



# Scoré-co-Léak Eorwarding

Théré and Back Ágain

Claudio Canella (@cc0x1f), Lukas Giner (@redrabbyte), Michael Schwarz (@misc0110) October 2, 2020

Graz University of Technology

CISPA Helmholtz Center for Information Security

#BHASIA @BLACKHATEVENTS





#### Claudio Canella

PhD student @ Graz University of Technology

- **У** @cc0x1f





#### Lukas Giner

PhD student @ Graz University of Technology

**梦** @redrabbyte

■ lukas.giner@iaik.tugraz.at





#### Michael Schwarz

Faculty @ CISPA Helmholtz Center for Information Security

- **y** @misc0110
- michael.schwarz@cispa.saarland



• How do loads handle previous stores?



- How do loads handle previous stores?
- How is Meltdown mitigated in hardware?



- How do loads handle previous stores?
- How is Meltdown mitigated in hardware?
- Can we abuse these mechanisms for new attacks?



- How do loads handle previous stores?
- How is Meltdown mitigated in hardware?
- Can we abuse these mechanisms for new attacks?
- Can we mitigate such attacks efficiently?





• Mental model of CPU is simple



- Mental model of CPU is simple
- Instructions are executed in program order



- Mental model of CPU is simple
- Instructions are executed in program order
- Pipeline stalls when stages are not ready



- Mental model of CPU is simple
- Instructions are executed in program order
- Pipeline stalls when stages are not ready
- If data is not cached, we need to wait



IF	ID	EX	МЕМ	WB					
	IF	ID	EX	МЕМ	WB				
		IF	ID	EX	МЕМ	WB			
			IF	ID	EX	МЕМ	WB		
				IF	ID	EX	MEM	WB	

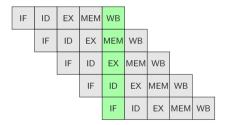
- Instructions are...
  - fetched (IF) from the L1 Instruction Cache



IF	ID	EX	МЕМ	WB					
	IF	ID	EX	МЕМ	WB				
		IF	ID	EX	МЕМ	WB			
			IF	ID	EX	МЕМ	WB		
				IF	ID	EX	MEM	WB	

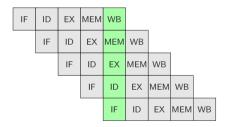
- Instructions are...
  - fetched (IF) from the L1 Instruction Cache
  - decoded (ID)





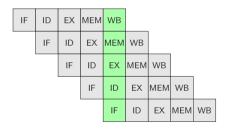
- Instructions are...
  - fetched (IF) from the L1 Instruction Cache
  - decoded (ID)
  - executed (EX) by execution units





- Instructions are...
  - fetched (IF) from the L1 Instruction Cache
  - decoded (ID)
  - executed (EX) by execution units
- Memory access is performed (MEM)





- Instructions are...
  - fetched (IF) from the L1 Instruction Cache
  - decoded (ID)
  - executed (EX) by execution units
- Memory access is performed (MEM)
- Architectural register file is updated (WB)





• No dependency between instructions? Why wait!





- No dependency between instructions? Why wait!
- Execute out of order





- No dependency between instructions? Why wait!
- Execute out of order, retire in order





- No dependency between instructions? Why wait!
- Execute out of order, retire in order
- What happens when an instruction faults?





- No dependency between instructions? Why wait!
- Execute out of order, retire in order
- What happens when an instruction faults?
- Undo out-of-order effects → instructions were transient





• Transient instructions are undone, what good are they?





- Transient instructions are undone, what good are they?
- Some changes are persistent, but invisible architecturally





- Transient instructions are undone, what good are they?
- Some changes are persistent, but invisible architecturally
- → encode data in microarchitectural state





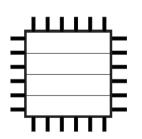
- Transient instructions are undone, what good are they?
- Some changes are persistent, but invisible architecturally
- → encode data in microarchitectural state
- The cache is a simple solution

