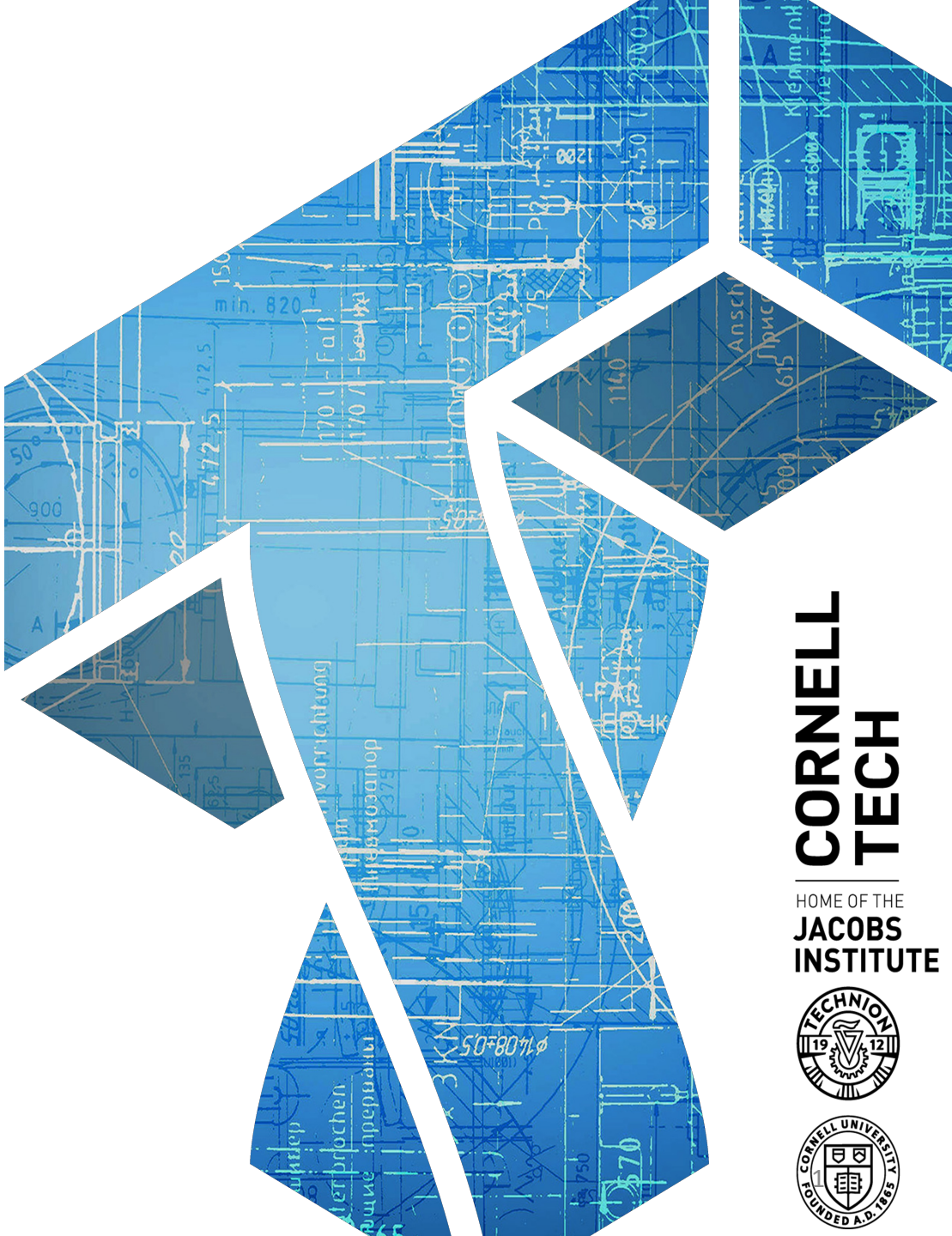
CS 5435: Containment & isolation

Instructor: Tom Ristenpart

https://github.com/tomrist/cs5435-fall2024



CORNELL TECH

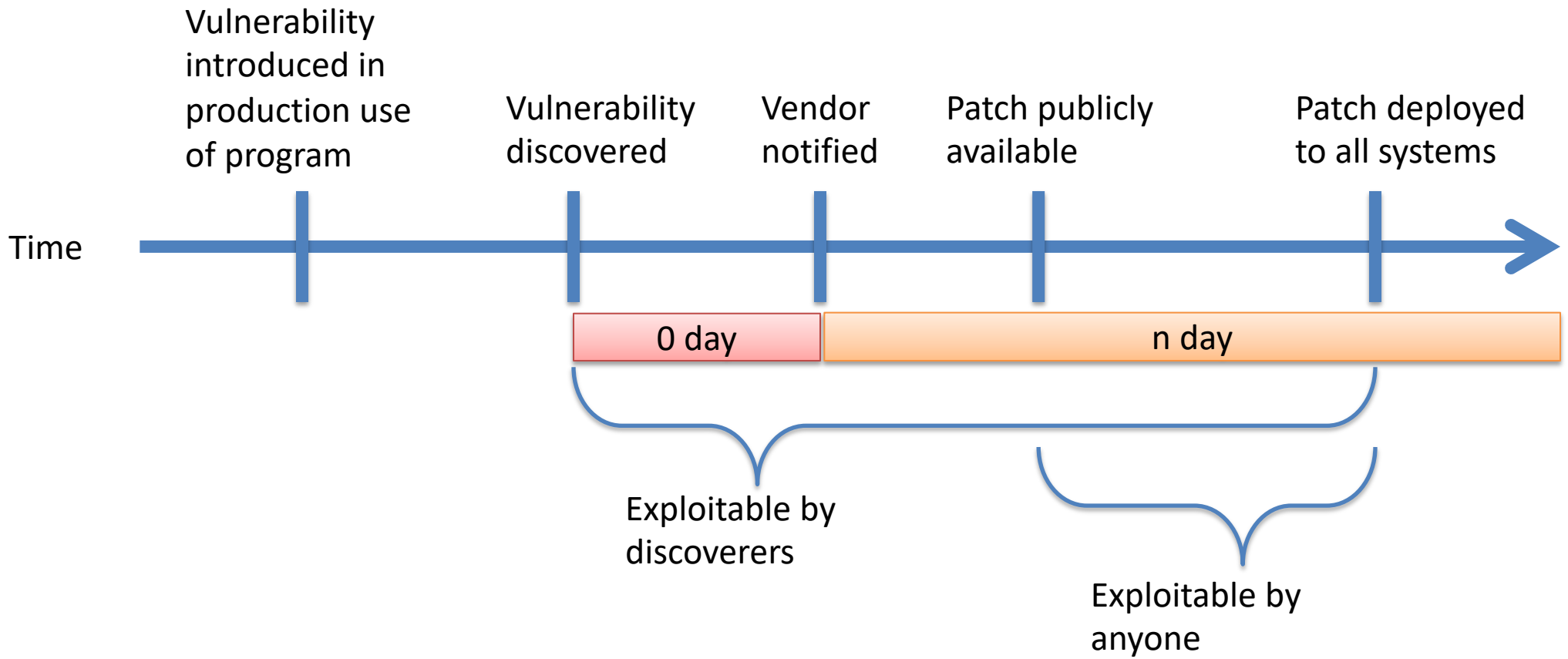
HOME OF THE JACOBS INSTITUTE

Vulnerabilities happen…

- Countermeasures to prevent exploitation
  - ASLR, canaries, W^X
- Discovering and patching vulnerabilities
  - Manual analysis, static analysis, fuzzing,

