Testing Software and Hardware against Speculation Contracts

Boris Köpf Microsoft Research

NUS CS Research Week 07/01/22

Based on Joint Work With

- Marco Guarnieri (IMDEA Software Institute)
- Jose Morales (IMDEA Software Institute)
- Oleksii Oleksenko (Microsoft Research)
- Jan Reineke (Saarland University)
- Andres Sanchez (EPFL)
- Mark Silberstein (Technion)
- Pepe Vila (Arm)

... and lots of discussions with & feedback from colleagues at Microsoft

Performance is Fundamental



CPUs Compilers Virtual machines Networks

. . .

minimize

time space energy



...

Performance-enhancing techniques...

- Caching
- Concurrency
- Deduplication
- Compression
- ...

... and their Impact on Security

- Caching
- Concurrency
- Deduplication
- Compression



Cache-timing attacks on AES

Loophole: Timing Attacks on Shared Event Loops in Chrome

Memory Deduplication as an Advanced Exploitation Vector

Spot me if you can: Uncovering spoken phrases in encrypted VoIP conversations



Example: Cache Side-Channel



2005: First attacks on AES (Bernstein/Shamir et al.)

2014-...: Highly effective attacks using shared caches (Yarom et al/...)

... and their Impact on Security

- Caching
- Concurrency
- Deduplication
- Compression
- Speculative execution

Cache-timing attacks on AES

Loophole: Timing Attacks on Shared Event Loops in Chrome

Memory Deduplication as an Advanced Exploitation Vector



Example: Speculative Leak



Spectre V1



void f(int x) if (x < A size) y = B[A[x]]</pre>













Covert Channels vs Side-channels

• Covert channels: Adversary = Sender & Receiver

on RSA etc.

1996

• Side-channels: Adversary = Eavesdropper

		Abstract	
A note on the Confinement	20 Years of Covert Channel Modelling	Covert channels emerged in mystery and departed in confusion.	
Problem	and Analysis	Spectrea	attacks
1973	1999	2018	
	Timing Attacks	Cache-timing	

attacks on AES

2005

100

Flush+Reload

2014

Countermeasures

- Software-based countermeasures
 - Insertion of speculation barriers, speculative load hardening, ...
 - Rely on (often implicit) assumptions about underlying hardware
- Hardware-based countermeasures
 - InvisiSpec (Micro 18), NDA (Micro 19), STT (Micro 19), SPT (Micro 21),...
 - Rely on (often implicit) assumptions about software



This Talk: Co-design for Secure Speculation



- 1. Checking software for contract compliance
- 2. Hardware-software contracts for secure speculation
- 3. Checking CPUs for contract compliance

Speculation Contracts

Speculation Contracts in a Nutshell

- Baseline contract: "constant-time programming":
 - Obligation on software: Make sure secrets don't affect loads, stores, branch targets
 - (Often implicit) obligation on hardware: Nothing *except* addresses of loads, stores, branch targets leaks
 - Technically: Make sure that all executions of a program that agree on addresses of loads, stores, and branch targets also agree on what a muarch attacker can observe
 - Note that both obligations are instance of *non-interference (NI)*: *P* satisfies NI ⇔ for all *h,h',I*: *P*(*h,I*)=*P*(*h',I*)
- Core idea: We generalize from executions and observations to capture the security properties of a wide range of speculation mechanisms

Examples of Contracts

Leakage of CPU without speculation

Leakage of CPU with branch speculation

- CT-Seq:
 - Observations: addresses of loads, stores, branch targets -
 - Executions: sequential in-order
- CT-Spec:
 - Observations: addresses of loads, stores, branch targets
 - Executions: sequential in-order + "mispredicted" branches up to a bound
- Arch-Seq
 - Observations: addresses of loads, stores, branch targets + data that is loaded
 - Executions: sequential in-order
- CT-Bpas, CT-Spec-Bpas,...

Only transiently loaded data is protected

What *is* a contract?