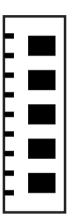
# Measuring Cache State



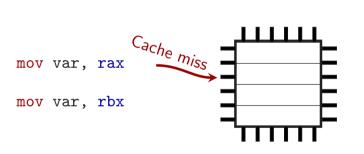
mov var, rax

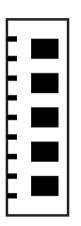
mov var, rbx



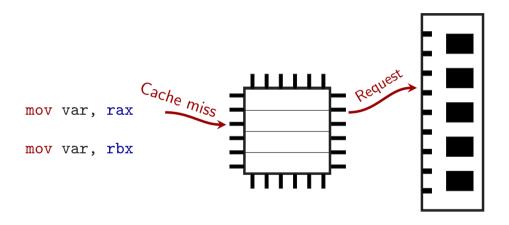




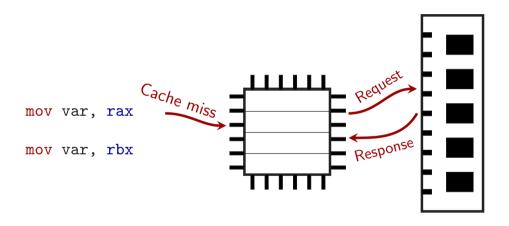




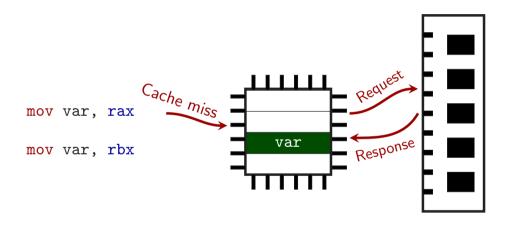




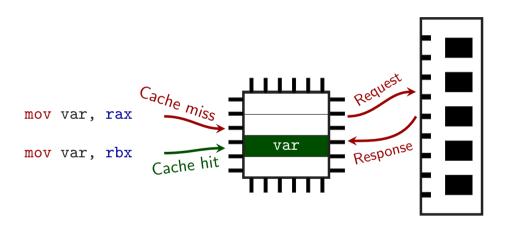






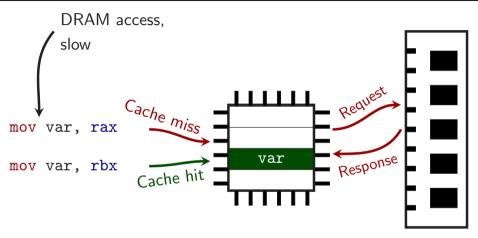




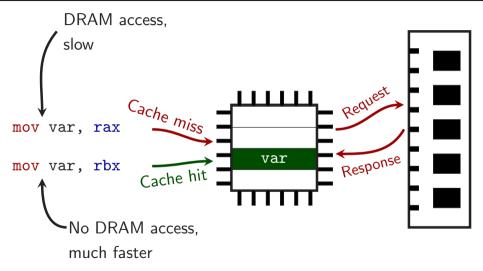


# Measuring Cache State

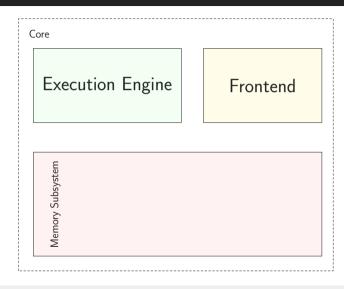




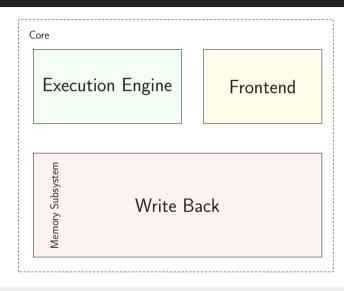




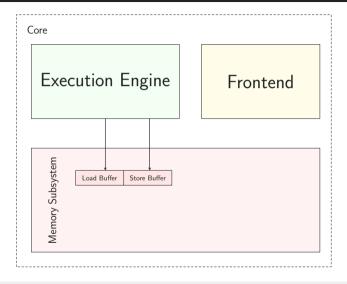




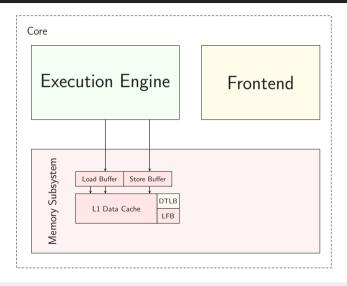




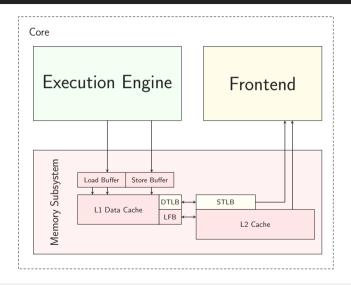




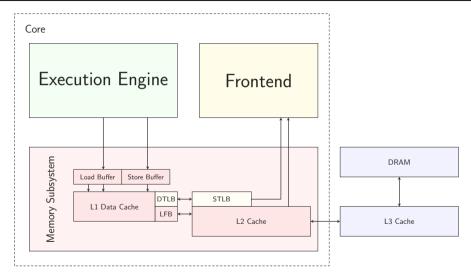




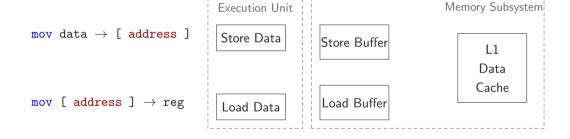




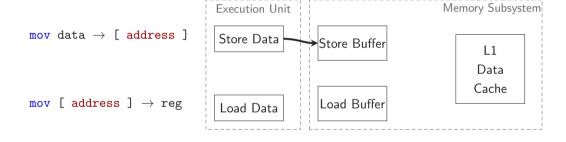




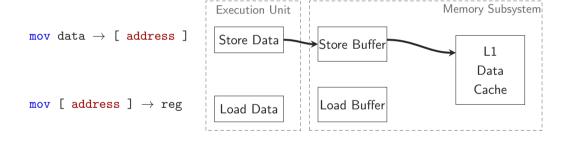




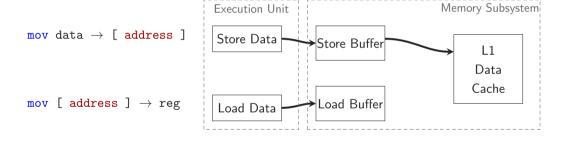




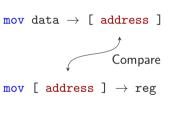


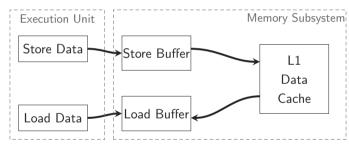




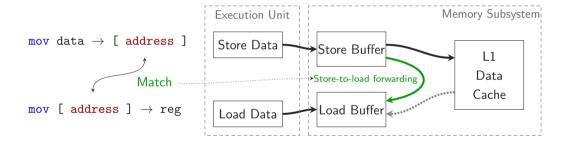












Addresses

Addresses



#### True Positive

Forwarding

True Negative

Forwarding





Addresses

Forwarding

Addresses

Forwarding













True Positive



Store-to-Leak











**True Negative** 









Forwarding

Forwarding



True Positive





**False Positive** 



Store-to-Leak

**Fallout** 



**True Negative** Addresses









Forwarding

Addresses

Forwarding























**True Negative** 







• Optimization during transient execution





- Optimization during transient execution
- Failures not visible architecturally





- Optimization during transient execution
- Failures not visible architecturally
- Store-to-Leak relies on correct matching





- Optimization during transient execution
- Failures not visible architecturally
- Store-to-Leak relies on correct matching
- Only stores to valid addresses are forwarded



#### **Intel Architecture Optimization Reference Manual**

[The store execution phase] Fills the store buffers with linear and physical address and data.



#### **Intel Architecture Optimization Reference Manual**

[The store execution phase] Fills the store buffers with linear and physical address and data. Once store address and data are known, the store data can be forwarded to the following load operations that need it. [...]





ullet Virtual address that is physically backed o valid





- ullet Virtual address that is physically backed o valid
- Permission checks are deferred





- Virtual address that is physically backed → valid
- Permission checks are deferred
- → Exploited in Meltdown attacks





- ullet Virtual address that is physically backed o valid
- Permission checks are deferred
- → Exploited in Meltdown attacks
- → Store-to-load forwarding on inaccessible addresses



\*kernel = 'X'