Patching and vulnerability lifecycles

Imagine an exploitable vulnerability in some widely used software

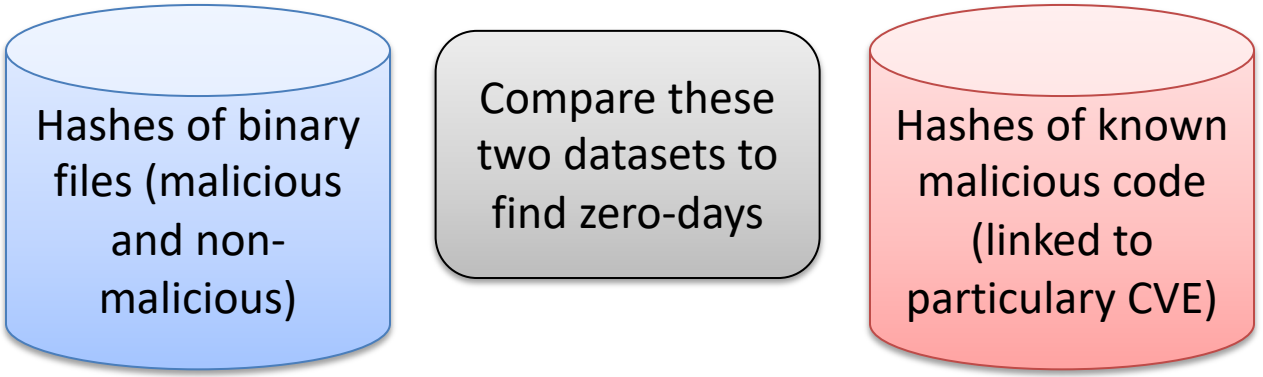


Studying zero days

- How can we study zero day exploitation?
- [Bilge, Dumitras 2012] study

Symantec

Gathers samples of binaries (with consent) from endpoint hosts

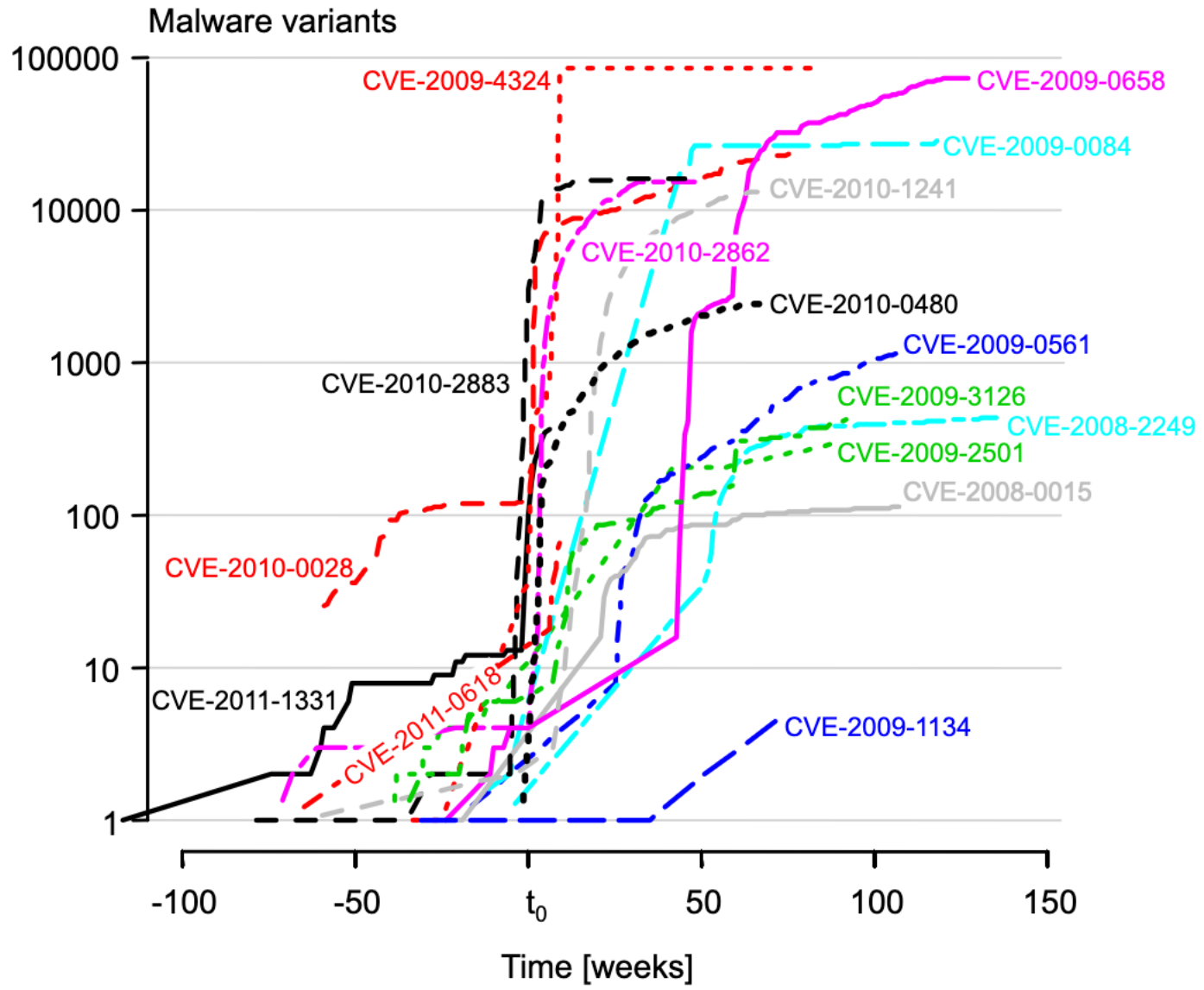


Includes date at which file observed on endpoint device

Studying zero days

- How can we study zero day exploitation?
- [Bilge, Dumitras 2012] study
  - Identified 18 zero-days, 11 new zero-days
  - Typical zero day attack lasts 312 days
  - Exploitation goes up by factor of 10^5 after public release

