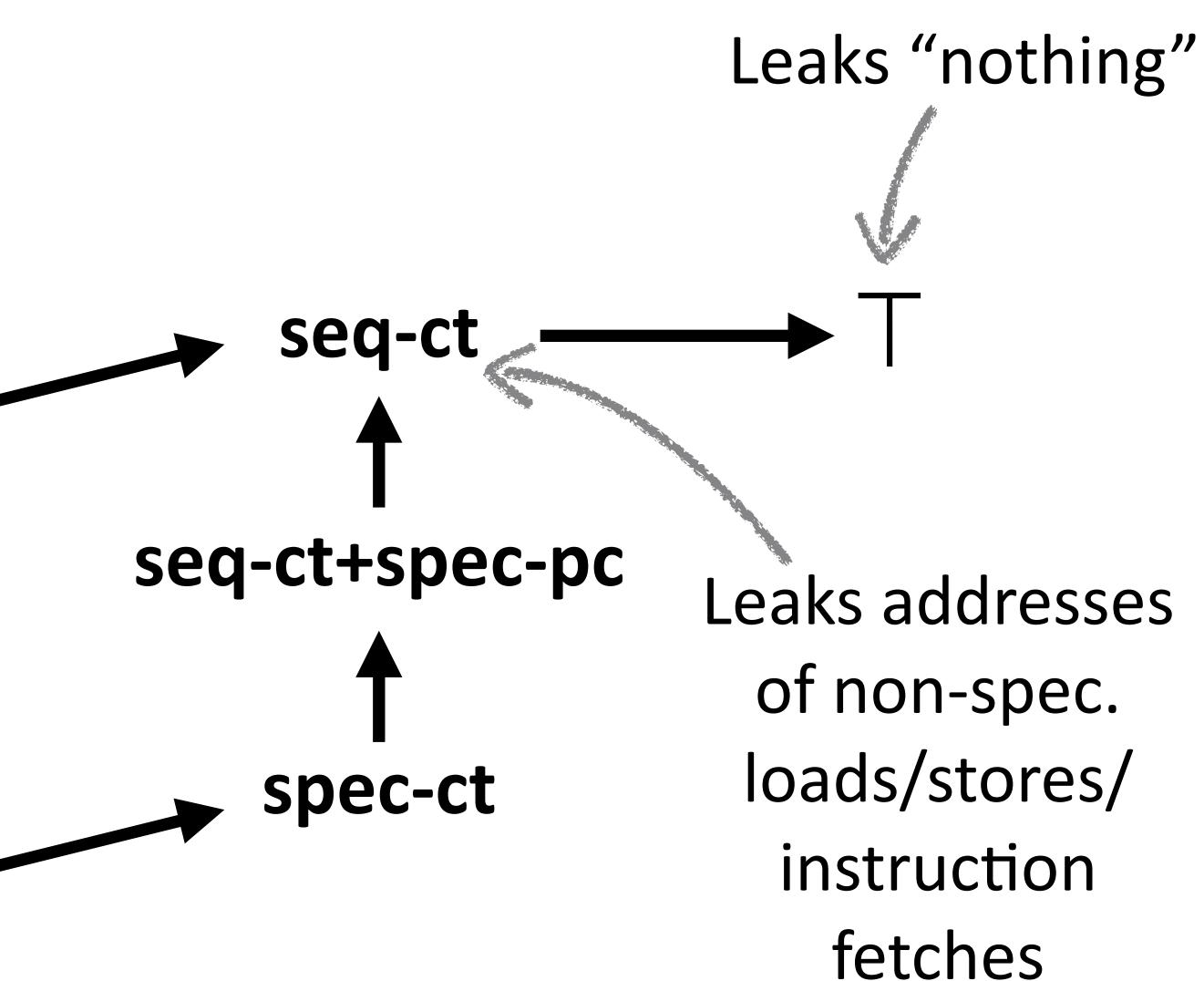




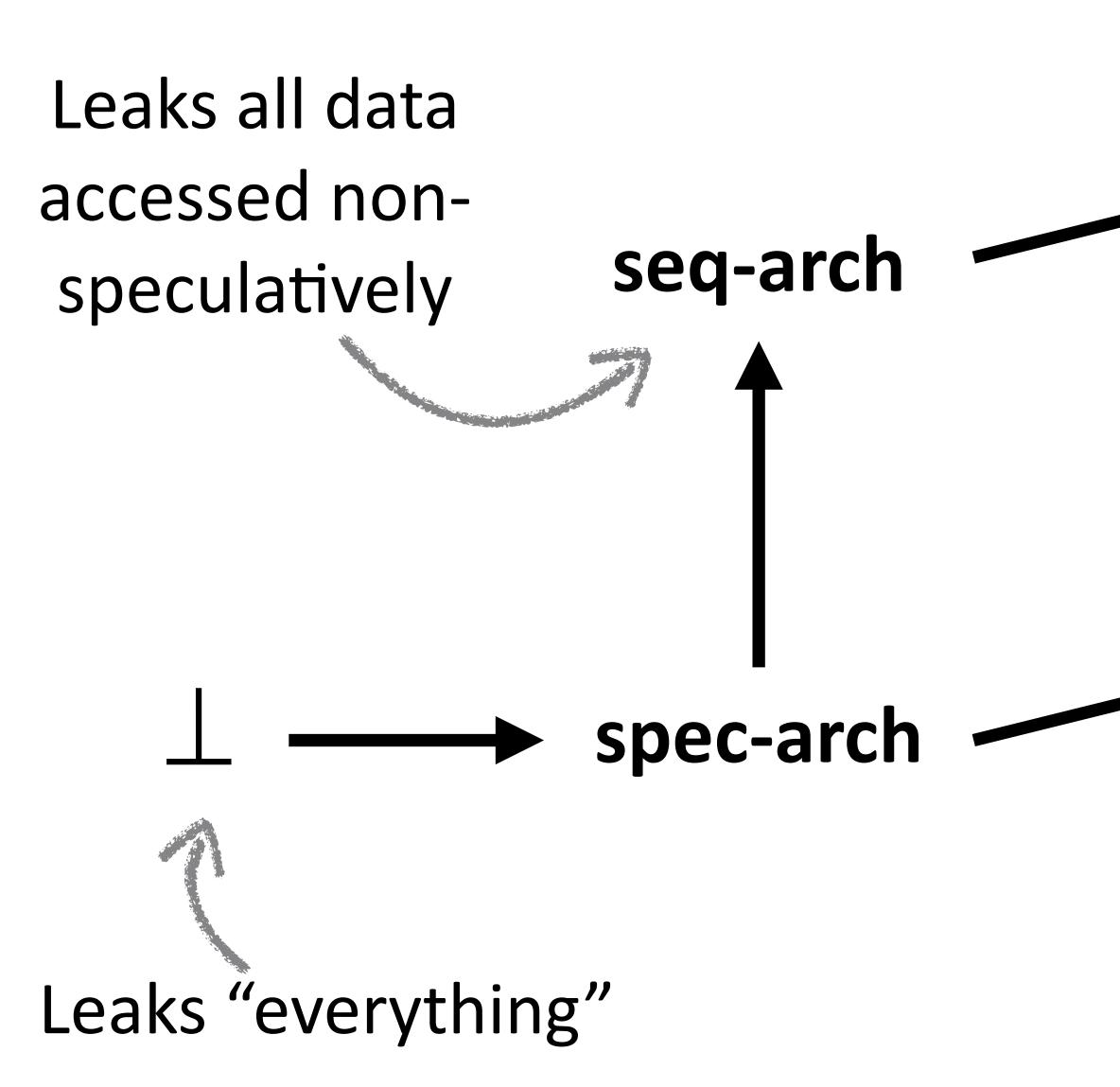


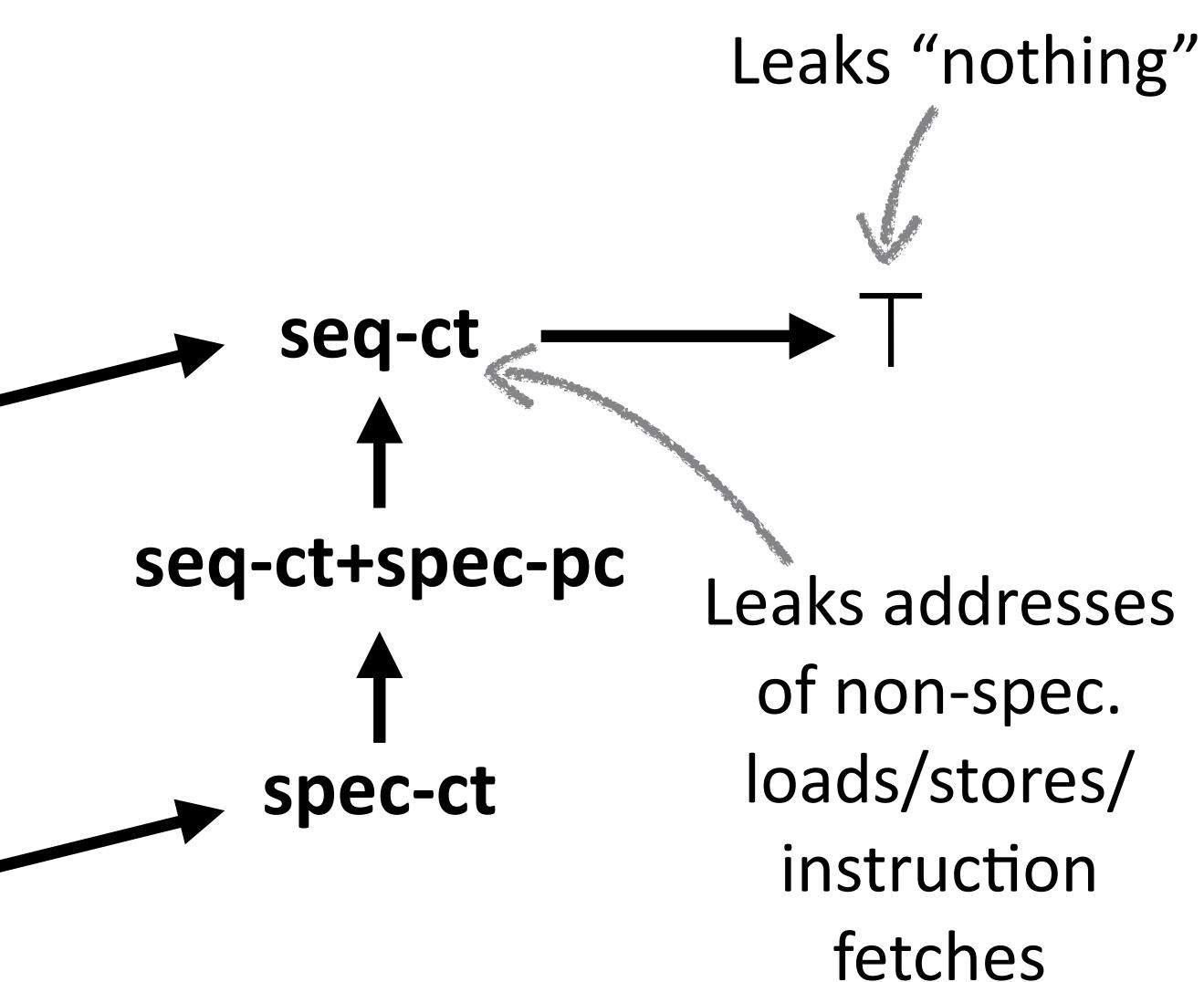


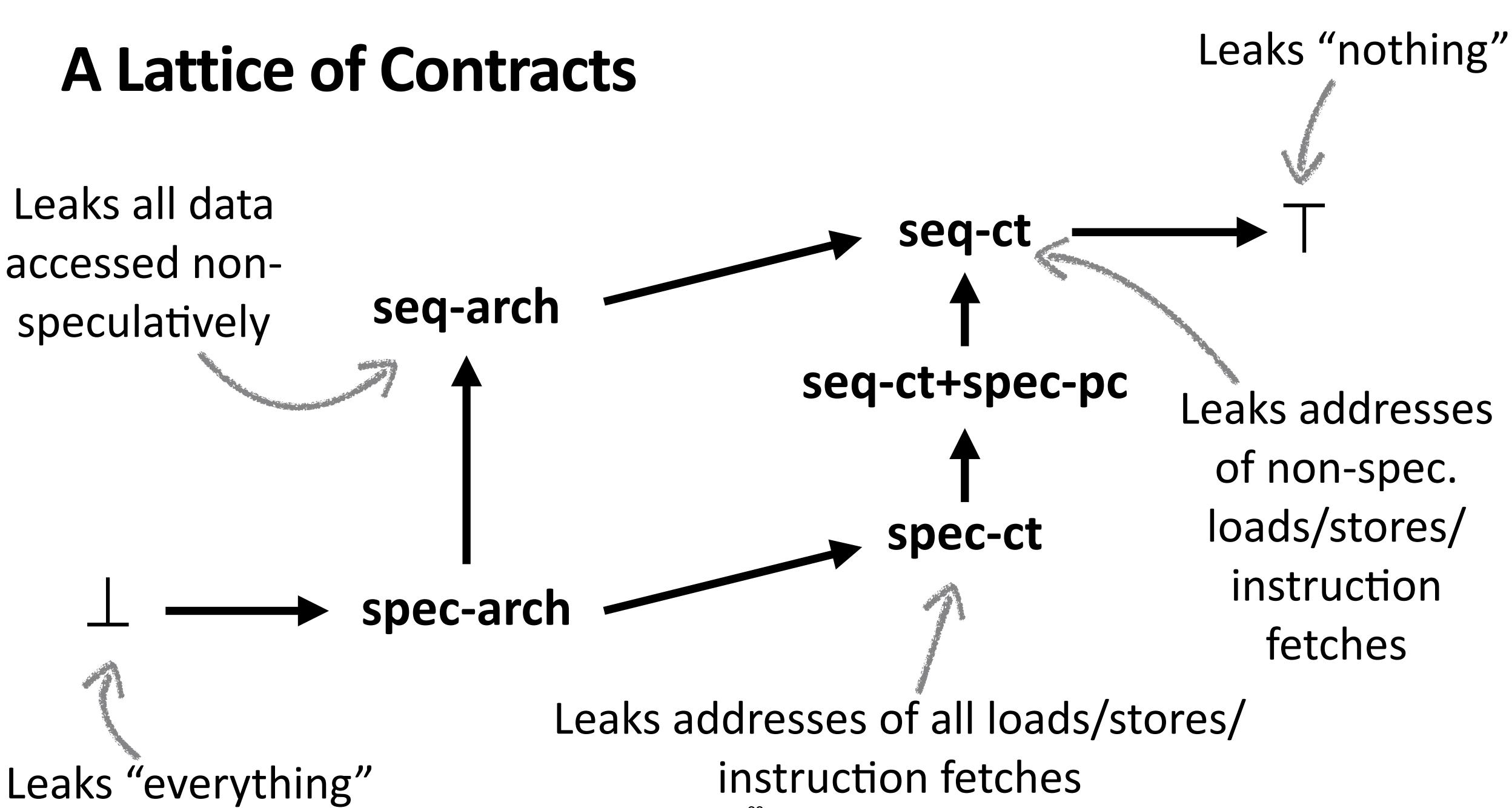




A Lattice of Contracts

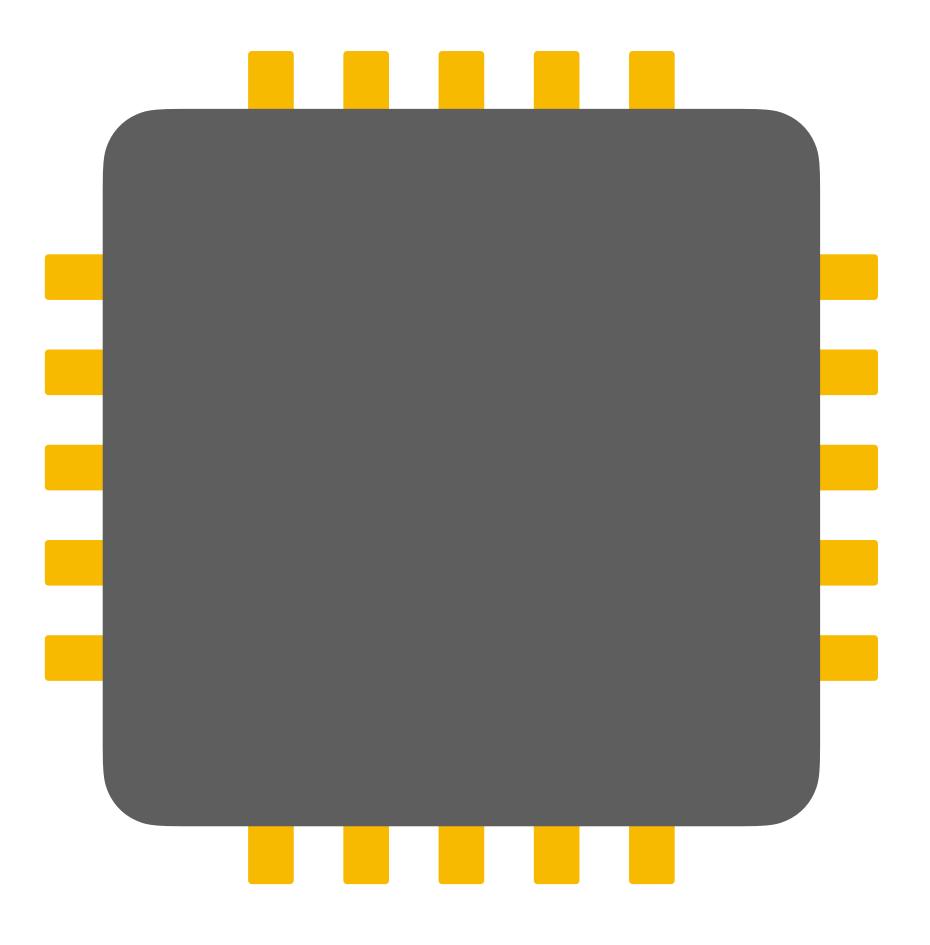