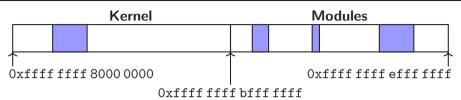




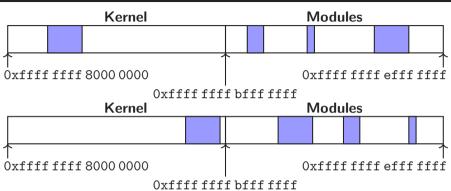




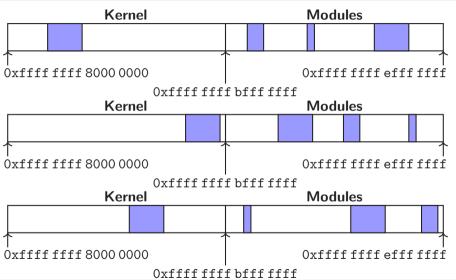




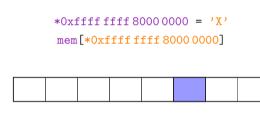


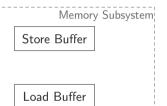








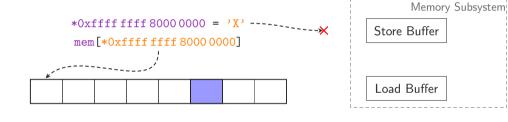




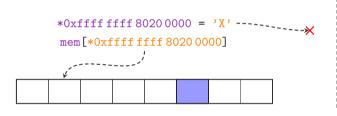


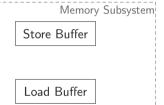






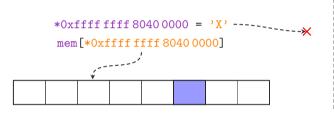


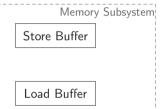






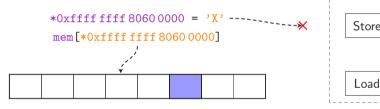


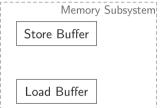








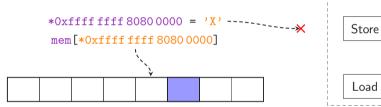


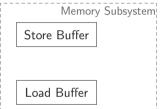


mem: ABCDEFGHIJKLMNOPQRSTUVWXYZ





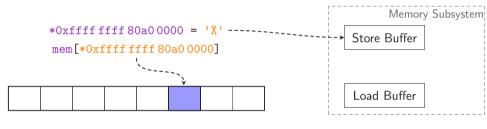




mem: ABCDEFGHIJKLMNOPQRSTUVWXYZ



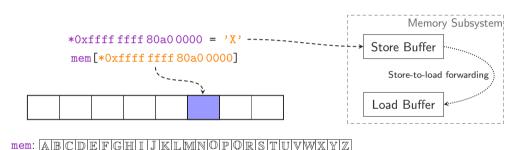




mem: ABCDEFGHIJKLMNOPQRSTUVWXYZ

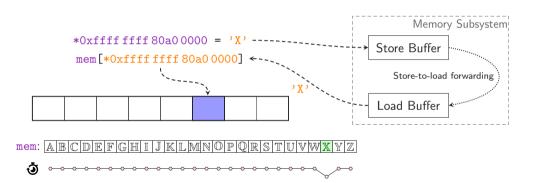


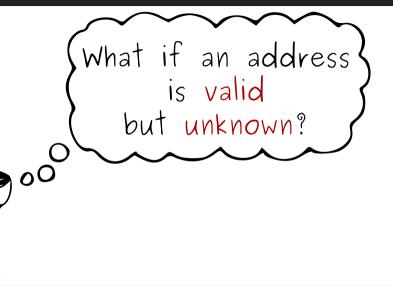




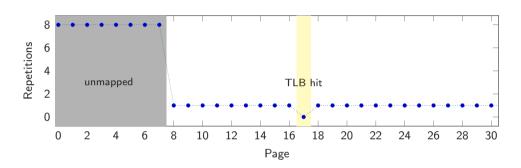
**5** 





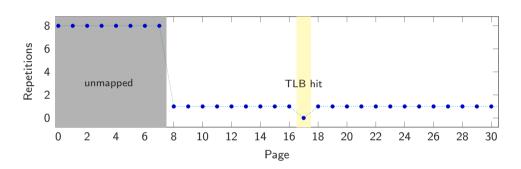






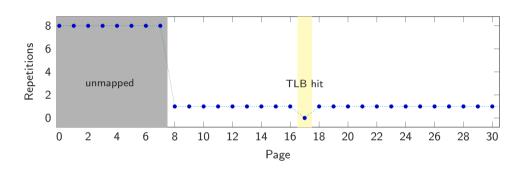
ullet Forward on first try o in TLB





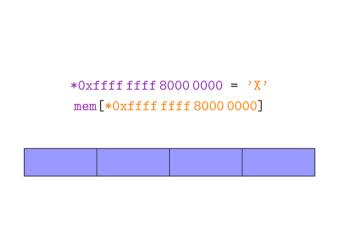
- ullet Forward on first try o in TLB
- ullet On second try o valid, not in TLB





- ullet Forward on first try o in TLB
- ullet On second try o valid, not in TLB
- $\bullet \ \ \mathsf{Not} \ \mathsf{forwarded} \to \mathsf{unmapped}$