- A contract is a labelled ISA semantics, where labels correspond to the information that programs are allowed to leak during execution
- ISA:

Syntax

(Expressions)	е	:=	$n \mid x \mid \ominus e \mid e_1 \otimes e_2 \mid ite(e_1, e_2, e_3)$
(Instructions)	i	:=	skip $ x \leftarrow e $ load $x, e $ store x, e
			$ $ jmp $e $ beqz $x, \ell $ spbarr
(Programs)	р	:=	$i \mid p_1; p_2$

Core rules for CT-SEQ

 $\frac{\text{LOAD}}{p(a(\mathbf{pc})) = \text{load } x, e \quad x \neq \mathbf{pc} \quad n = \llbracket e \rrbracket(a)}{\langle m, a \rangle \xrightarrow{\text{load } n} \langle m, a [\mathbf{pc} \mapsto a(\mathbf{pc}) + 1, x \mapsto m(n)] \rangle}$

$$\frac{\text{BEQZ-SAT}}{p(a(\mathbf{pc})) = \mathbf{beqz} \ x, \ell \qquad a(x) = 0}{\langle m, a \rangle \xrightarrow{\mathbf{pc} \ \ell} \langle m, a[\mathbf{pc} \mapsto \ell] \rangle}$$

 $[p]_{ct}^{seq}(\sigma)$ = trace of observations

Core Rules for CT-Spec

$$\frac{STEP}{p(\sigma(\mathbf{pc})) \neq \mathbf{beqz} \ x, \ell} \quad \sigma \xrightarrow{\tau}_{ct}^{seq} \sigma'}{\langle \sigma, \omega + 1 \rangle \cdot s \xrightarrow{\tau}_{ct}^{spec} \langle \sigma', \omega \rangle \cdot s} \qquad \qquad \begin{array}{c} \text{ROLLBACK} \\ s = \langle \sigma', \omega' \rangle \cdot s' \\ \hline \langle \sigma, 0 \rangle \cdot s \xrightarrow{\mathbf{pc} \ \sigma'(\mathbf{pc}) \ spec} \\ \hline \langle \sigma, 0 \rangle \cdot s \xrightarrow{\mathbf{pc} \ \sigma'(\mathbf{pc}) \ ct} s \end{array} \qquad \qquad \begin{array}{c} \text{BARRIER} \\ p(\sigma(\mathbf{pc})) = \mathbf{spbarr} \quad \sigma \xrightarrow{\tau}_{ct}^{seq} \sigma' \\ \hline \langle \sigma, \omega + 1 \rangle \cdot s \xrightarrow{\tau}_{ct}^{spec} \langle \sigma', 0 \rangle \cdot s \end{array}$$

BRANCH

$$\frac{p(\sigma(\mathbf{pc})) = \mathbf{beqz} \ x, \ell \quad \ell_{correct} = \begin{cases} \ell & \text{if } \sigma(x) = 0 \\ \sigma(\mathbf{pc}) + 1 & \text{otherwise} \end{cases}}{\epsilon_{mispred} \left\{ \ell, \sigma(\mathbf{pc}) + 1 \right\} \setminus \ell_{correct}} \quad \omega_{mispred} = \begin{cases} w & \text{if } \omega = \infty \\ \omega & \text{otherwise} \end{cases}}{\epsilon_{mispred} \left\{ \sigma(\mathbf{pc} \mapsto \ell_{mispred}), \omega_{mispred} \right\} \cdot \left\{ \sigma(\mathbf{pc} \mapsto \ell_{correct}), \omega \right\} \cdot s \end{cases}}$$

Contracts form a Lattice



 $\llbracket \cdot \rrbracket_2 \rightarrow \llbracket \cdot \rrbracket_1$ means $\llbracket \cdot \rrbracket_2$ leaks more information than $\llbracket \cdot \rrbracket_1$

Checking Programs for Contract Compliance



Checking Programs for Contract Compliance

- Contracts pose a verification condition on software:
 - Make sure secrets don't affect contract traces
 - What is "secret" is defined by a policy $\boldsymbol{\pi}$

Definition 3 $(p \vdash NI(\pi, \llbracket \cdot \rrbracket))$. Program *p* is *non-interferent* w.r.t. contract $\llbracket \cdot \rrbracket$ and policy π if for all initial architectural states $\sigma, \sigma': \sigma \simeq_{\pi} \sigma' \Rightarrow \llbracket p \rrbracket(\sigma) = \llbracket p \rrbracket(\sigma')$.