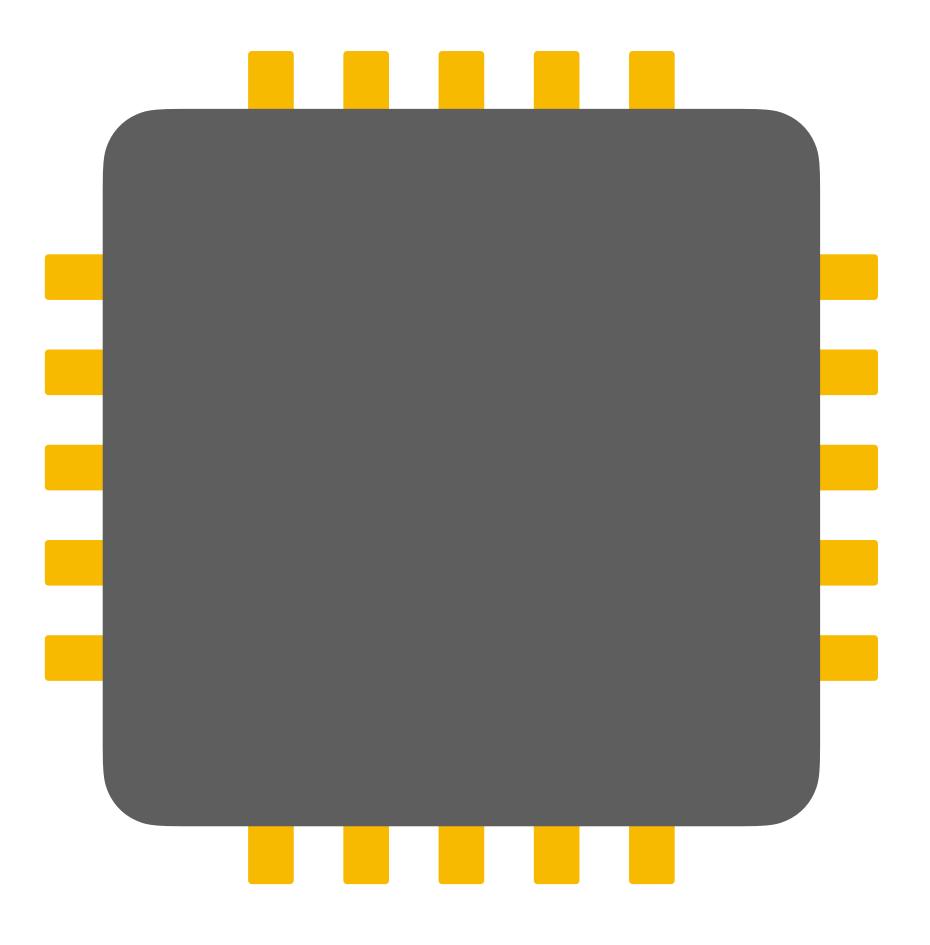




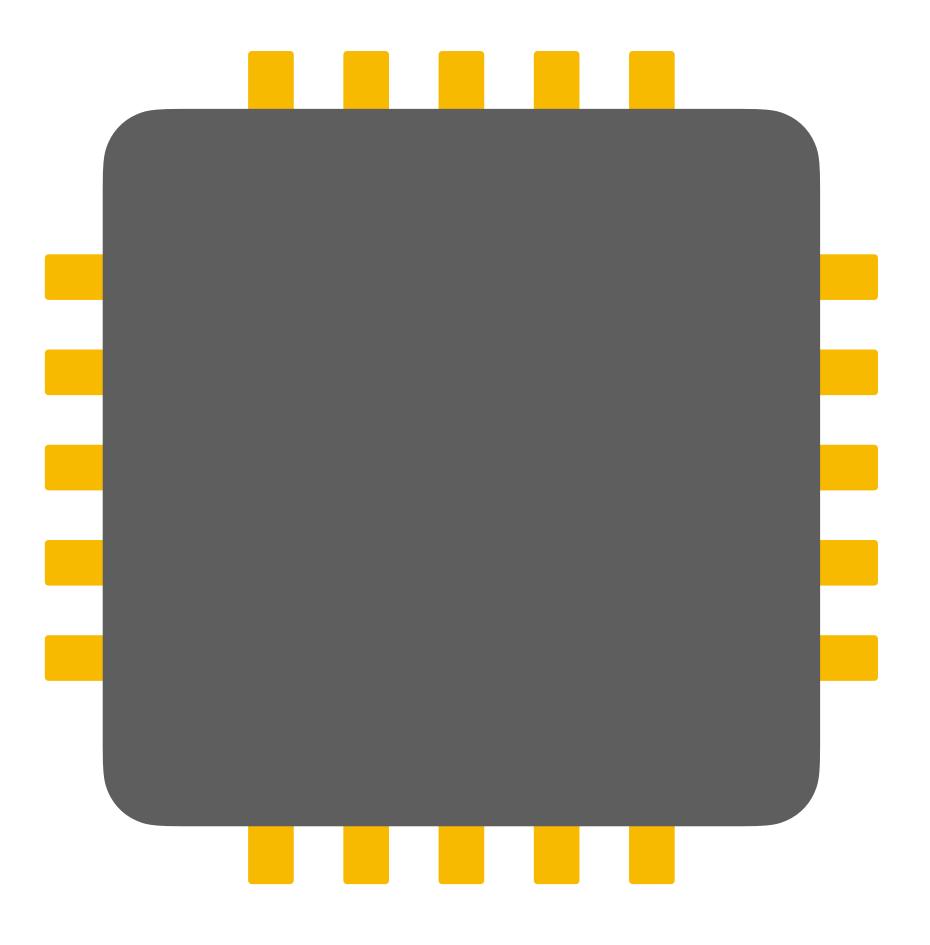


Hardware Countermeasures



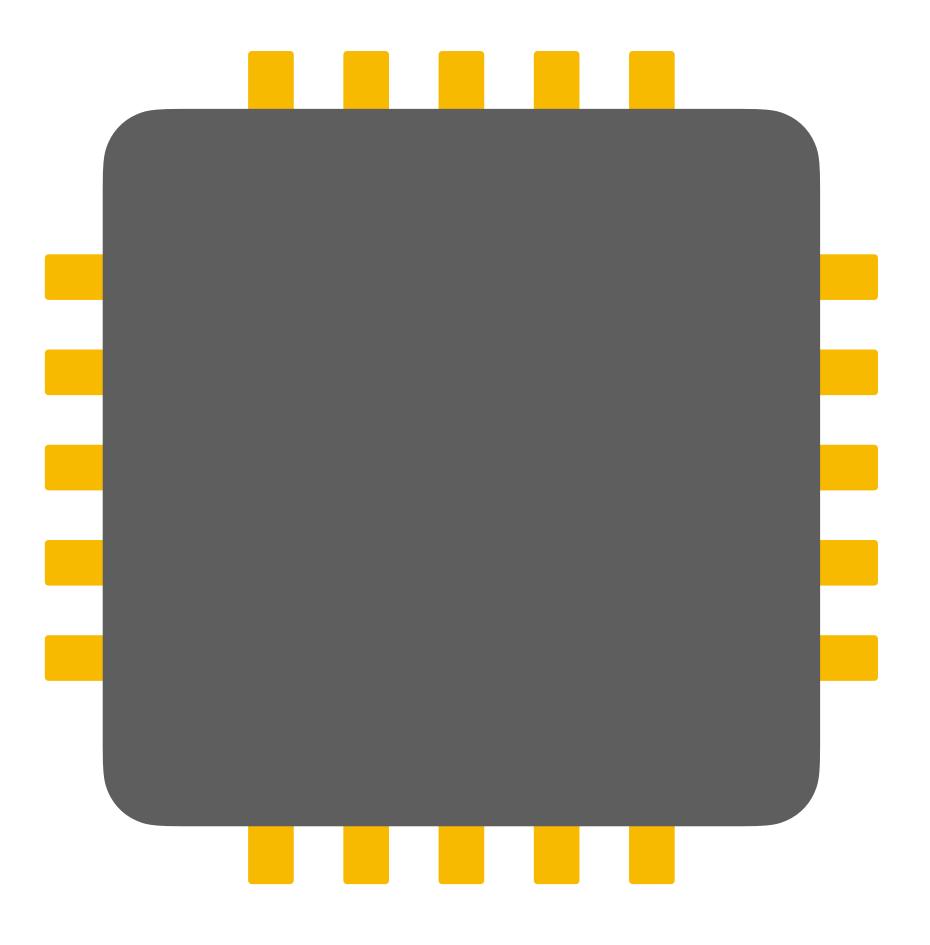


3-stage pipeline (fetch, execute, retire)



3-stage pipeline (fetch, execute, retire)