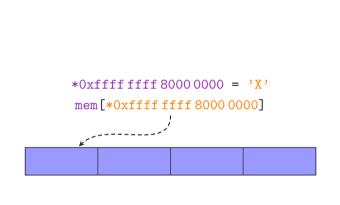


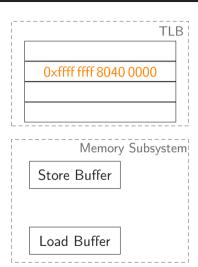


Memory Subsystem
Store Buffer

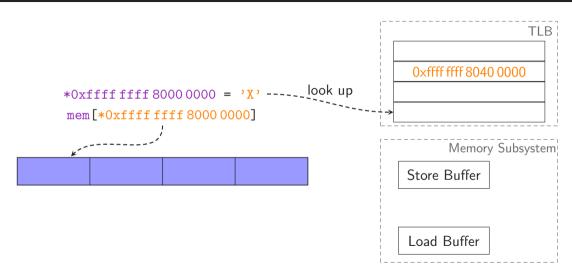
Load Buffer



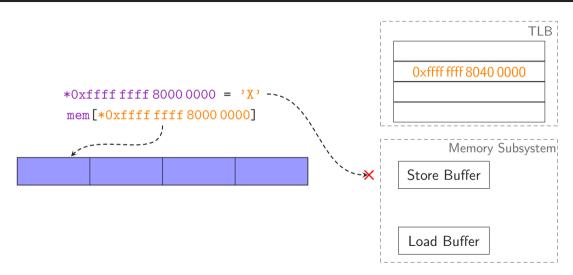




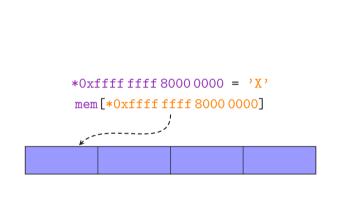


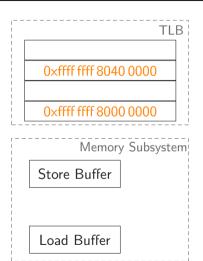




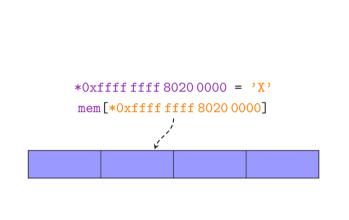


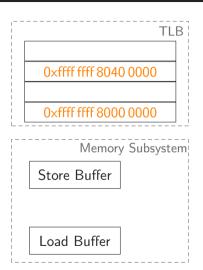




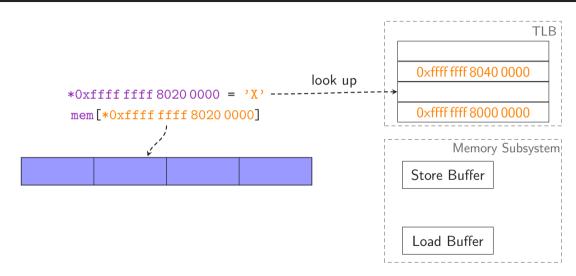




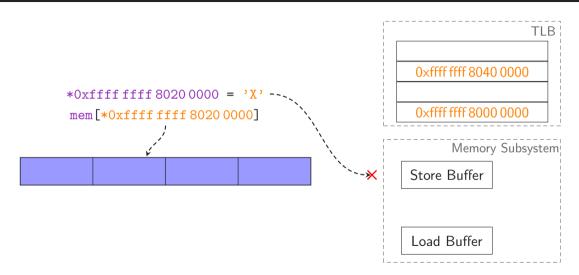




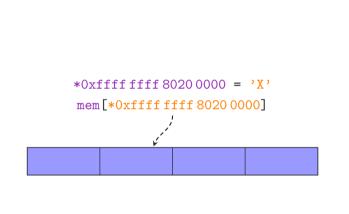


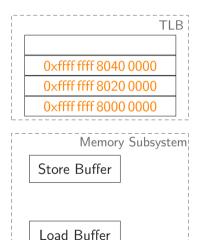




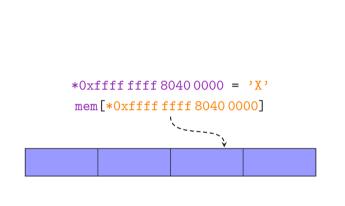


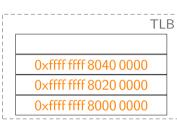








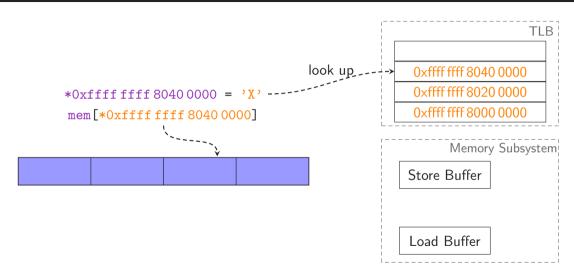




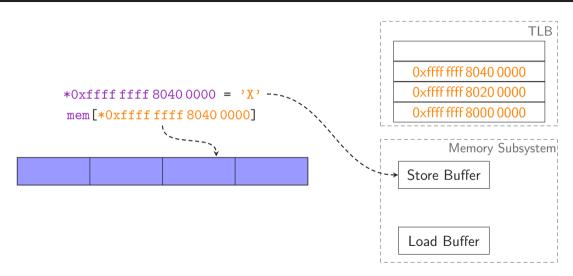
Memory Subsystem
Store Buffer

Load Buffer

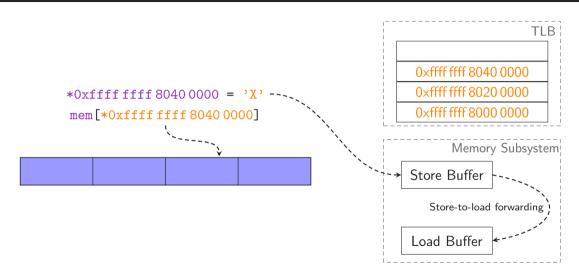




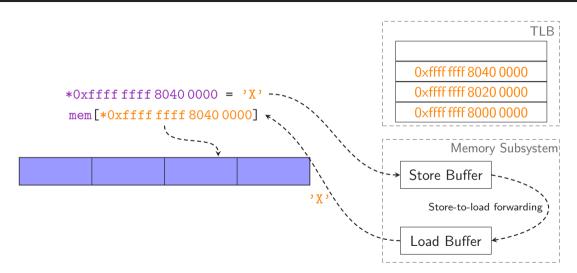




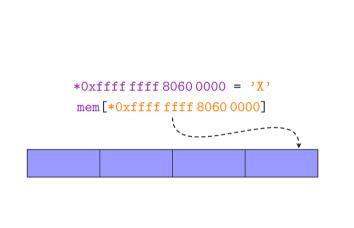


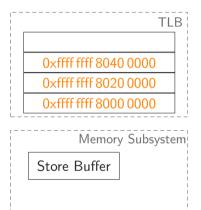






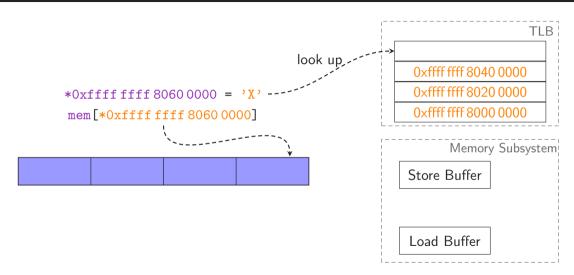




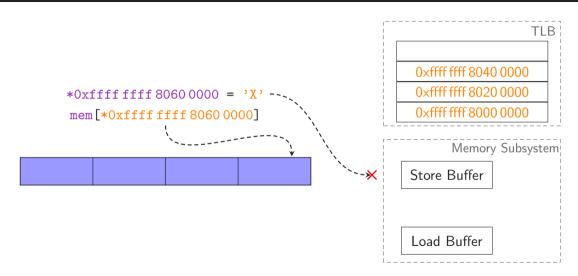


Load Buffer

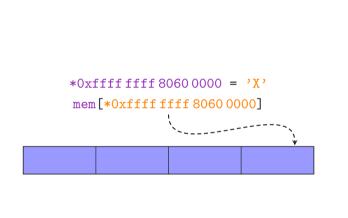


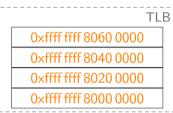












Memory Subsystem
Store Buffer

Load Buffer



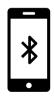
Kernel			Modules						



	Kernel		Modul	es	
			(*)		



	Kernel		Mod	ules	
			*		





	Kernel	-	Modules	
			<b> </b>	





	Kernel	-	Modules	
			*	



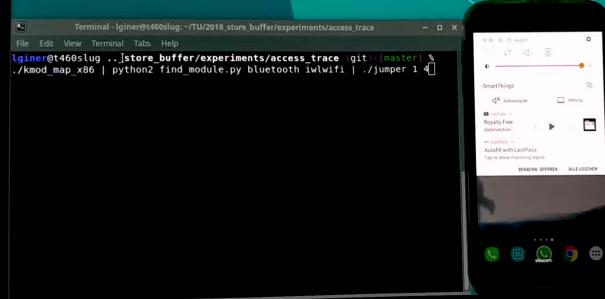


	Kernel		Modules	
			<b>®</b>	



lginer@t460slug ..\_store\_buffer/experiments/access\_trace (git)-[master] % ./kmod\_map\_x86