Studyin



(b) Malware variants exploiting zero-day vulnerabilities before and after disclosure (time = t_0).

Zero day markets

- Bug bounty programs vs. zero day markets
  - Pwn2own competition by Google
- Companies monetizing zero-days by selling them to governments

"We wouldn't share this with Google for even $1 million," says Bekrar. "We don't want to give them any knowledge that can help them in fixing this exploit or other similar exploits. We want to keep this for our customers."

- Vupen CEO

TECH • SECURITY

Meet The Hackers Who Sell Spies The Tools To Crack Your PC (And Get Paid Six-Figure Fees)

Andy Greenberg Former Staff

Covering the worlds of data security, privacy and hacker culture.

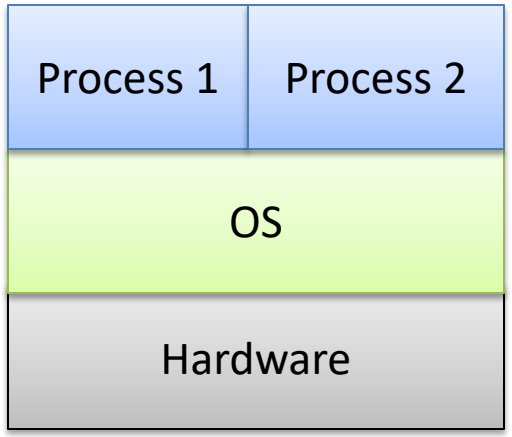
Mar 21, 2012, 09:08am EDT

Exploitable vulnerabilities aren’t going away.
How to do we limit their ill-effects?

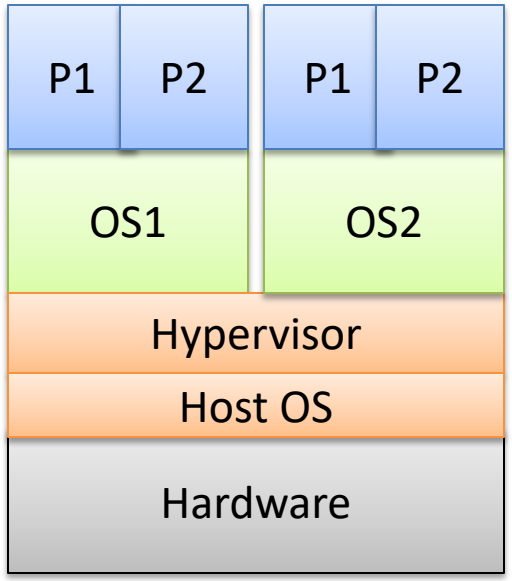
Containment and isolation

- We can try to use software/hardware mechanisms to contain exploited programs
- File system jails early example:
  - chroot: change apparent root directory for a process
  - jail (in FreeBSD): chroot + individual hostname/IP/root
- Containers
  - OS isolation of different process families
- Virtual machines
  - Different guest operating systems and programs isolated by hypervisor