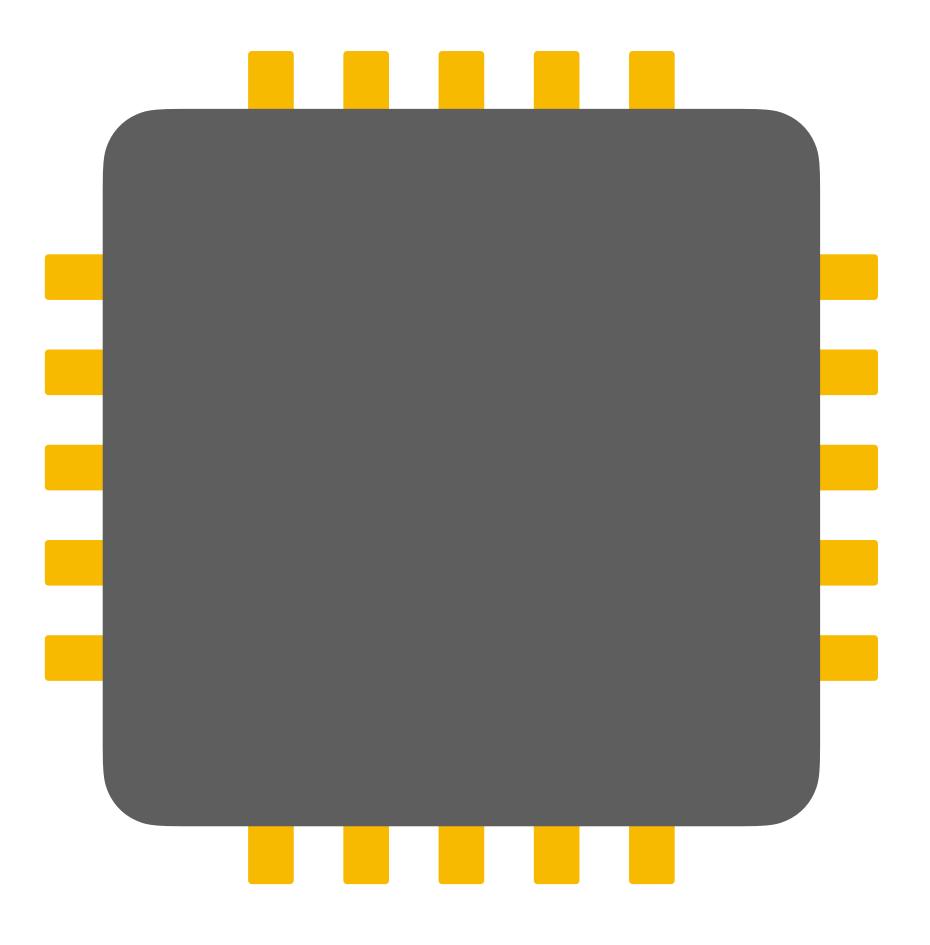
Speculative and out-of-order execution



3-stage pipeline (fetch, execute, retire)

Speculative and out-of-order execution

Parametric in *branch predictor* and *memory hierarchy*

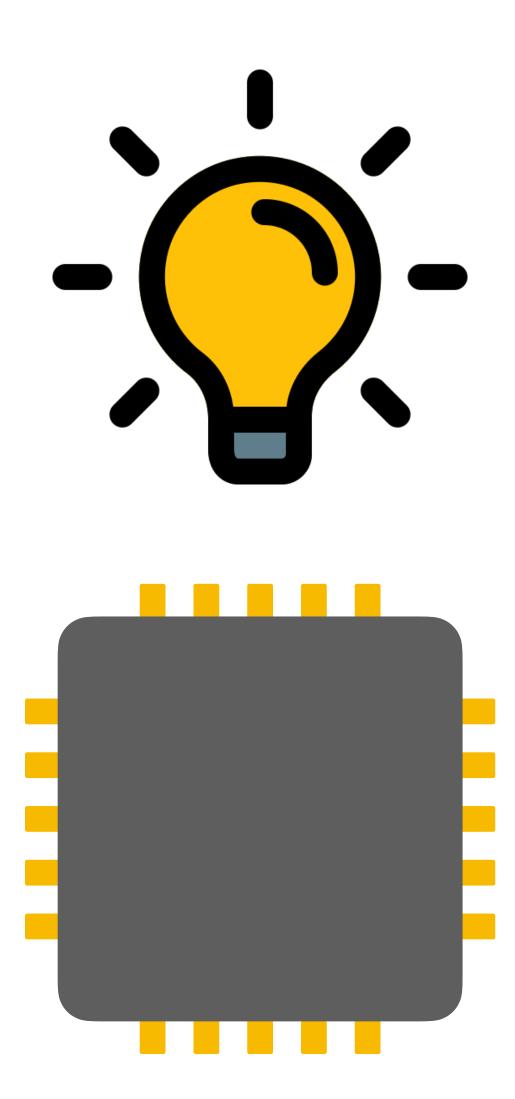


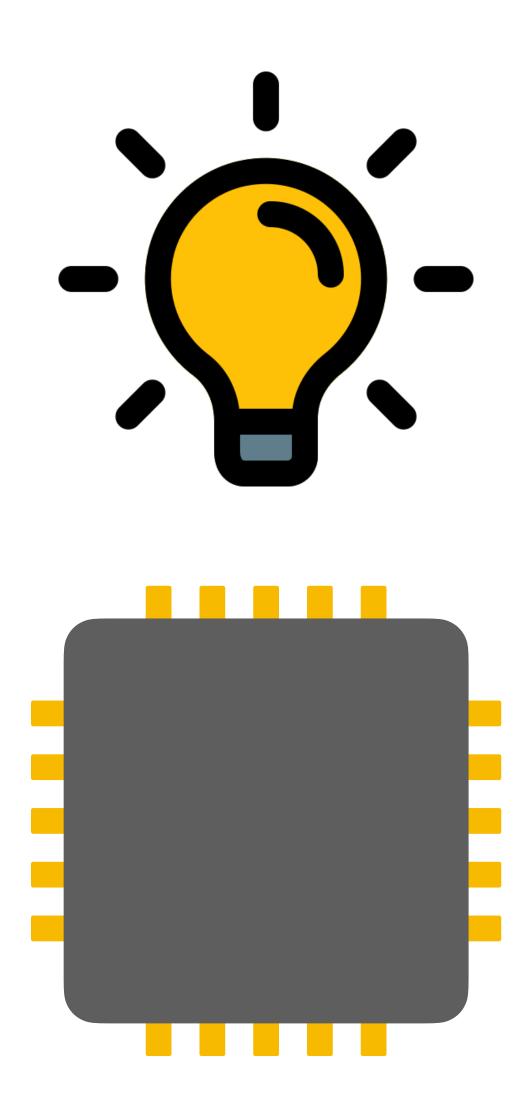
3-stage pipeline (fetch, execute, retire)

Speculative and out-of-order execution

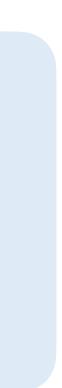
Parametric in *branch predictor* and *memory hierarchy*

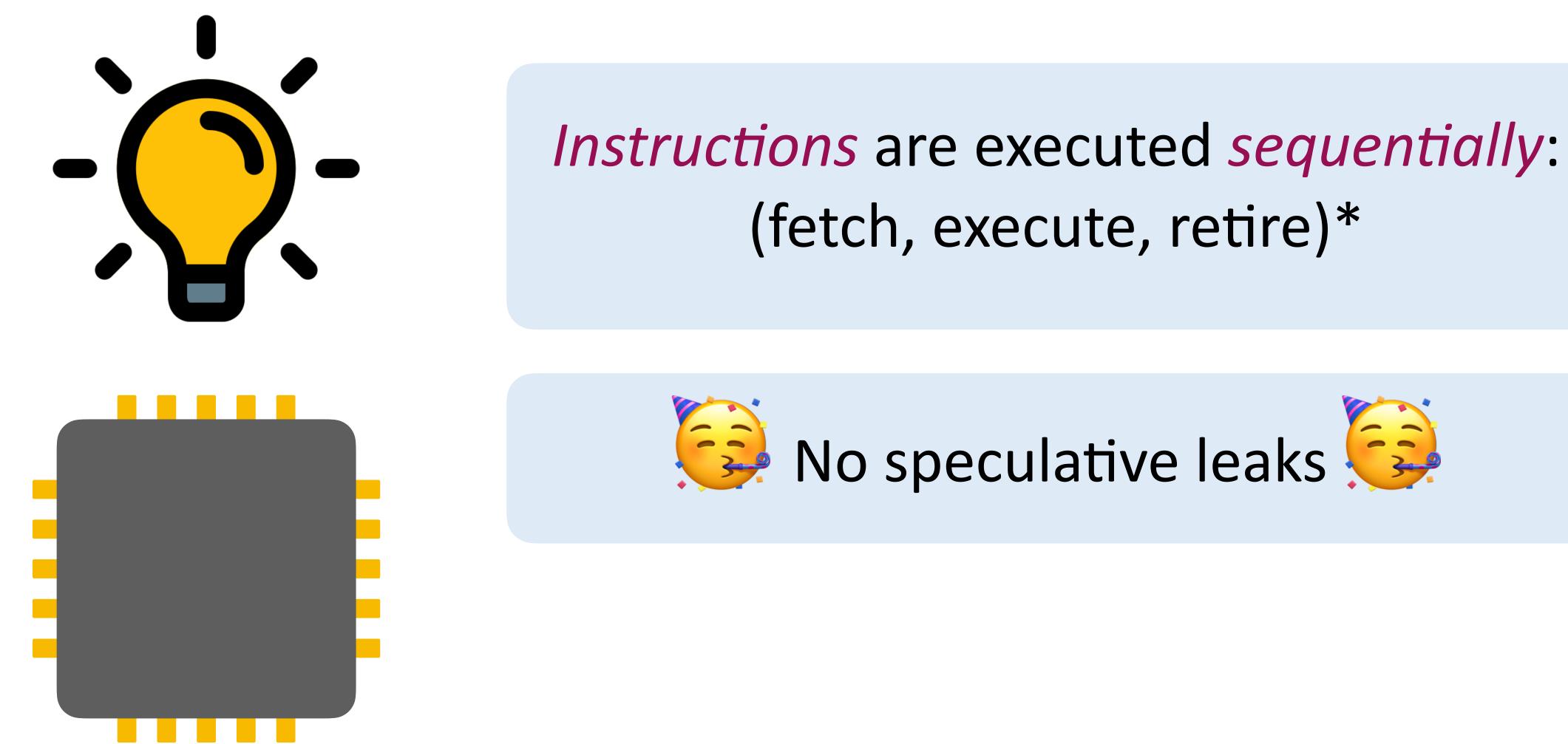
Different *schedulers* for different countermeasures