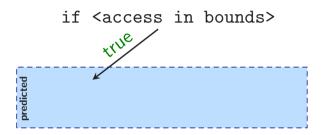
File Edit View Terminal Tabs Help



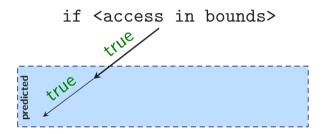
if <access in bounds>

predicted

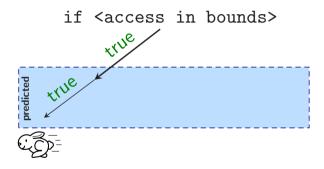




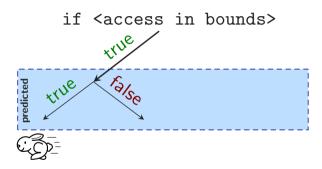




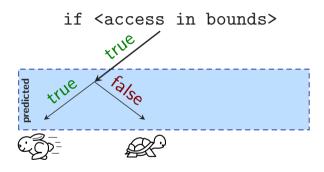




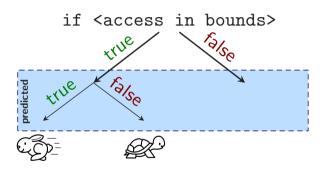




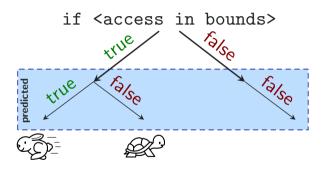




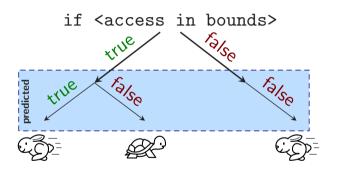




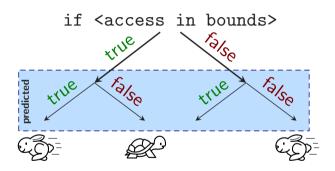




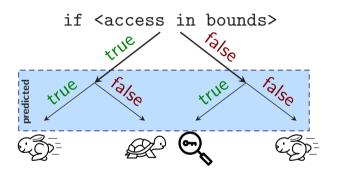


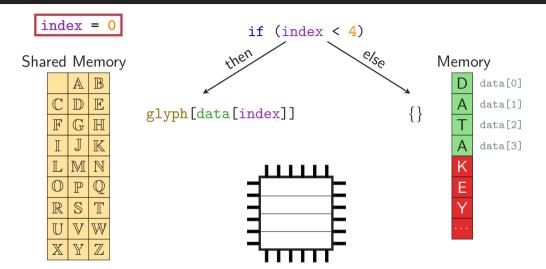


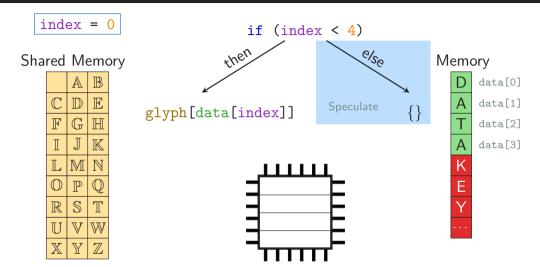




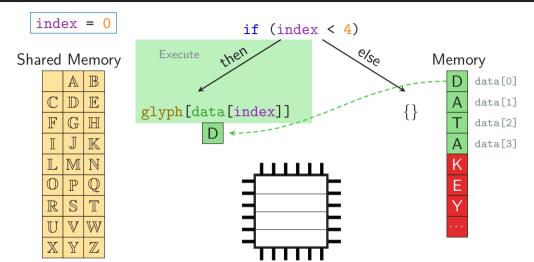


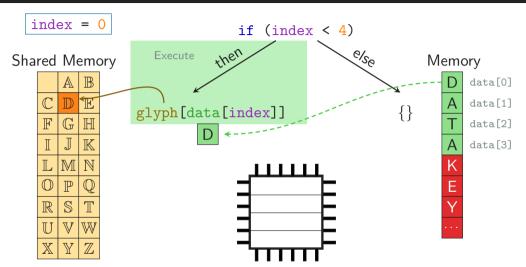




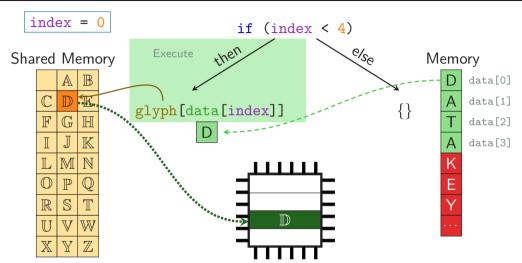


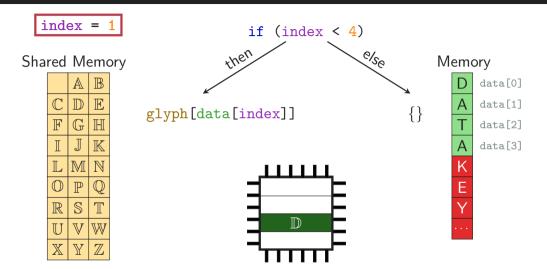


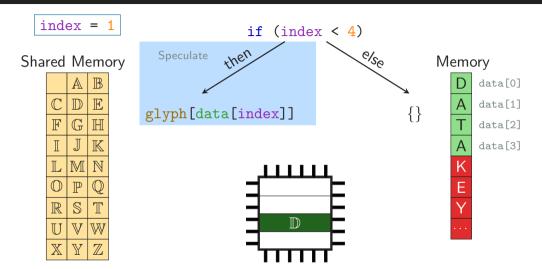




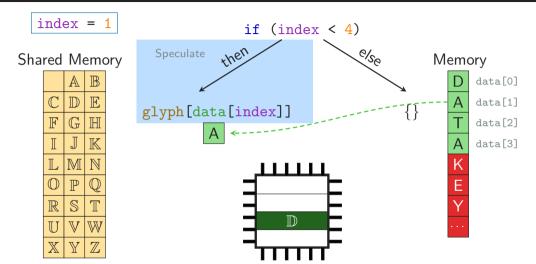




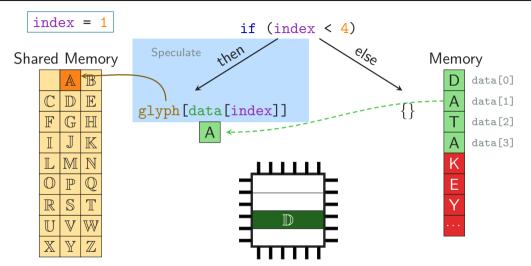




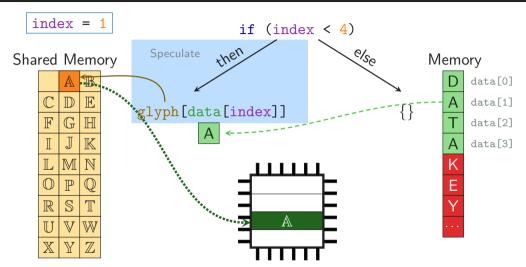


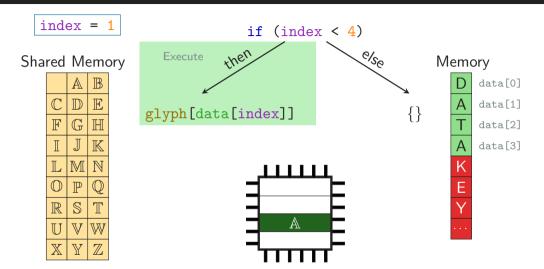