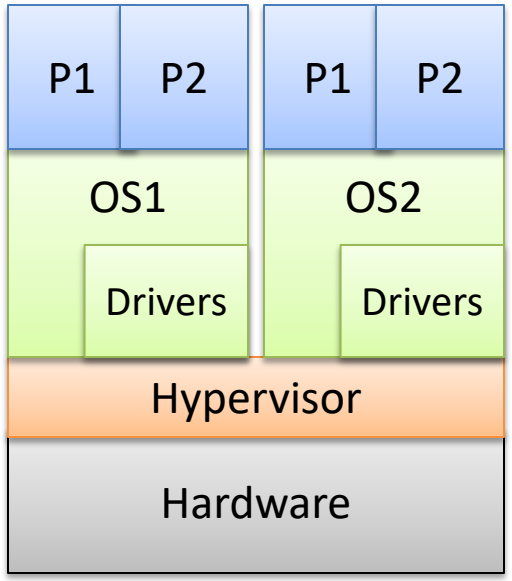
Virtualization



No virtualization



Full virtualization + Type 2



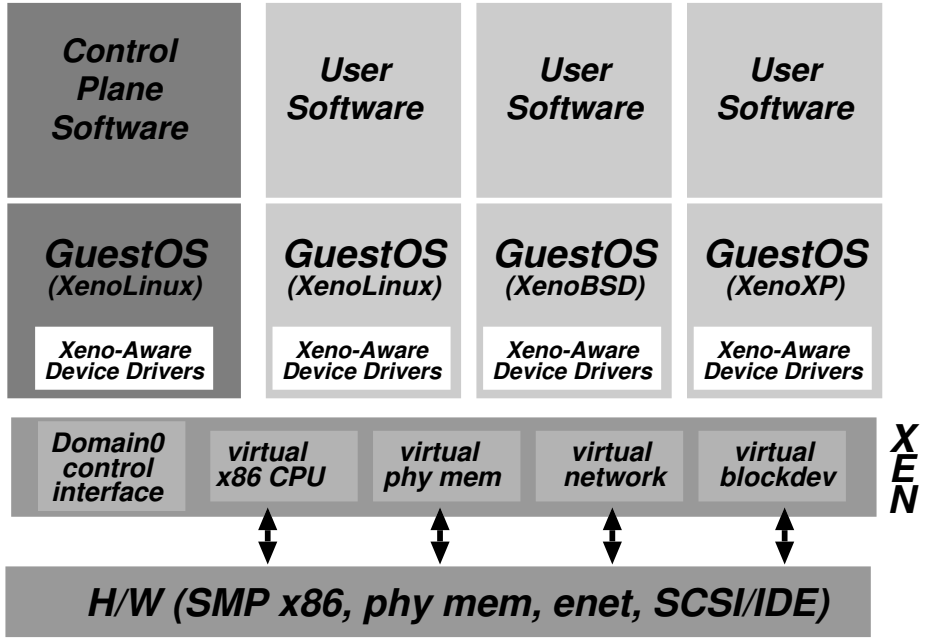
Paravirtualization + Type 1

Type-1: Hypervisor runs directly on hardware
Type-2: Hypervisor runs on host OS

Full virtualization: unmodified guest OS
Paravirtualization: modified guest OS

Xen

- 2003: academic paper
  - "Xen and the Art of Virtualization"
- Paravirtualization
  - Modified guest OS
  - Hypercalls (like system calls but for guest OS) to hypervisor
  - Each guest given 1 or more VCPUs
- Why paravirtualization?
  - Performance improvements
- VMWare, VirtualBox, HyperV