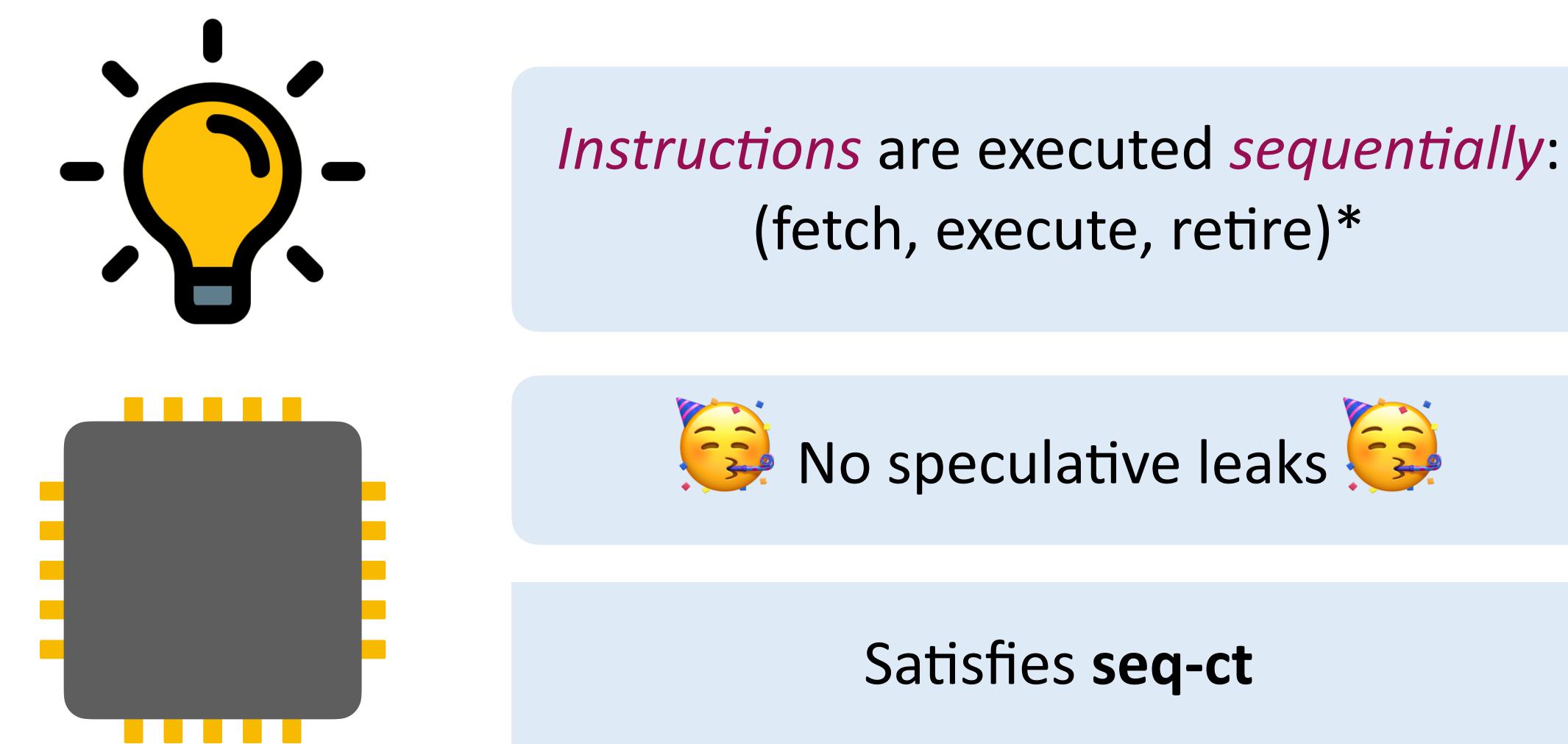


Instructions are executed sequentially: (fetch, execute, retire)*

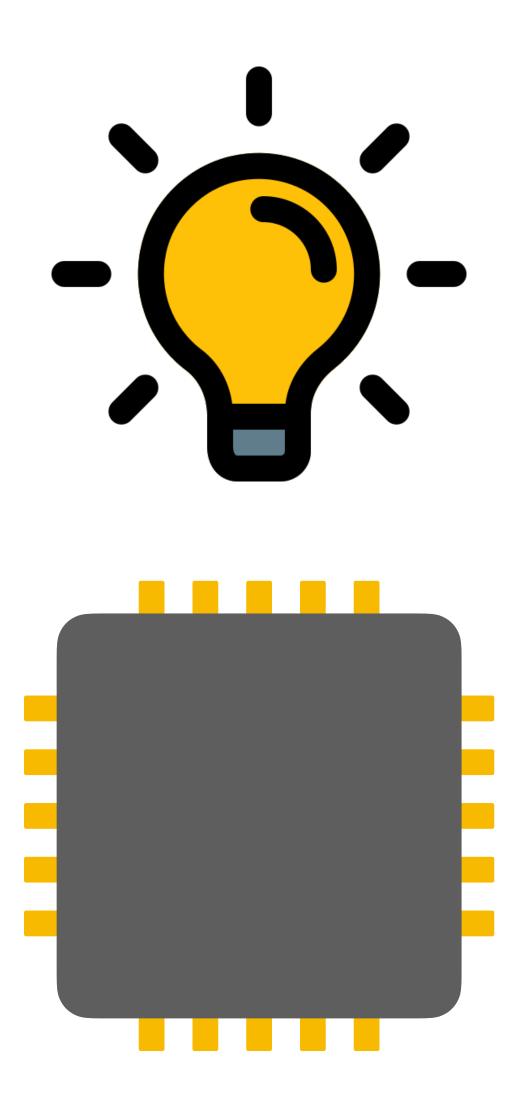


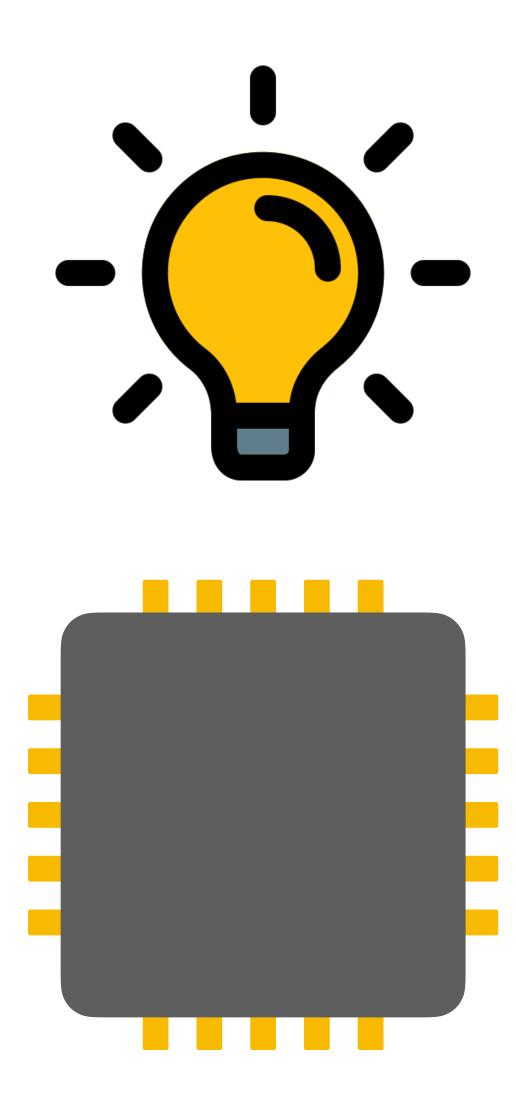




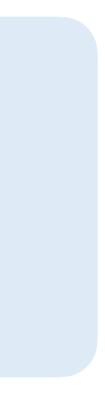


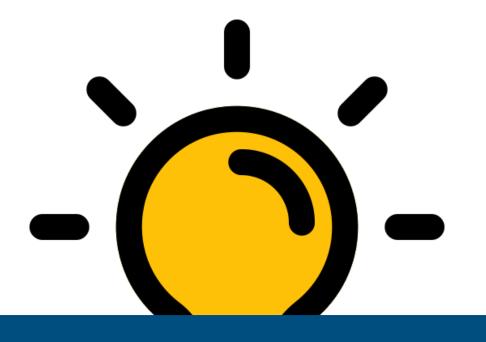






Delaying loads until all sources of speculation are resolved







Security guarantees?

if (x < A_size) z = A[x] y = B[z]</pre>

if (x < A_size) z = A[x] y = B[z]</pre>

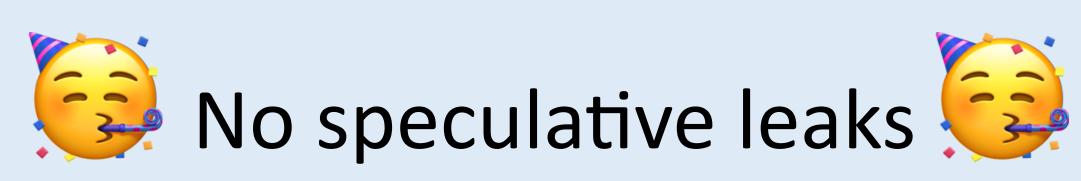
if (x < A_size) z = A[x] y = B[z]</pre>

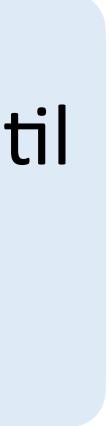
A[x] and B[z] delayed until x < A size is resolved</pre>



if (x < A_size) z = A[x] y = B[z]</pre>

A[x] and B[z] delayed until x < A size is resolved</pre>

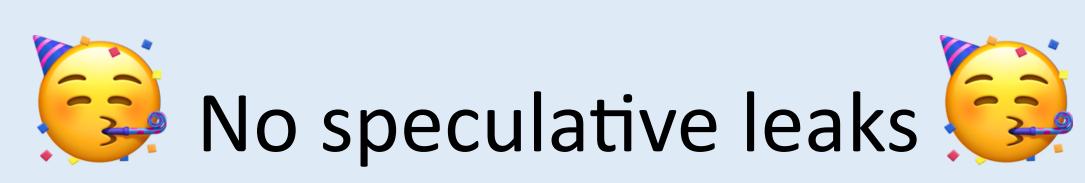


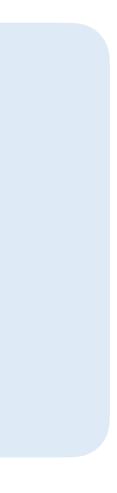




z = A[x] if (x < A_size) y = B[z]</pre>

B[z] delayed until x < A_size is resolved</pre>

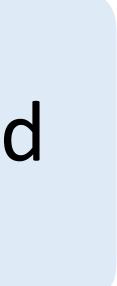






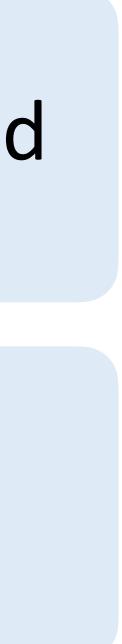
z = A[x] if (x < A_size) if (z==0) skip</pre>

if (z=0) is not delayed



if (z==0) is not delayed

Program speculatively leaks **A**[**x**]



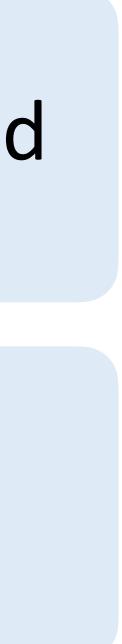
z = A[x]if (x < A size) if (z==0)skip

Observation: Can only leak data accessed non-speculatively

if (z==0) is not delayed

Program speculatively leaks A[x]





z = A[x] if (x < A size) if (z==0) skip</pre>

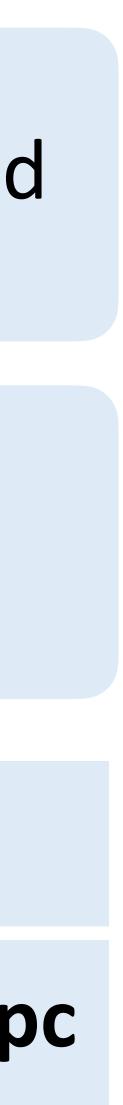
Observation: Can only leak data accessed non-speculatively

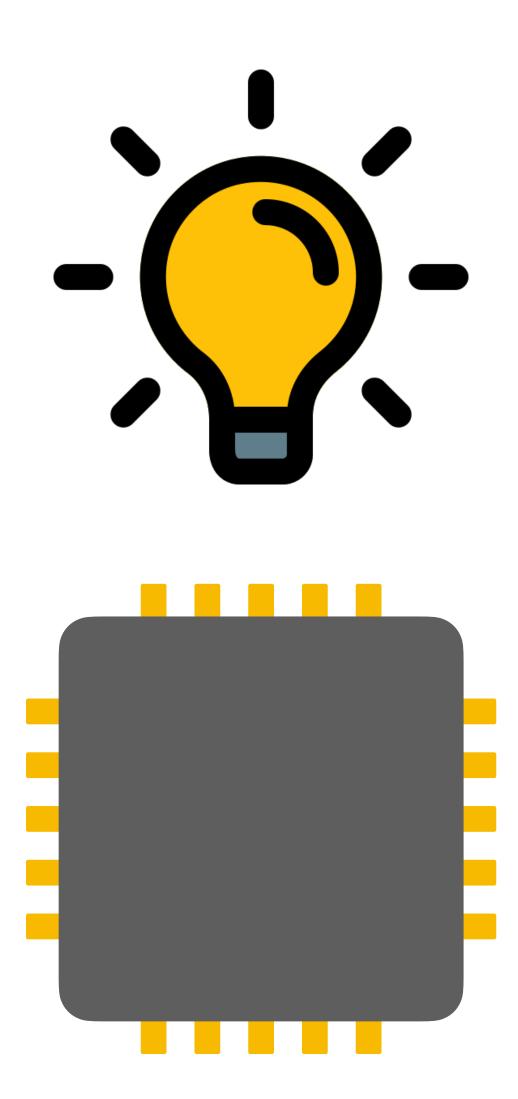
if (z==0) is not delayed

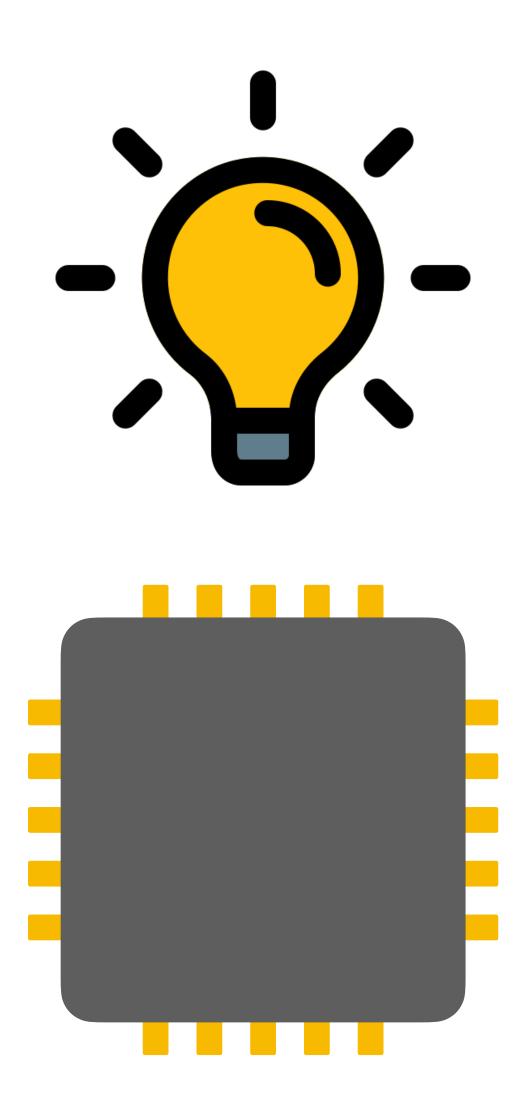
Program speculatively leaks **A**[**x**]