- "Constant-time programming": secret is part of architectural state
- "Sandboxing": secret is memory that is *not accessed* during in-order execution

 $\llbracket p \rrbracket_{\operatorname{arch}}^{\operatorname{seq}}(\sigma) = \llbracket p \rrbracket_{\operatorname{arch}}^{\operatorname{seq}}(\sigma') \Rightarrow \llbracket p \rrbracket(\sigma) = \llbracket p \rrbracket(\sigma')$





Tools for Checking Software

- Spectector
- Specfuzz
- Binsec/Haunted
- Pitchfork
- KleeSpectre
- SpecuSym
- ...

[2105.05801] SoK: Practical Foundations for Software Spectre Defenses (arxiv.org)

Spectector

- 1. Spectector symbolically executes a program wrt to a contract semantics to obtain pairs of (Path condition, Observation trace)
- 2. We query Z3 whether, for all σ , σ' that satisfy the path condition, we have

$$\llbracket p \rrbracket_{\operatorname{arch}}^{\operatorname{seq}}(\cdot) = \llbracket p \rrbracket_{\operatorname{arch}}^{\operatorname{seq}}(\sigma') \Rightarrow \llbracket p \rrbracket_{\operatorname{ct}}^{\operatorname{spec}}(\sigma) = \llbracket p \rrbracket_{\operatorname{ct}}^{\operatorname{spec}}(\sigma')$$

Illustration: Kocher's Examples

- Ex 1: Vanilla Spectre 1
- Ex 2: Move leak to local function
- Ex 3: Local function that can't be inlined
- Ex 4: Left-shift y
- Ex 5: Use y as initial value of for loop
- Ex 6: Check bounds with mask rather than <
- Ex 7: Compare against last known-good value
- Ex 8: Use ?: operator
- Ex 9: Use separate value to communicate safety check
- Ex 10: Leak comparison result
- Ex 11: Use memcmp() to read memory for the leak
- Ex 12: Make index sum of two parameters
- Ex 13: Move safety check in inline function
- Ex 14: Invert lower bits of x
- Ex 15: Pass pointer to the length

1 if (y < size)
2 temp &= B[A[y] * 512];</pre>

Spectre Mitigations in Microsoft's C/C++ Compiler

	VISUAL C++ ICC							CLANG						
Ex.	U	NP	F	EN	UI	NP	FI	EN	U	NP	FI	EN	SI	LH
	-00	-02	-00	-02	-00	-02	-00	-02	-00	-02	-00	-02	-00	-02
01	0	0	•	•	0	0	٠	٠	0	0	•	•	•	•
02	0	0	•	•	0	0	•	•	0	0	•	•	•	•
03	0	0	•	0	0	0	•	•	0	0	•	•	•	٠
04	0	0	0	0	0	0	•	•	0	0	•	•	•	•
06	0	0	0	0	0	0	•	•	0	0	•	•	•	•
07	0	0	0	0	0	0	•	•	0	0	•	•	•	•
08	0	•	0	•	0	•	•	•	0	•	•	•	•	•
09	0	0	0	0	0	0	•	•	0	0	•	•	•	•
10	0	0	0	0	0	0	•	•	0	0	•	•	•	0
11	0	0	0	0	0	0	•	•	0	0	•	•	•	•
12	0	0	0	0	0	0	•	•	0	0	•	•	•	•
13	0	0	0	0	0	0	•	•	0	0	•	•	•	•
14	0	0	0	0	0	0	•	•	0	0	•	•	•	•
15	0	0	0	0	0	0	•	•	0	0	•	•	0	•