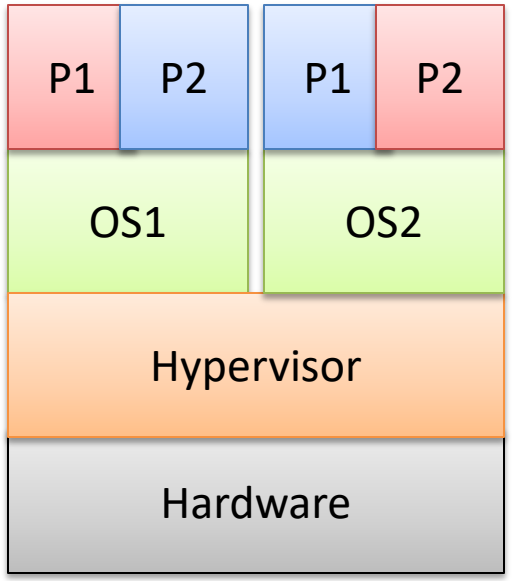


Example VM Use Cases

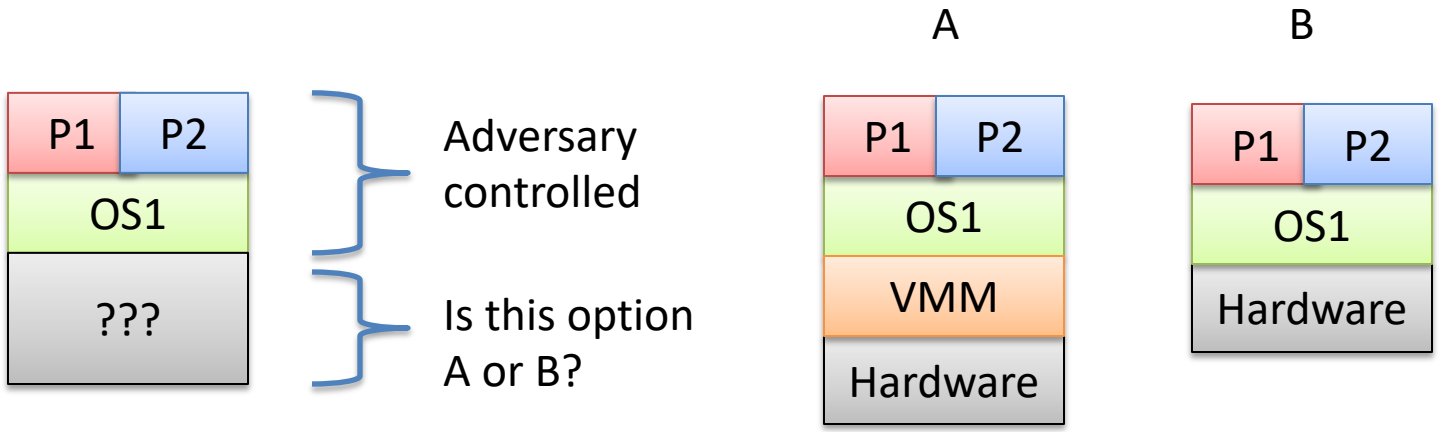
- Legacy support (e.g., IBM VM/370 from 1970s)
- Development
- Server consolidation
- Sandboxing / containment
- Cloud computing Infrastructure-as-a-Service

Study of malware

- Researchers use VMs to study malware
- Example of VM sandboxing
  - Hypervisor must contain malicious code
- How would you evade analysis as a malware writer?
  - split personalities



VMM Transparency



- Adversary can detect if:
  - Paravirtualization
  - Logical discrepancies
    - Expected CPU behavior vs virtualized
    - Red pill (Store Interrupt Descriptor Table instr)
  - Timing discrepancies
    - Slower use of some resources

Garfinkel et al. "Compatibility is not transparency: VMM Detection Myths and Reality"

A VMWare detection approach

```
MOV EAX,564D5868 <-- "VMXh"
MOV EBX,0
MOV ECX,0A
MOV EDX,5658 <-- "VX"
IN EAX,DX <-- Check for VMWare
CMP EBX,564D5868
```

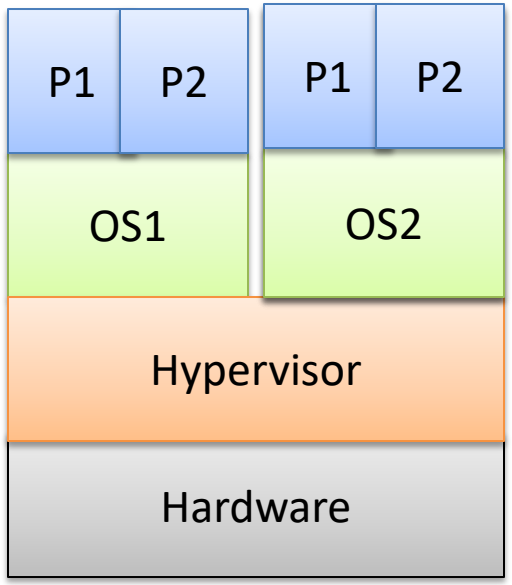
IN instruction used by VMWare to facilitate host-to-guest communication