Satisfies seq-arch

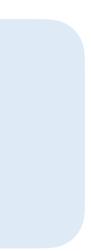
Satisfies seq-ct+spec-pc

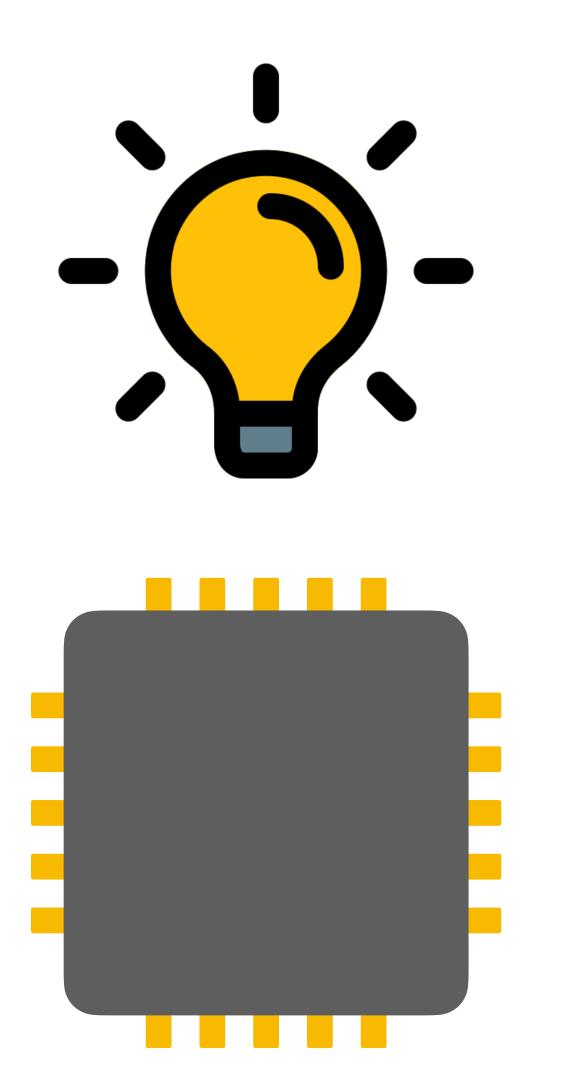






Taint speculatively loaded data

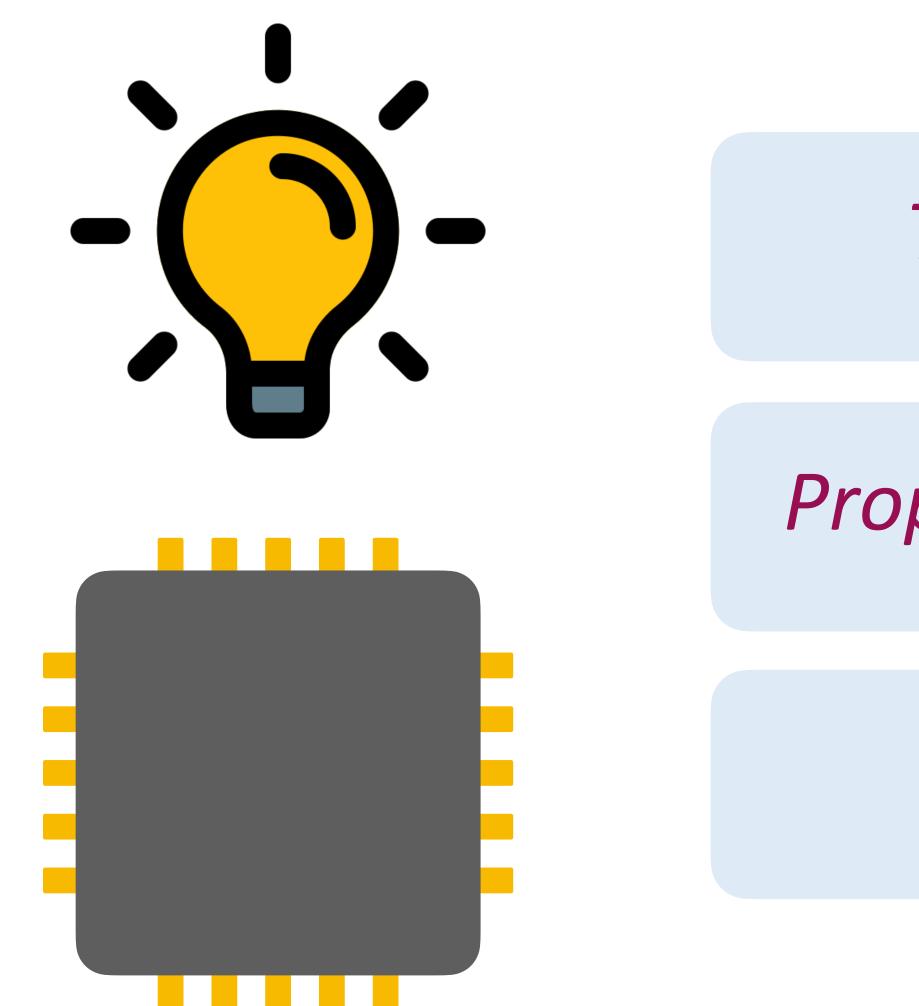




Taint speculatively loaded data

Propagate taint through computation





Taint speculatively loaded data

Propagate taint through computation

Delay tainted operations







Taint speculatively loaded data

Security guarantees?

Delay tainted operations

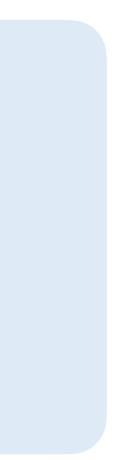


if (x < A_size) z = A[x] y = B[z]</pre>

if (x < A_size) z = A[x] y = B[z]

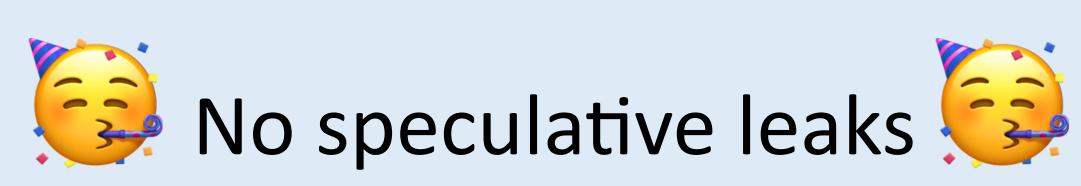
if (x < A_size) z = A[x] y = B[z]

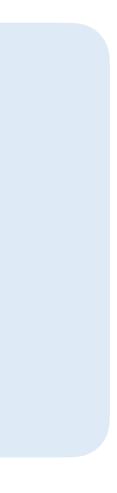
A [x] tainted as unsafe
B [z] delayed until
A [x] is safe



if (x < A_size) z = A[x] y = B[z]

A [x] tainted as unsafe
B [z] delayed until
A [x] is safe





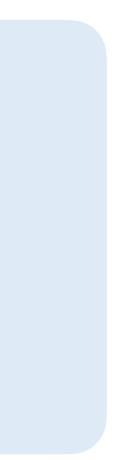


z = A[x] if (x < A_size) y = B[z]</pre>

z = A[x] if (x < A size) y = B[z]</pre>

z = A[x] if (x < A size) y = B[z]</pre>

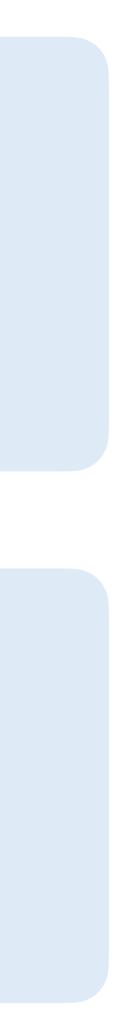
A[x] tagged as safe B[z] not delayed



z = A[x] if (x < A size) y = B[z]</pre>

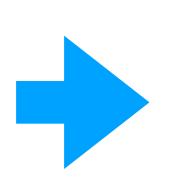
A [x] tagged as safe B [z] not delayed

Program speculatively leaks **A**[**x**]



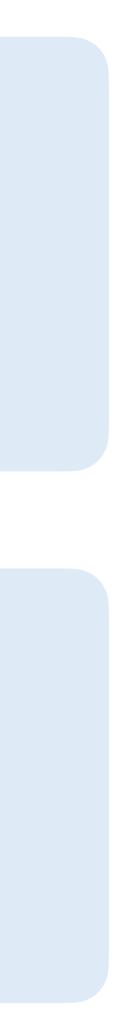
z = A[x] if (x < A size) y = B[z]</pre>

Also satisfies seq-arch



A [x] tagged as safe B [z] not delayed

Program speculatively leaks **A**[**x**]



No Countermeasures [The World until 2018]

if (x < A_size) z = A[x] y = B[z]</pre>

No Countermeasures [The World until 2018]

if (x < A_size) z = A[x] y = B[z]</pre>

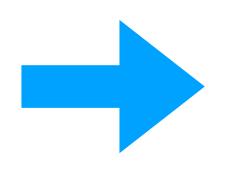
Leaks addressed of speculative and non-speculative accesses



No Countermeasures [The World until 2018]

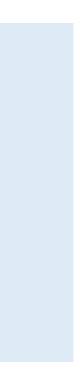
if (x < A_size) z = A[x] y = B[z]</pre>

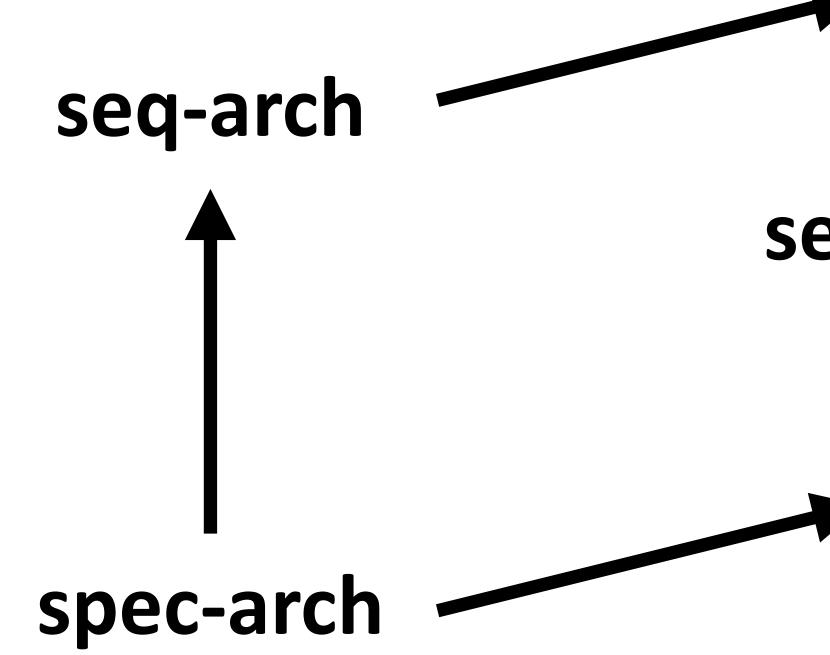
Leaks addressed of speculative and non-speculative accesses