1 if (y < size)
2 temp &= B[A[y] * 512];</pre>

Clang V7.0.0 -O2 Speculative Load Hardening



1	if (y < size)
2		if $(A[y] == k)$
3		temp &= B[0];

Clang V7.0.0 -O2 Speculative Load Hardening

	-			
		1	mov	size, %rdx
		2	mov	y, %rbx
		3	mov	\$0, %rax
		4	cmp	%rbx , %rdx
		5	jbe	END
		6	cmovbe	\$-1, %rax
		7	or	%rax, %rbx
		8	mov	k, %rcx
We detect that A[0xFFFF]		9	cmp	%rcx, A(%rbx)
can leak via control flow		10	jne	END
	1	11	cmovne	\$-1, %rax
	1	12	mov	B, %rcx
	1	13	and	<pre>%rcx, temp</pre>
	1	14	jmp	END

1

temp &= B[A[y<size?(y+1):0]*512];

Intel ICC V19.0.0.117 -O2 w/ speculation barriers

	1	mov	y, %rdi
	2	lea	1(%rdi), %rdx
	3	mov	size, %rax
	4	xor	%rcx, %rcx
	5	cmp	%rax, %rdi
	6	cmovb	%rdx, %rcx
	7	mov	temp, %r8b
	8	mov	A(%rcx), %rsi
ICC inserts spurious fence	9	shl	\$9, %rsi
	- 10	lfence	
	11	and	B(<mark>%rsi),</mark> %r8b
	12	mov	%r8b, temp

-

Checking CPUs for Contract Compliance



Checking CPUs for Contract Compliance

- A CPU satisfies a contract if programs do not leak more information to a microarchitectural adversary than what the contract specifies
 - For all programs, whenever two executions agree on contract traces, they must also agree on hardware traces
- What is a "CPU", what are "hardware traces"?
 - 1. "CPU" is an operational semantics with uarch components; hardware traces are obtained as a projection
 - Captures simple out-of-order CPU with 3-stage pipeline
 - Operates on registers, main memory, and reorder buffer
 - Stubs for caches, branch predictors, scheduler
 - 2. "CPU" is a fabricated chip; hardware traces are given by side-channel attack (e.g. Prime+Probe on L1D)

Contracts for Mechanisms for Secure Speculation



Testing Black-box CPUs against Speculation Contracts

- Key observation: checking contract compliance can be done in a black-box fashion
 - For all programs, whenever two executions agree on contract traces, they must also agree on hardware traces
- Challenges:
 - How to cope with the intractable search space?
 - How to implement "contracts" for a realistic ISA?
 - How to obtain deterministic hardware traces?

Revizor (ASPLOS '22)



- Test case generator: Creates DAG, adds terminators to blocks, populates with random instructions (from specified subsets) and operands (from specified subsets)
- Input generator: Generates random 32 bit numbers for registers, flags, and memory (1 or 2 pages)
- **Model:** Unicorn (QEMU-based), instrumented to collect traces + explore mispredicted branches
- **Executor:** Prime+Probe (on L1D) and Prime+Probe+assists (clear page table bit), based on nanoBench
 - Priming: Run each test case in a loop with different pseudorandom inputs to ensure muarch state is primed in a diverse but deterministic fashion

Results

	Target 1	Target 2	Target 3	Target 4	Target 5	Target 6	Target 7	Target 8
CPU				Skylake	Coffee Lake			
V4 patch		off			on		on	
Instruction Set	AR	AR+MEM	AR+MEM+VAR	AR+MEM+VAR	AR+MEM+CB	AR+MEM+CB+VAR	AI	R+MEM
Executor Mode	Prime+Probe							Probe+Assist
Table 2: Description of the experimental setups.								