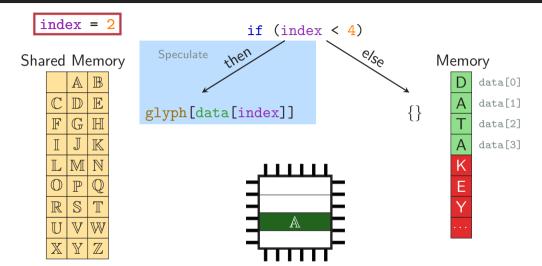


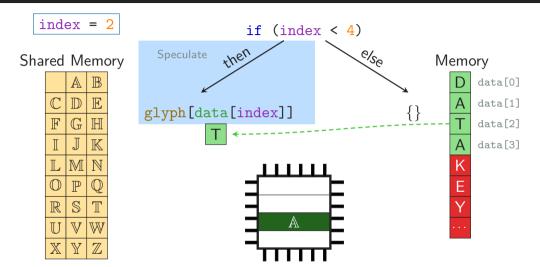


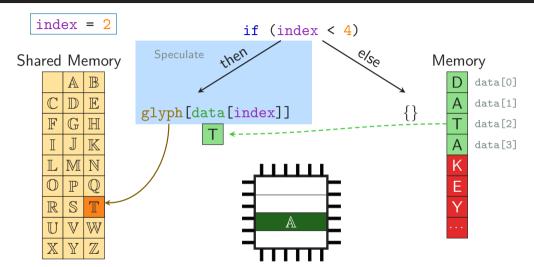




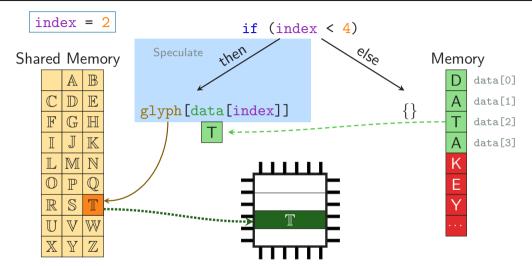




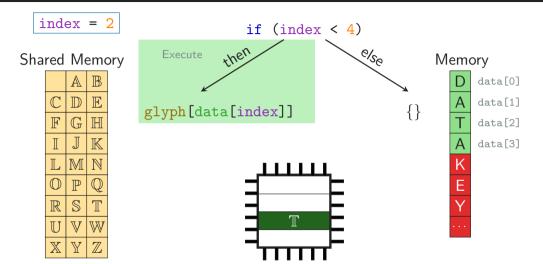


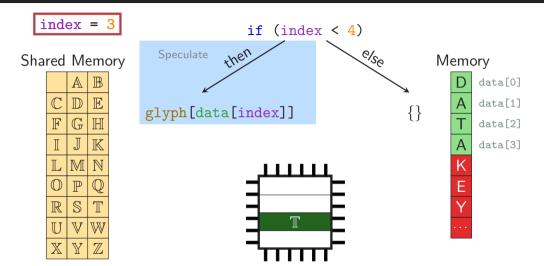


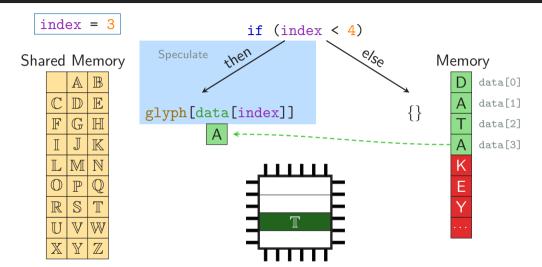


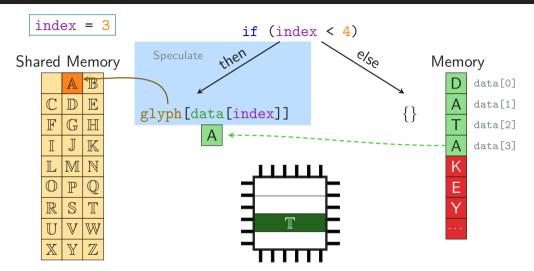




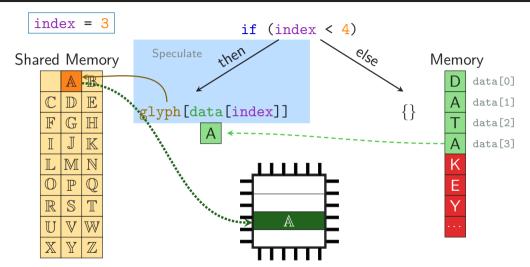


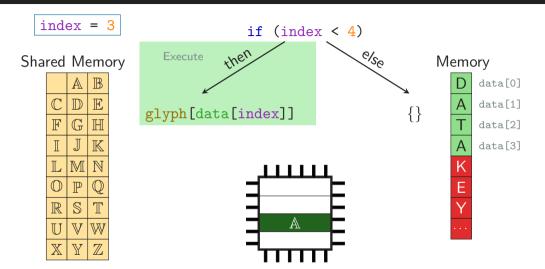




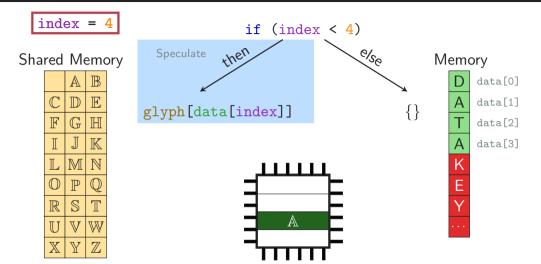


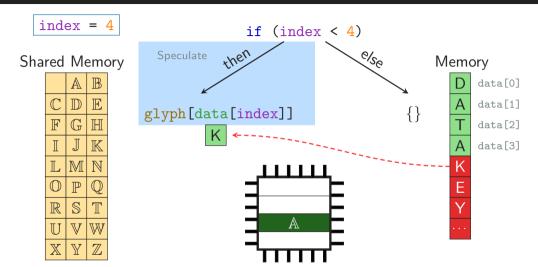


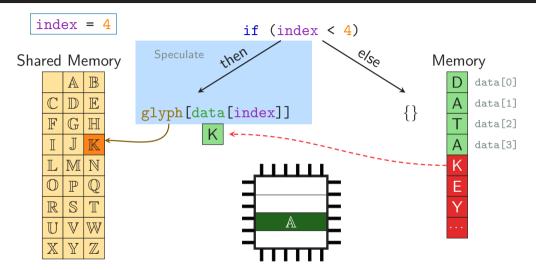




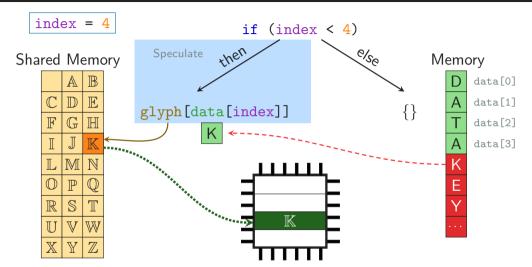


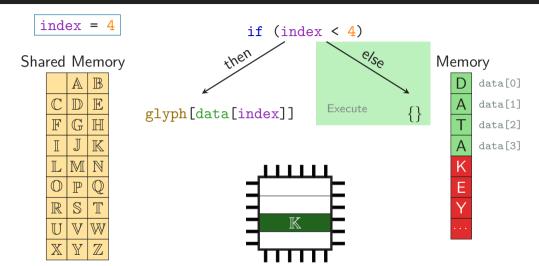






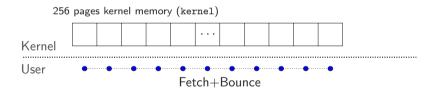






## **Speculative Fetch+Bounce**





### **Speculative Fetch+Bounce**



```
if (x < len(array))
    y = kernel[array[x] * 4096]

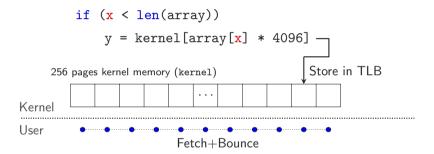
256 pages kernel memory (kernel)

Kernel

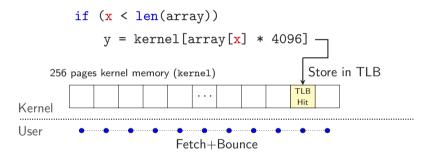
User

Fetch+Bounce</pre>
```