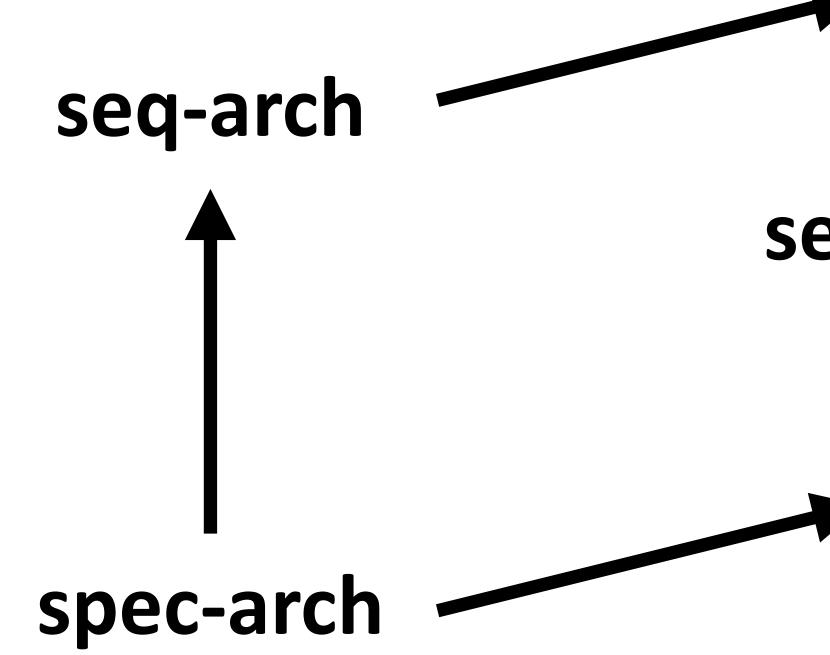
Satisfies spec-ct





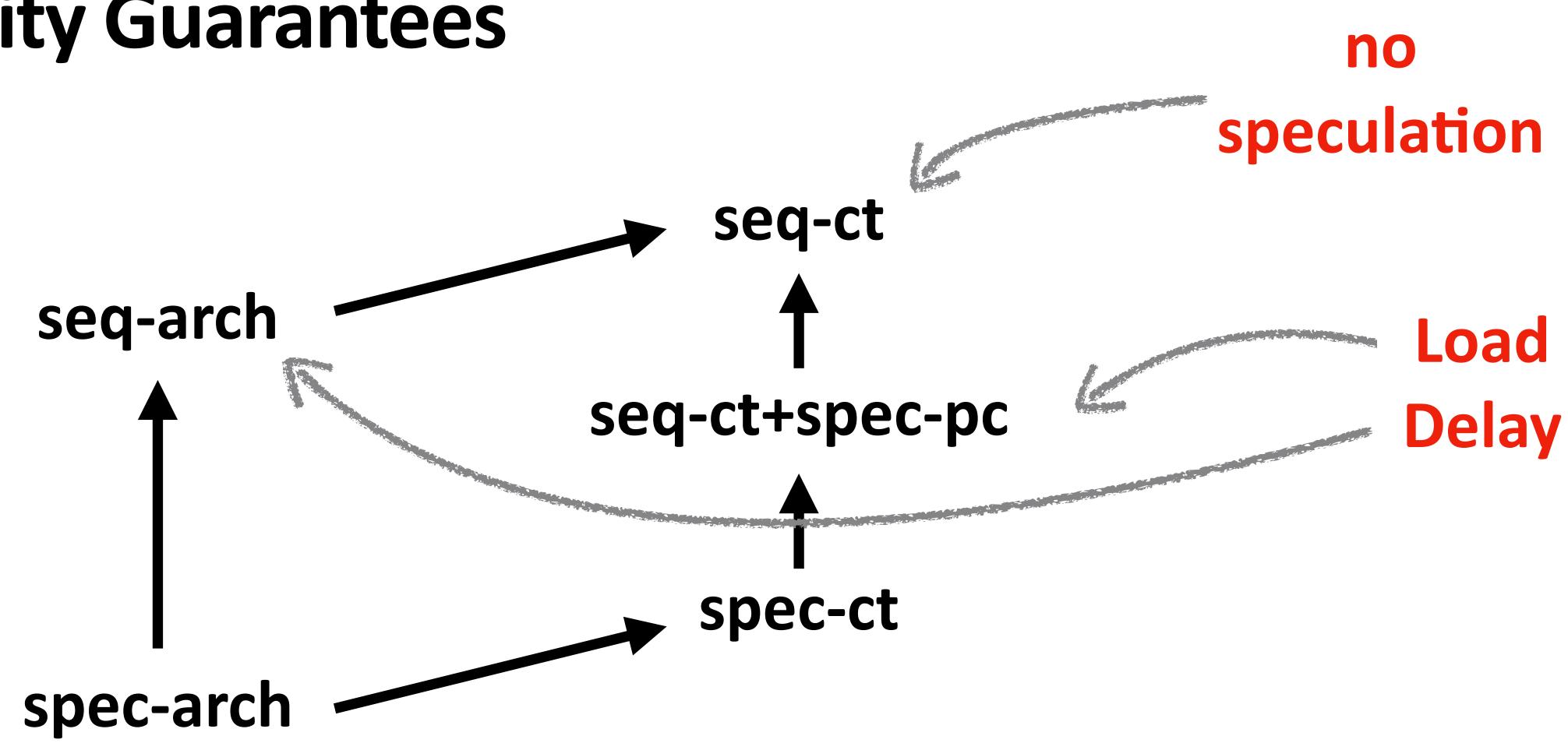


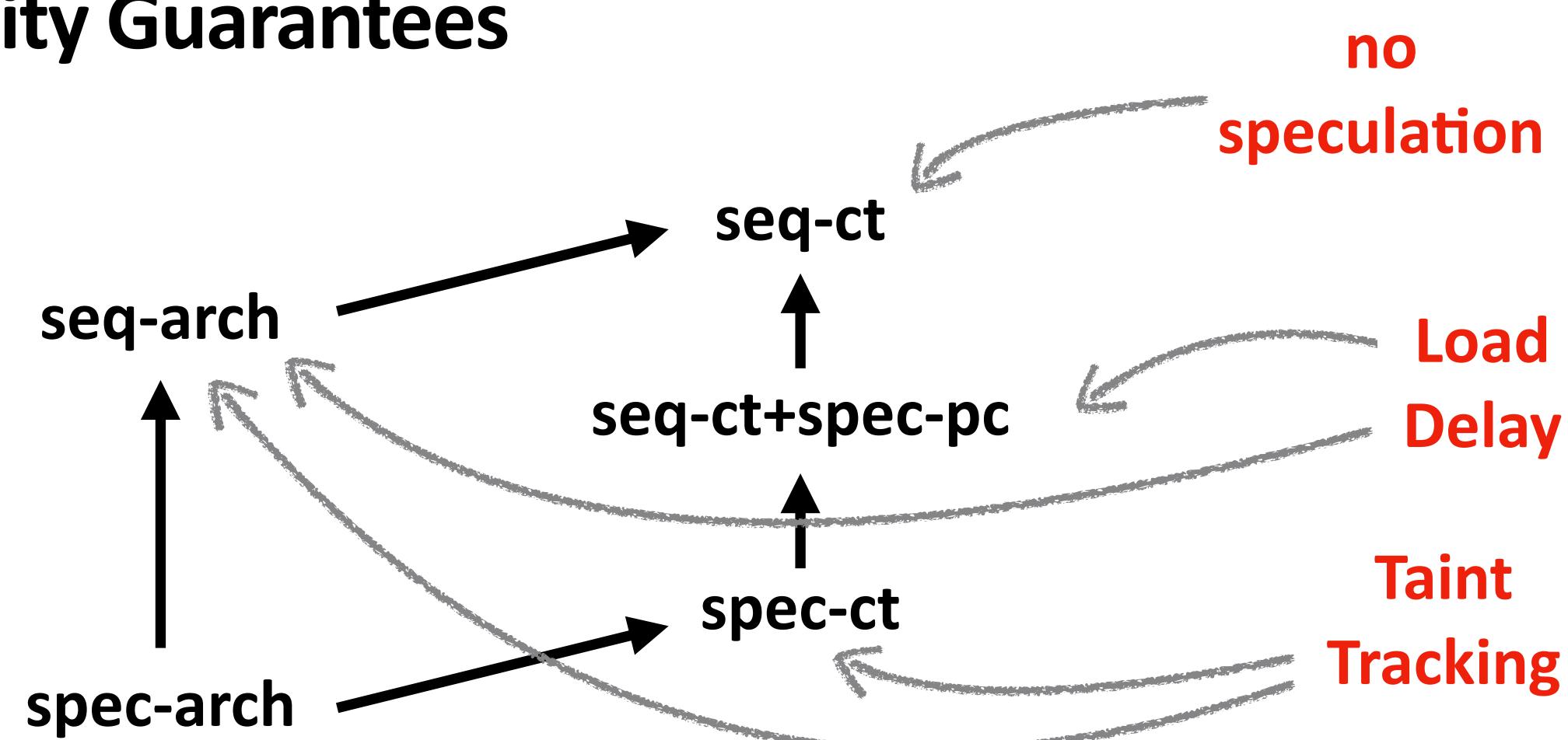
seq-ct f seq-ct+spec-pc f spec-ct

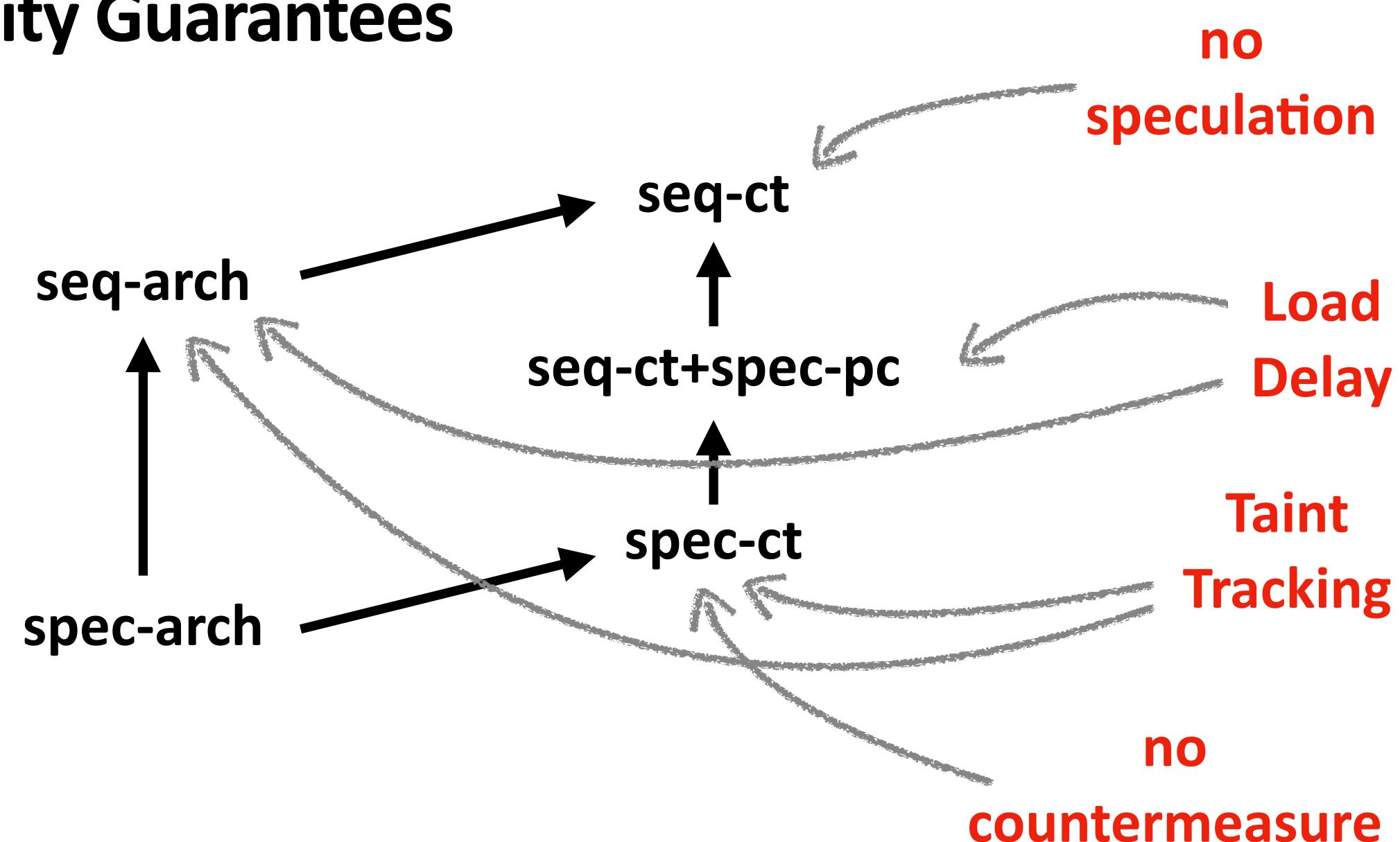


seq-ct f seq-ct+spec-pc f spec-ct

no speculation







Secure Programming

Program p is **non-interferent** wrt contract $\llbracket \cdot \rrbracket$ and policy π if for all arch. states σ , σ' : if $\sigma \approx_{\pi} \sigma'$ then $\llbracket p \rrbracket (\sigma) = \llbracket p \rrbracket (\sigma')$

Program p is **non-interferent** wrt contract $\llbracket \cdot rbrace$ and policy π if for all arch. states σ , σ' : if $\sigma \approx_{\pi} \sigma'$ then $[p](\sigma) = [p](\sigma')$

Specify secret data



Program *p* is *non-interferent* wrt contract $|\cdot|$ and policy π if for all arch. states σ , σ' : if $\sigma \approx_{\pi} \sigma'$ then $[p](\sigma) = [p](\sigma')$

Specify secret data



Program *p* is *non-interferent* wrt contract $|\cdot|$ and policy π if for all arch. states σ , σ' : if $\sigma \approx_{\pi} \sigma'$ then $[p](\sigma) = [p](\sigma')$

Specify secret data



Program p is **non-interferent** wrt contract $\left[\cdot \right]$ and policy π if for all arch. states σ , σ' : if $\sigma \approx_{\pi} \sigma'$ then $[p](\sigma) = [p](\sigma')$

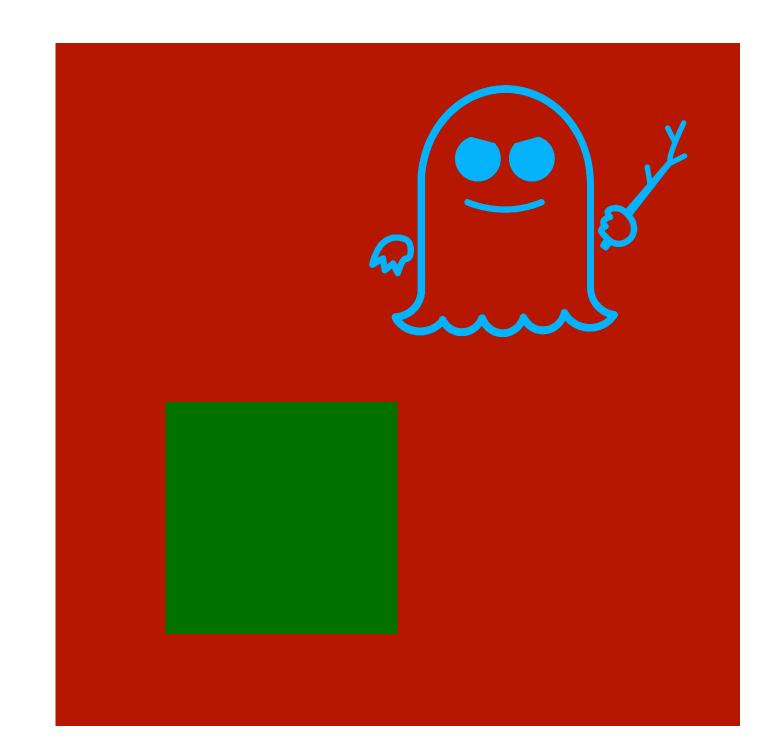
Theorem

Specify secret data

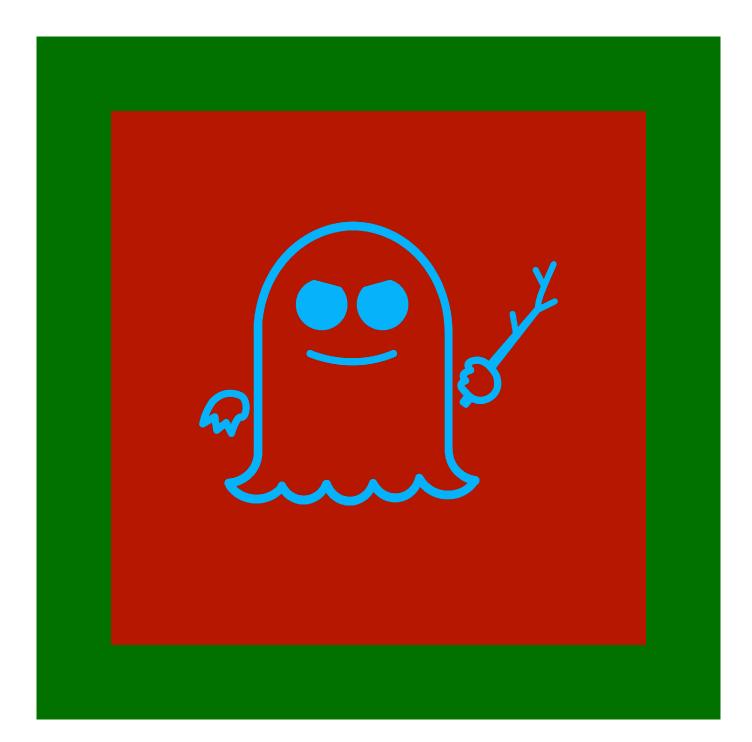
If p is **non-interferent** wrt contract $|\cdot|$ and policy π , and hardware { • } satisfies [•], then p is *non-interferent* wrt hardware $\{\cdot\}$ and policy π