	Target 1	Target 2	Target 3	Target 4	Target 5	Target 6	Target 7	Target 8
CT-SEQ	×	✓ (V4)	✓ (V4)	×	✓ (V1)	✓ (V1)	✓ (MDS)	✓ (LVI-Null)
CT-BPAS	\times^*	×	✓ (V4-var**)	\times^*	✓ (V1)	✓ (V1)	✓ (MDS)	✓ (LVI-Null)
CT-COND	\times^*	✓ (V4)	✓ (V4)	\times^*	×	✓ (V1-var**)	✓ (MDS)	✓ (LVI-Null)
CT-COND-BPAS	\times^*	\times^*	✓ (V4-var**)	\times^*	\times^*	✓ (V1-var**)	✓ (MDS)	✓ (LVI-Null)

* we did not repeat the experiment as a stronger contract was already satisfied.

** the violation represents a novel speculative vulnerability.

Table 3: Testing results. \checkmark means Revizor detected a violation; \times means Revizor detected no violations within 24h of testing. In parenthesis are Spectre-type vulnerabilities revealed by the detected violations.

- AR: in-register arithmetic, including logic and bitwise;
- MEM: memory operands and loads/stores;
- VAR: variable-latency operations (divisions).
- CB: conditional branches;

Detected Subtleties

• New variants of V1 & V4

1 b = variable_latency(a)
2 if (...) # misprediction
3 c = array[b] # executed if the latency is short

• Speculative stores can modify the cache on Coffee Lake (but likely not on Skylake)

Detection Speed

Time-to-violation

	Detection time								
Contract-permitted	V4-type	V1-type	MDS-type	LVI-type					
leakage	(Target 2)	(Target 5)	(Target 7)	(Target 8)					
None	73'25" (.7)	4'51" (.9)	5'35" (.7)	7'40" (1.1)					
V4	N/A	3'48" (.7)	6'37" (.8)	3'06" (1.0)					
V1	140'42" (.6)	N/A	7'03" (.8)	3'22" (.3)					

Summary

- We propose **HW/SW Contracts** as a framework for specifying security of speculative execution.
 - Captures unprotected speculation, InvisiSpec, Speculative Taint Tracking, ...
 - Can be used as a basis for secure programming
- We built **Spectector**, a tool to detect speculative leaks in software
- We built **Revizor**, a tool to test CPUs against contracts
 - Revizor generates random code snippets to find contract violations
 - Automatically surfaces V1, V4, LVI, MDS on x86 (Skylake and Coffee Lake)
- Many avenues for future work, including coverage, white-box analysis, and more expressive contracts

Links

- <u>Hardware-Software Contracts for Secure Speculation Microsoft</u> <u>Research</u> (IEEE S&P '21)
- <u>Spectector: Principled Detection of Speculative Information Flows</u> -<u>Microsoft Research</u> (IEEE S&P '20)
- <u>Revizor: Testing Black-box CPUs against Speculation Contracts</u> (arxiv.org) (ASPLOS '22)
- <u>Full Time Opportunities: Researcher (Side-channel Attacks and Defenses) in Cambridge | Research at Microsoft</u>

Illustration: Kocher's Examples

- Ex 1: Vanilla Spectre 1
- Ex 2: Move leak to local function
- Ex 3: Local function that can't be inlined
- Ex 4: Left-shift y
- Ex 5: Use y as initial value of for loop
- Ex 6: Check bounds with mask rather than <
- Ex 7: Compare against last known-good value
- Ex 8: Use ?: operator
- Ex 9: Use separate value to communicate safety check
- Ex 10: Leak comparison result
- Ex 11: Use memcmp() to read memory for the leak
- Ex 12: Make index sum of two parameters
- Ex 13: Move safety check in inline function
- Ex 14: Invert lower bits of x
- Ex 15: Pass pointer to the length