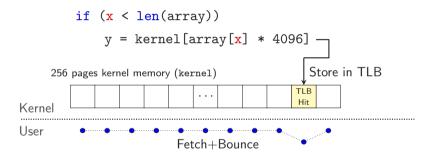














• Exploits deferred permission check in out-of-order execution



- Exploits deferred permission check in out-of-order execution
- Can read any kernel address



- Exploits deferred permission check in out-of-order execution
- Can read any kernel address
- Kernel maps all physical memory





- Exploits deferred permission check in out-of-order execution
- Can read any kernel address
- Kernel maps all physical memory
- → Read arbitrary memory by encoding in cache



### **Assumptions**

1. Stalling CPU might be too costly



#### **Assumptions**

- 1. Stalling CPU might be too costly
- 2. Stalling might require redesigning parts of CPU pipeline



#### **Assumptions**

- 1. Stalling CPU might be too costly
- 2. Stalling might require redesigning parts of CPU pipeline

#### **Hypothesis**

Load is executed, returned value is zeroed out on faults





Two tests:

1. Perform Meltdown attack

## Verifying the Hypothesis





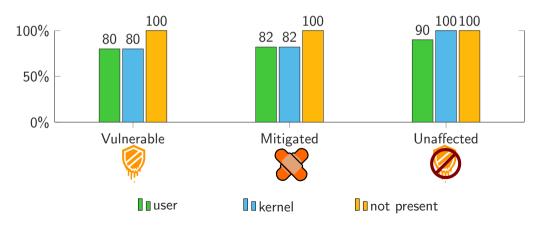
#### Two tests:

- 1. Perform Meltdown attack
- 2. Use Performance Counters

# Verifying the Hypothesis



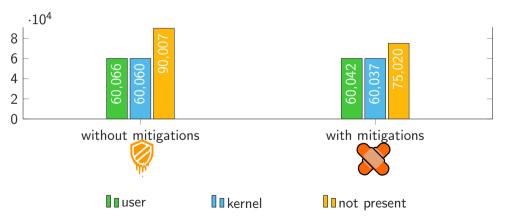
- Intel: CYCLE\_ACTIVITY.STALLS\_MEM\_ANY
- AMD: Dispatch Stalls



# Verifying the Hypothesis

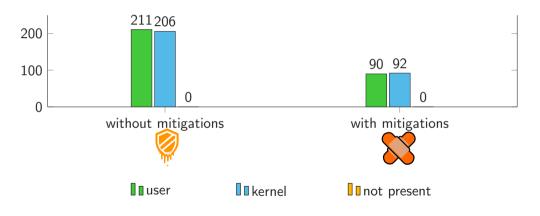


- Track number of issued  $\mu OPs$  on the load ports
- UOPS\_DISPATCHED\_PORT.PORT\_2, UOPS\_DISPATCHED\_PORT.PORT\_3



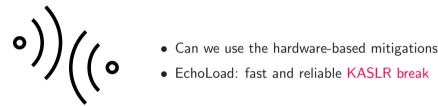


• L1D\_PEND\_MISS.PENDING\_CYCLES





 $\bullet\,$  Can we use the hardware-based mitigations for an attack?



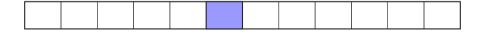
- Can we use the hardware-based mitigations for an attack?



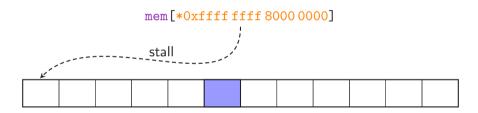
- Can we use the hardware-based mitigations for an attack?
- EchoLoad: fast and reliable KASLR break
- Encodes the returned value in the cache



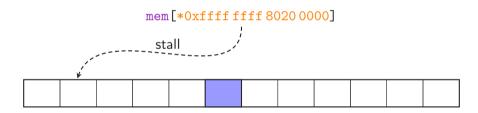
mem[\*0xfffffffff 8000 0000]



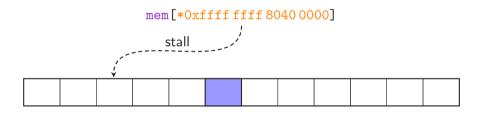








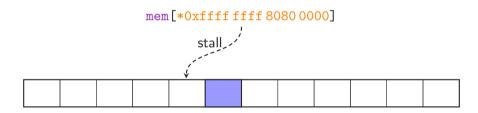




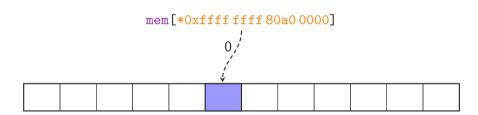












### Performance of EchoLoad



СРИ		Speculation		TSX	Segfault		
i7-6700K	Time (F-Score)	63 µs	(0.999)	48 µs	(1.000)	133 μs (1.000)	)
i9-9900K	Time (F-Score)	33 µs	(1.000)	29 µs	(1.000)	86 µs (1.000)	ļ
Xeon Silver 4208	Time (F-Score)	51 µs	(0.994)	40 µs	(1.000)	127 μs (1.000)	

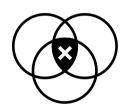
### Performance of EchoLoad



СРИ		Speculation		TSX	Segfault		
i7-6700K	Time (F-Score)	63 µs	(0.999)	48 µs	(1.000)	133 µs	(1.000)
i9-9900K	Time (F-Score)	33 µs	(1.000)	29 µs	(1.000)	86 µs	(1.000)
Xeon Silver 4208	Time (F-Score)	51 µs	(0.994)	40 µs	(1.000)	127 µs	(1.000)

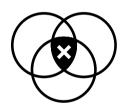
Works in SGX and JavaScript





Timing difference

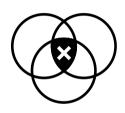




#### Timing difference

• between mapped and unmapped pages

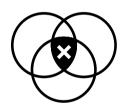




#### Timing difference

- between mapped and unmapped pages
- for different page sizes





#### Timing difference

- between mapped and unmapped pages
- for different page sizes
- between executable and non-executable pages

# Fake Load Address REsponse (FLARE)

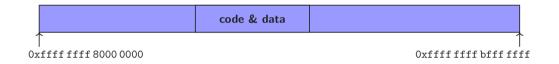




Current Linux design

# Fake Load Address REsponse (FLARE)

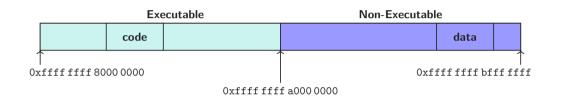




Step 1: Mitigating difference between mapped and unmapped pages

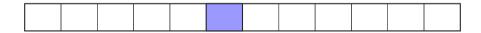
# Fake Load Address REsponse (FLARE)