Two Flavors of Secure Programming

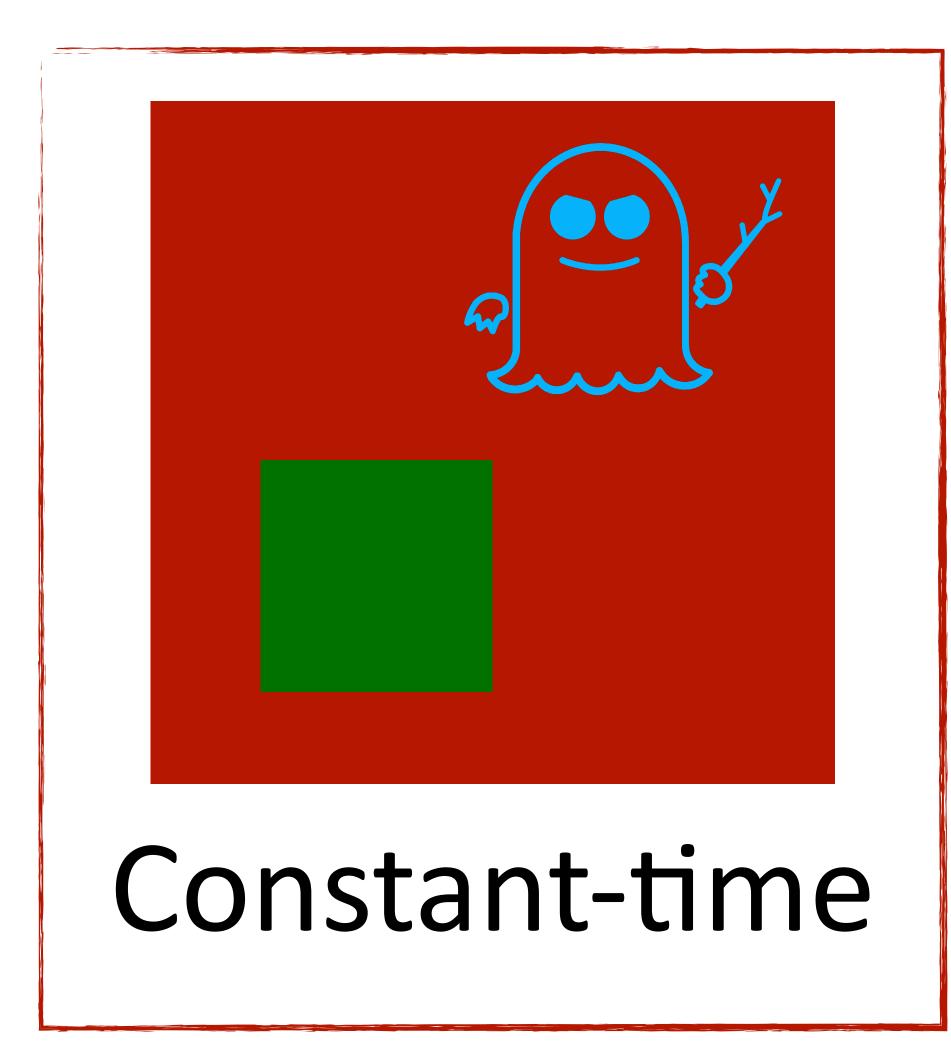


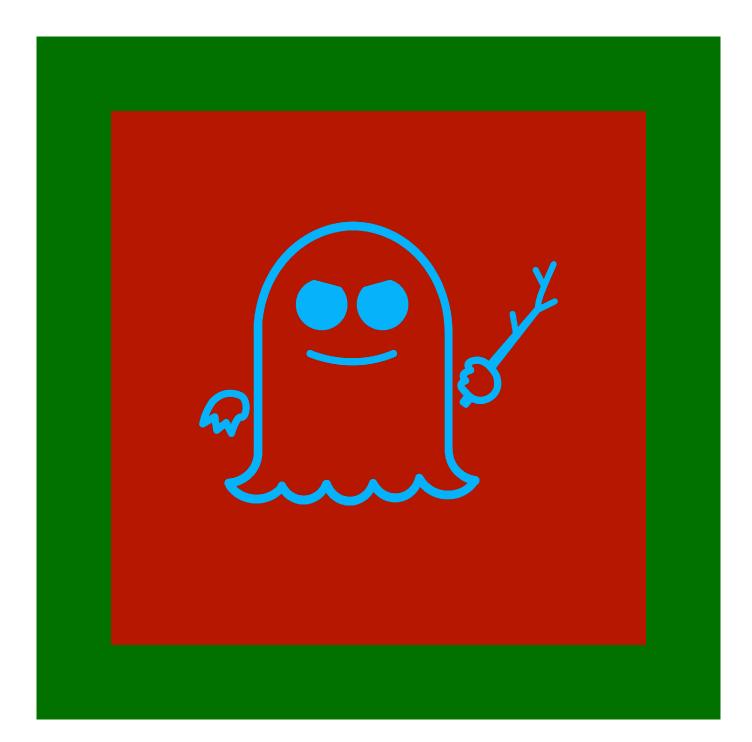
Constant-time



Sandboxing

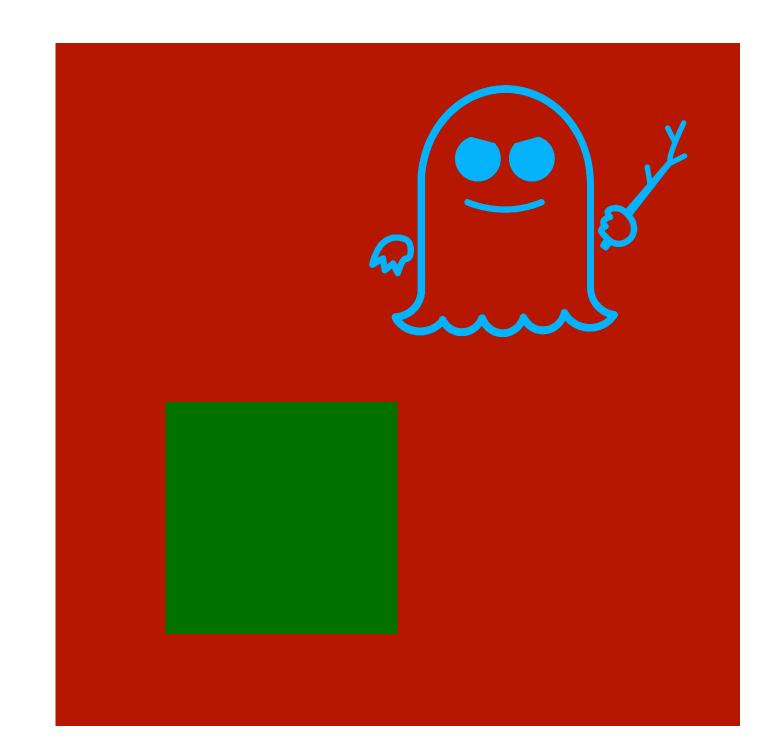
Two Flavors of Secure Programming



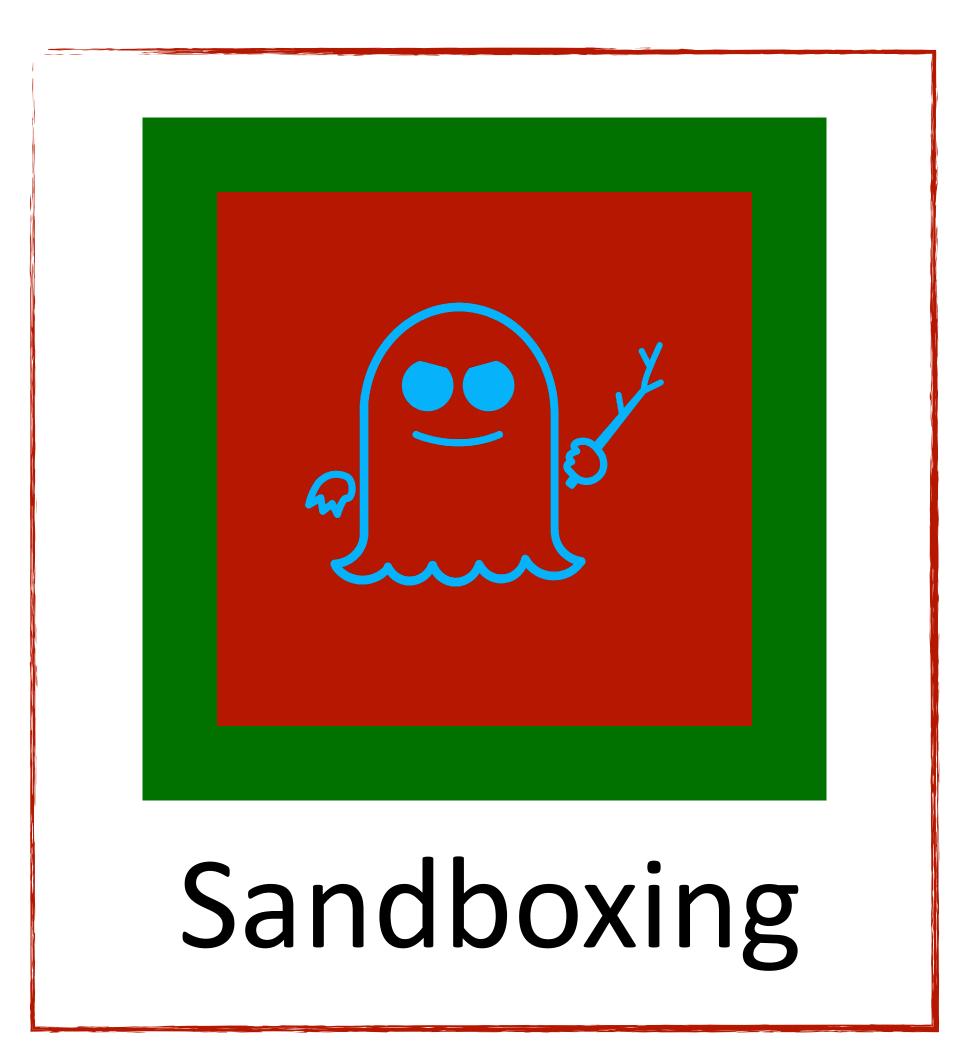


Sandboxing

Two Flavors of Secure Programming



Constant-time



Traditional CT wrt policy $\pi \equiv$ non-interference wrt seq-ct and π

Control-flow and memory accesses do not depend on secrets

Traditional CT wrt policy $\pi \equiv$ non-interference wrt seq-ct and π

Control-flow and memory accesses do not depend on secrets

Traditional CT wrt policy $\pi \equiv$ non-interference wrt seq-ct and π

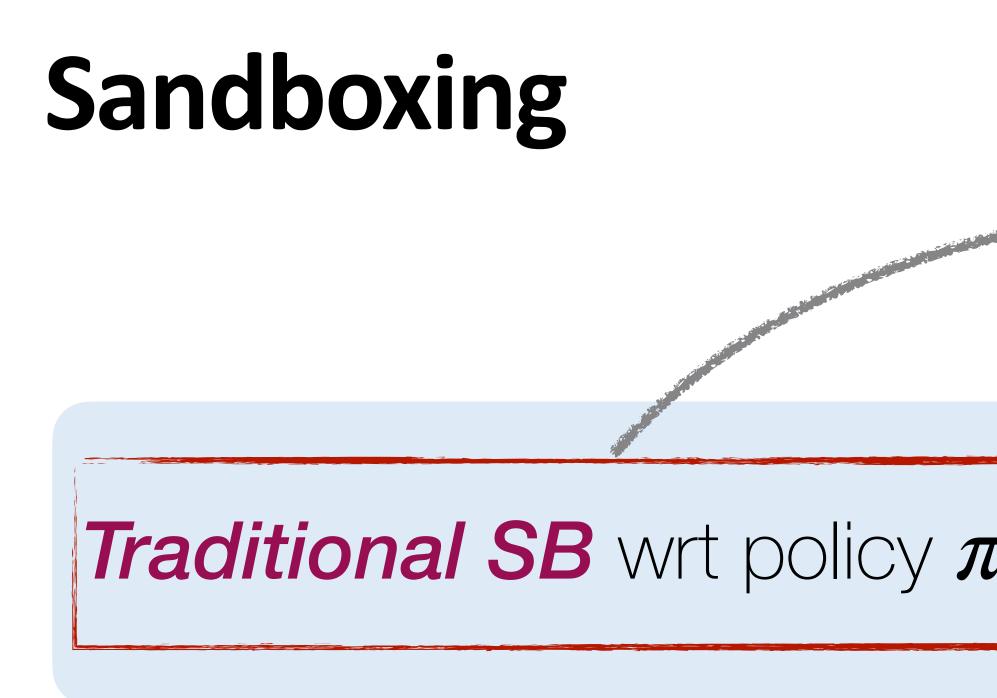
General CT wrt π and $|\cdot| \equiv$ non-interference wrt $|\cdot|$ and π