i if (y < size)
temp &= B[A[y] * 512];</pre>

	VISUAL C++ ICC							CLANG						
Ex.	U	NP	F	EN	UI	NP	FI	EN	U	NP	FI	EN	SI	LH
	-00	-02	-00	-02	-00	-02	-00	-02	-00	-02	-00	-02	-00	-02
01	0	0	•	•	0	0	٠	٠	0	0	•	•	•	•
02	0	0	•	•	0	0	•	•	0	0	•	•	•	•
03	0	0	•	0	0	0	•	•	0	0	•	•	•	٠
04	0	0	0	0	0	0	•	•	0	0	•	•	•	•
06	0	0	0	0	0	0	•	•	0	0	•	•	•	•
07	0	0	0	0	0	0	•	•	0	0	•	•	•	•
08	0	•	0	•	0	•	•	•	0	•	•	•	•	•
09	0	0	0	0	0	0	•	•	0	0	•	•	•	•
10	0	0	0	0	0	0	•	•	0	0	•	•	•	0
11	0	0	0	0	0	0	•	•	0	0	•	•	•	•
12	0	0	0	0	0	0	•	•	0	0	•	•	•	•
13	0	0	0	0	0	0	•	•	0	0	•	•	•	•
14	0	0	0	0	0	0	•	•	0	0	•	•	•	•
15	0	0	0	0	0	0	•	•	0	0	•	•	0	•

1 if (y < size)
2 temp &= B[A[y] * 512];</pre>

Clang V7.0.0 -O2 Speculative Load Hardening



1	if (y < size)
2		if $(A[y] == k)$
3		temp &= B[0];

Clang V7.0.0 -O2 Speculative Load Hardening

	-			
		1	mov	size, %rdx
		2	mov	y, %rbx
		3	mov	\$0, %rax
		4	cmp	%rbx , %rdx
		5	jbe	END
		6	cmovbe	\$-1, %rax
		7	or	%rax, %rbx
		8	mov	k, %rcx
We detect that A[0xFFFF]		9	cmp	%rcx, A(%rbx)
can leak via control flow		10	jne	END
	1	11	cmovne	\$-1, %rax
	1	12	mov	B, %rcx
	1	13	and	<pre>%rcx, temp</pre>
	1	14	jmp	END

1

temp &= B[A[y<size?(y+1):0]*512];

Intel ICC V19.0.0.117 -O2 w/ speculation barriers

	1	mov	y, %rdi
	2	lea	1(%rdi), %rdx
	3	mov	size, %rax
	4	xor	%rcx, %rcx
	5	cmp	%rax, %rdi
	6	cmovb	%rdx, %rcx
	7	mov	temp, %r8b
	8	mov	A(%rcx), %rsi
ICC inserts spurious fence	9	shl	\$9, %rsi
	- 10	lfence	
	11	and	B(<mark>%rsi),</mark> %r8b
	12	mov	%r8b, temp

-

Core Idea 2: Checking Security

1. We symbolically execute speculative semantics to obtain pairs of (Path condition, Observation trace) load, store, pc, start, rollback

size, %rax mov y, %rbx mov \$0, %rdx mov 3 %rbx, %rax cmp 4 jbe END 5 cmovbe \$-1, %rdx A(%rbx), %rax mov 7 \$9, %rax shl 8 or %rdx, %rax 9 B(%rax), %rax mov 10 %rdx, %rax or 11 and %rax, temp 12

start \cdot rollback $\cdot \tau$ when y < sizestart $\cdot \tau \cdot$ rollback when $y \ge size$

 $\tau = \mathbf{loadO} (A + y) \cdot \mathbf{loadO} (B + (A[y] * 512) || mask)$

mask = ite(y < size, 0x0, 0xFF..FF)</pre>

Checking Contract Compliance vai

- 1. We symbolically execute program to obtain pairs of (Path condition, Observation Trace)
- 2. We query Z3 whether, for all public *lo* and all secret *hi*, *hi'* that satisfy the path condition, we have *lo*: size, y, A, B

hi: A[y]

• Obs (*lo,hi*) = Obs(*lo,hi*') implies Obs^{Spec}(*lo,hi*) = Obs^{Spec}(*lo,hi*') —