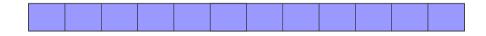


Step 2: Mitigating difference between executable and non-executable pages

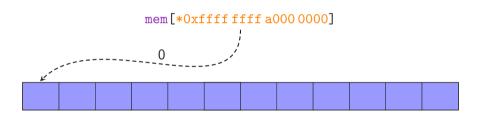




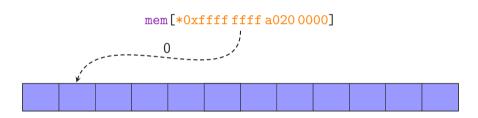




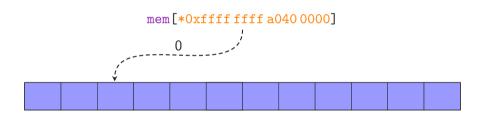




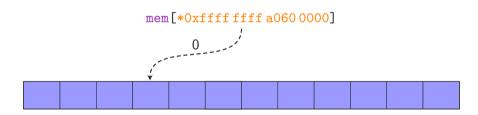




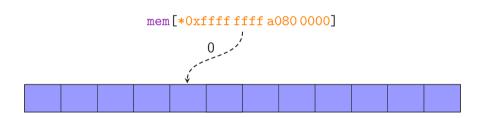




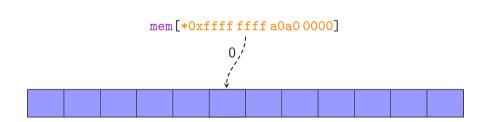




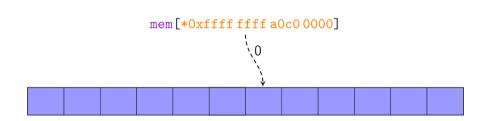




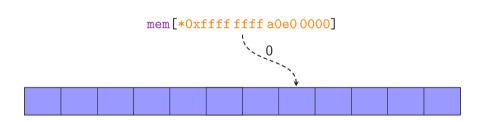




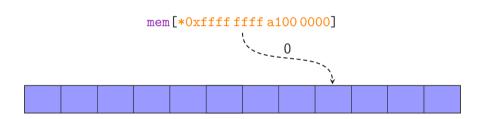




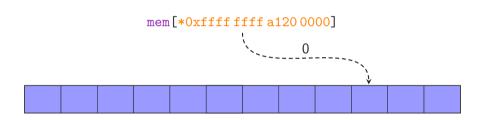




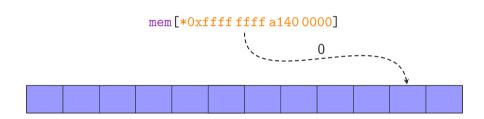




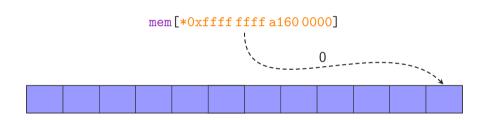


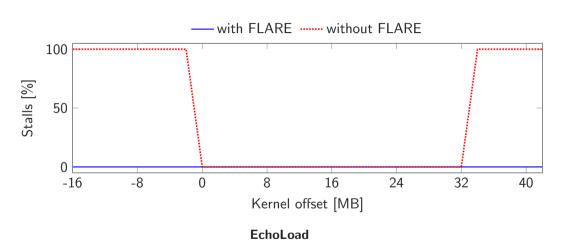


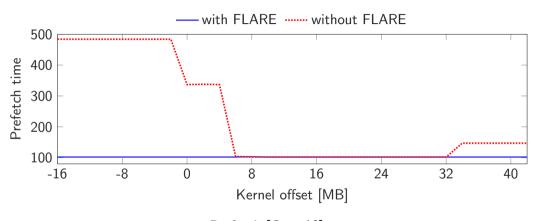




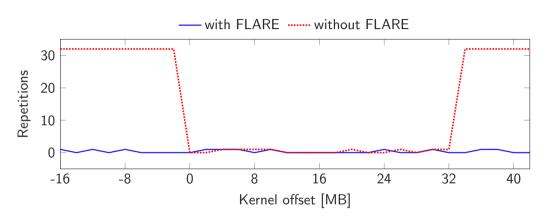




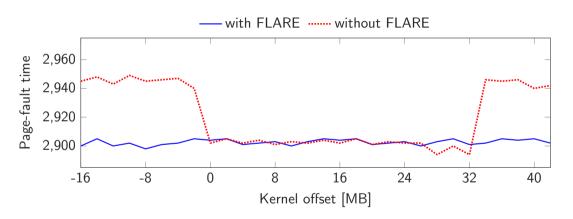




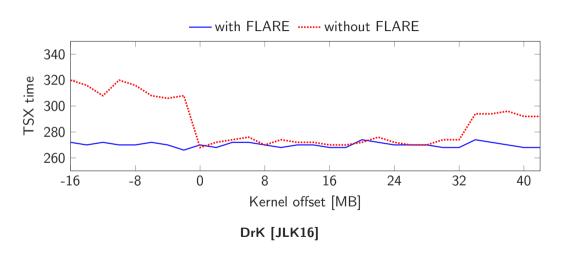
Prefetch [Gru+16]

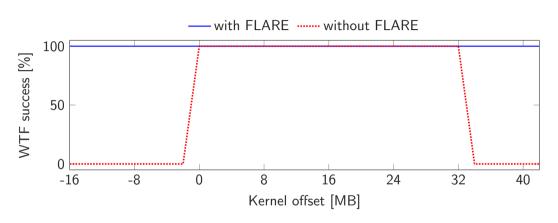


Data Bounce [Sch+19]



Double page fault [HWH13]





Fallout [Can+19]



You can find our proof-of-concept implementation of FLARE on:

• https://github.com/IAIK/FLARE



• Optimizations introduce security problems





- Optimizations introduce security problems
- Mitigations can overlook edge-cases



- Optimizations introduce security problems
- Mitigations can overlook edge-cases
- Once again requires software workaround





## Scoré-co-Léak Eorwarding

Chêrê and Back Âgan

Claudio Canella (@cc0x1f), Lukas Giner (@redrabbyte), Michael Schwarz (@misc0110) October 2, 2020

Graz University of Technology

CISPA Helmholtz Center for Information Security

#BHASIA @BLACKHATEVENTS

## References

C. Canella, D. Genkin, L. Giner, D. Gruss, M. Lipp, M. Minkin, D. Moghimi,
F. Piessens, M. Schwarz, B. Sunar, J. Van Bulck, and Y. Yarom. Fallout: Leaking
Data on Meltdown-resistant CPUs In: CCS 2019

- D. Gruss, C. Maurice, A. Fogh, M. Lipp, and S. Mangard. Prefetch Side-Channel Attacks: Bypassing SMAP and Kernel ASLR. In: CCS. 2016.
- R. Hund, C. Willems, and T. Holz. Practical Timing Side Channel Attacks against Kernel Space ASLR. In: S&P. 2013.
- Y. Jang, S. Lee, and T. Kim. Breaking Kernel Address Space Layout Randomization with Intel TSX. In: CCS. 2016.
  - M. Schwarz, C. Canella, L. Giner, and D. Gruss. Store-to-Leak Forwarding: Leaking Data on Meltdown-resistant CPUs. In: arXiv:1905.05725 (2019).