Sandboxing

Sandboxing

Traditional SB wrt policy $\pi \equiv$ non-interference wrt seq-arch and π

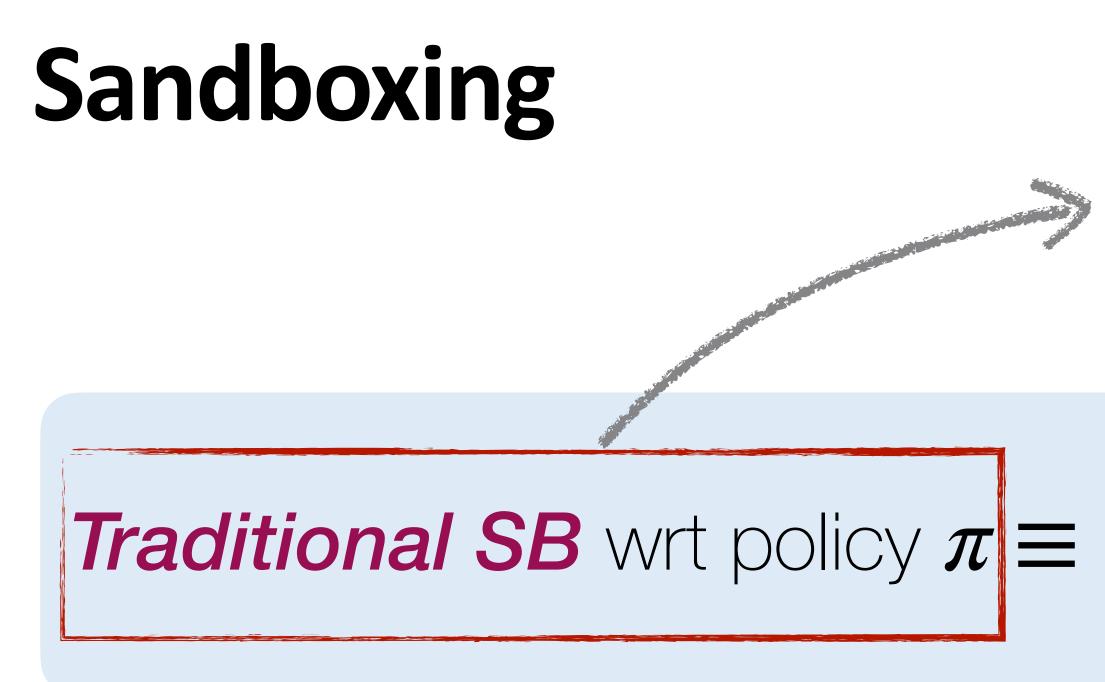




Programs never access high memory locations (out-of-sandbox)

Traditional SB wrt policy $\pi \equiv$ non-interference wrt seq-arch and π





General SB wrt π and $\llbracket \cdot \rrbracket \equiv$ Traditional SB wrt π + non-interference wrt π and $\llbracket \cdot \rrbracket$

Programs never access high memory locations (out-of-sandbox)

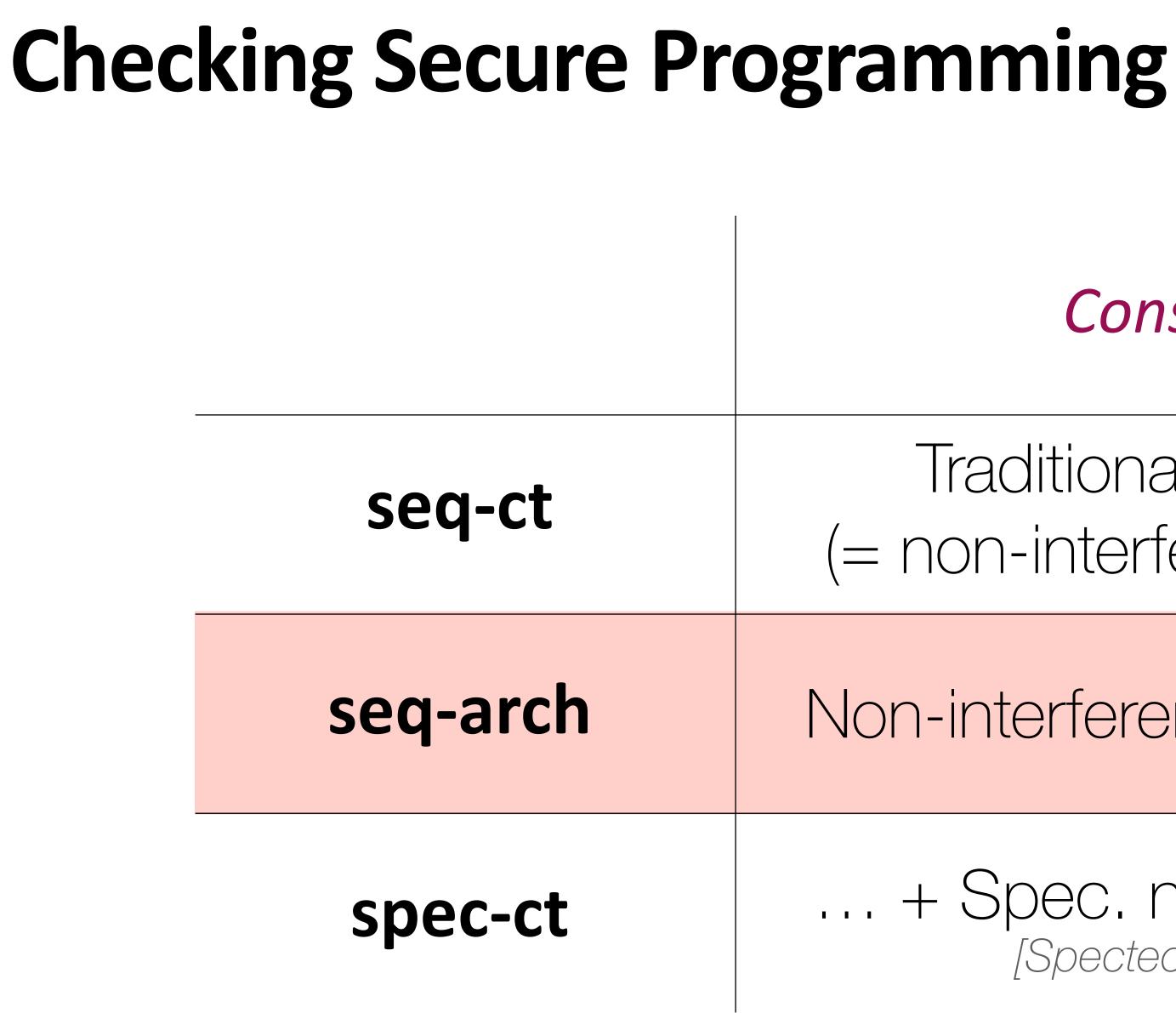
Traditional SB wrt policy $\pi \equiv$ non-interference wrt seq-arch and π





Checking Secure Programming Constant-time Traditional constant-time seq-ct (= non-interference wrt **seq-ct**) seq-arch Non-interference wrt seq-arch ... + Spec. non-interference spec-ct [Spectector, S&P'20]



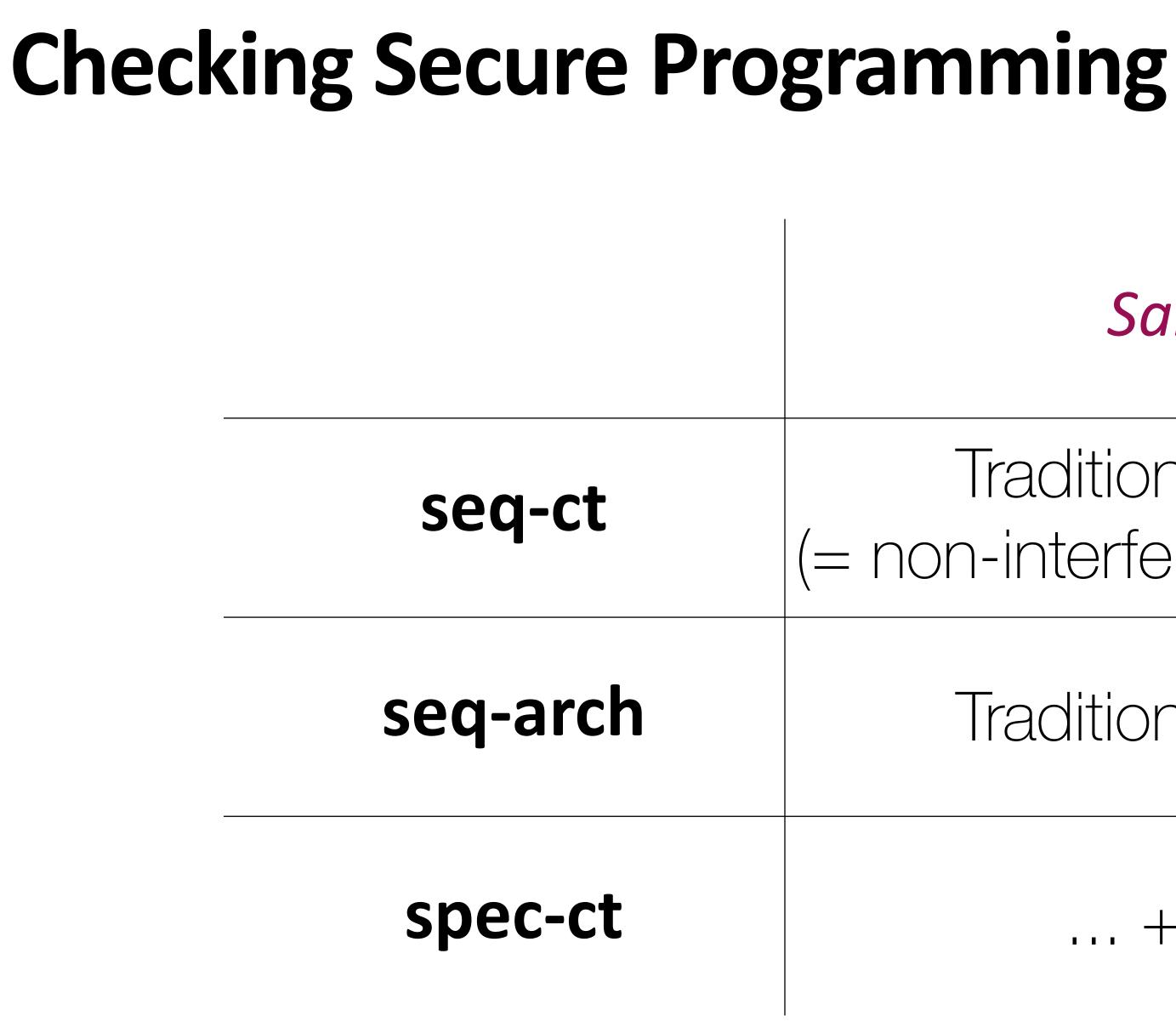


Constant-time

Traditional constant-time (= non-interference wrt **seq-ct**)

Non-interference wrt seq-arch

... + Spec. non-interference [Spectector, S&P'20]



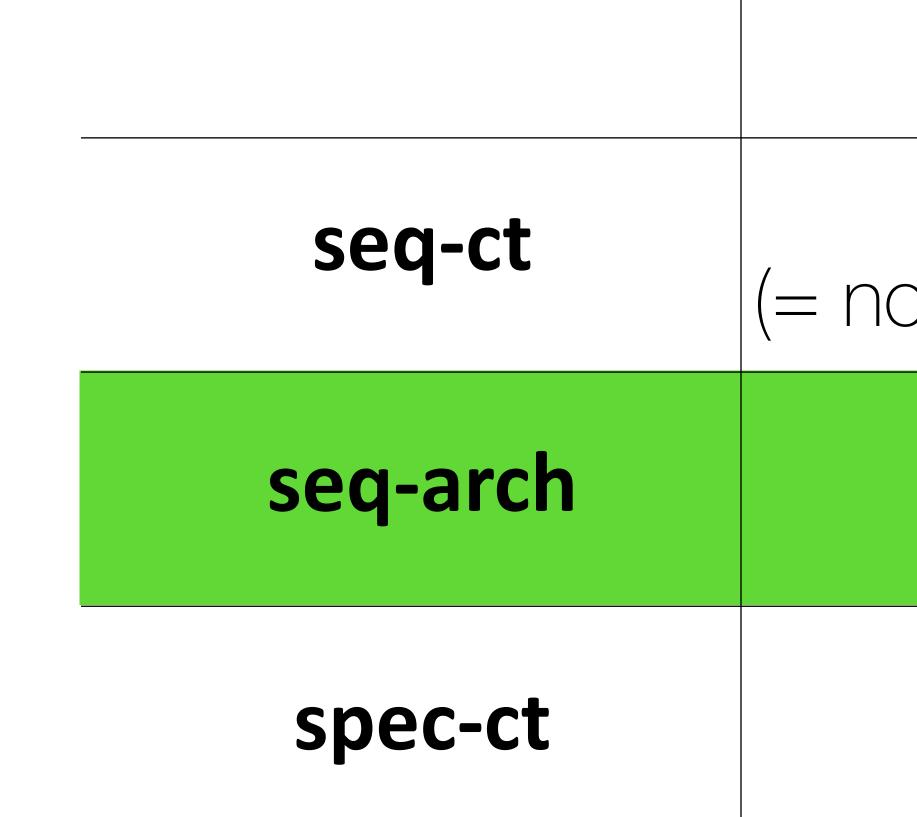
Sandboxing

Traditional sandboxing (= non-interference wrt **seq-arch**)

Traditional sandboxing

... + weak SNI

Checking Secure Programming





Sandboxing

Traditional sandboxing (= non-interference wrt **seq-arch**)

Traditional sandboxing

... + weak SNI



Conclusions

Need to rethink hardware-software contracts with security and safety in mind!

Need to rethink hardware-software contracts with security and safety in mind!

Should strive for simple and mechanism-independent contracts.

Need to rethink hardware-software contracts with security and safety in mind!

Find out more in our paper: M. Guarnieri, B. Köpf, J. Reineke, and P. Vila S&P (Oakland) 2021

Should strive for simple and mechanism-independent contracts.

Hardware-Software Contracts for Secure Speculation