VMWare:
places VMXh in EBX
Physical:
processor exception

From
http://handlers.sans.org/tliston/ThwartingVMDetection_Liston_Skoudis.pdf

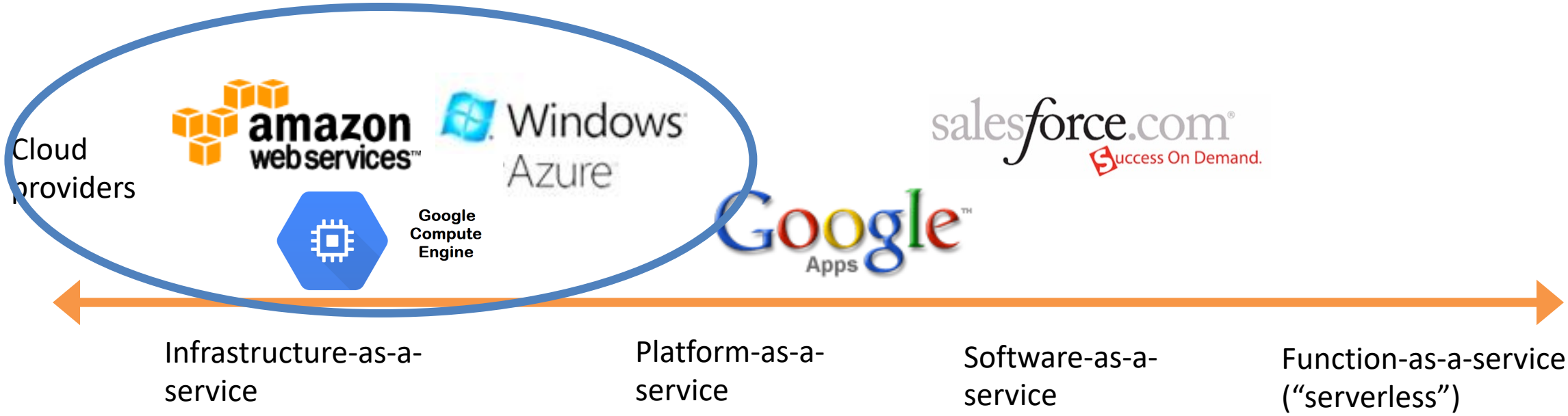
Server consolidation

- Consolidation
  - Use VMs to optimize use of hardware
  - Pack as many VMs onto each server as possible
  - Turn off other servers
- Threat model?
  - Containment
  - Isolation
  - Assume guests are/can be compromised



Cloud computing

NIST: Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.



A simplified model of public IaaS cloud computing

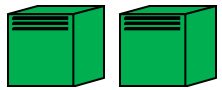
Users run Virtual Machines (VMs) on cloud provider's infrastructure

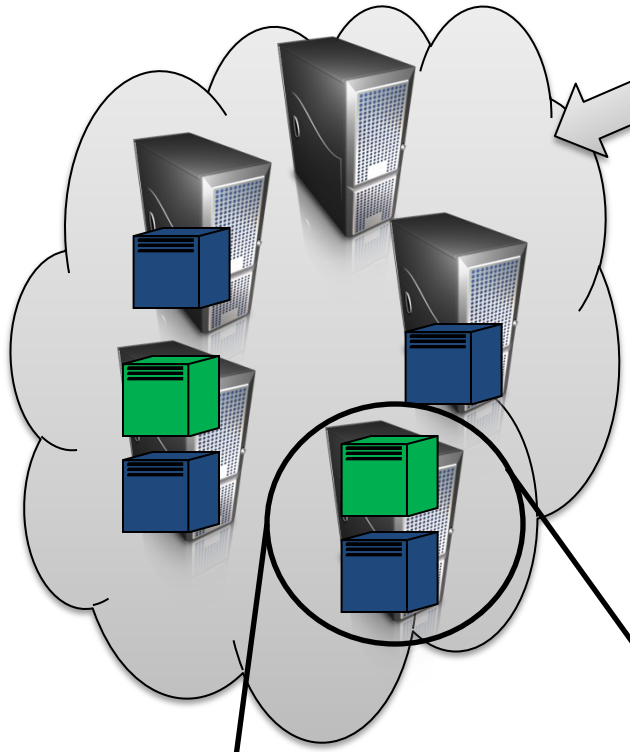
User A

virtual machines (VMs)

User B

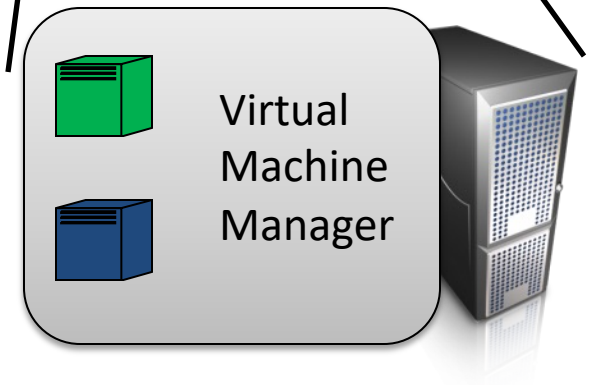
virtual machines (VMs)





Owned/operated by cloud provider

Google Compute Engine

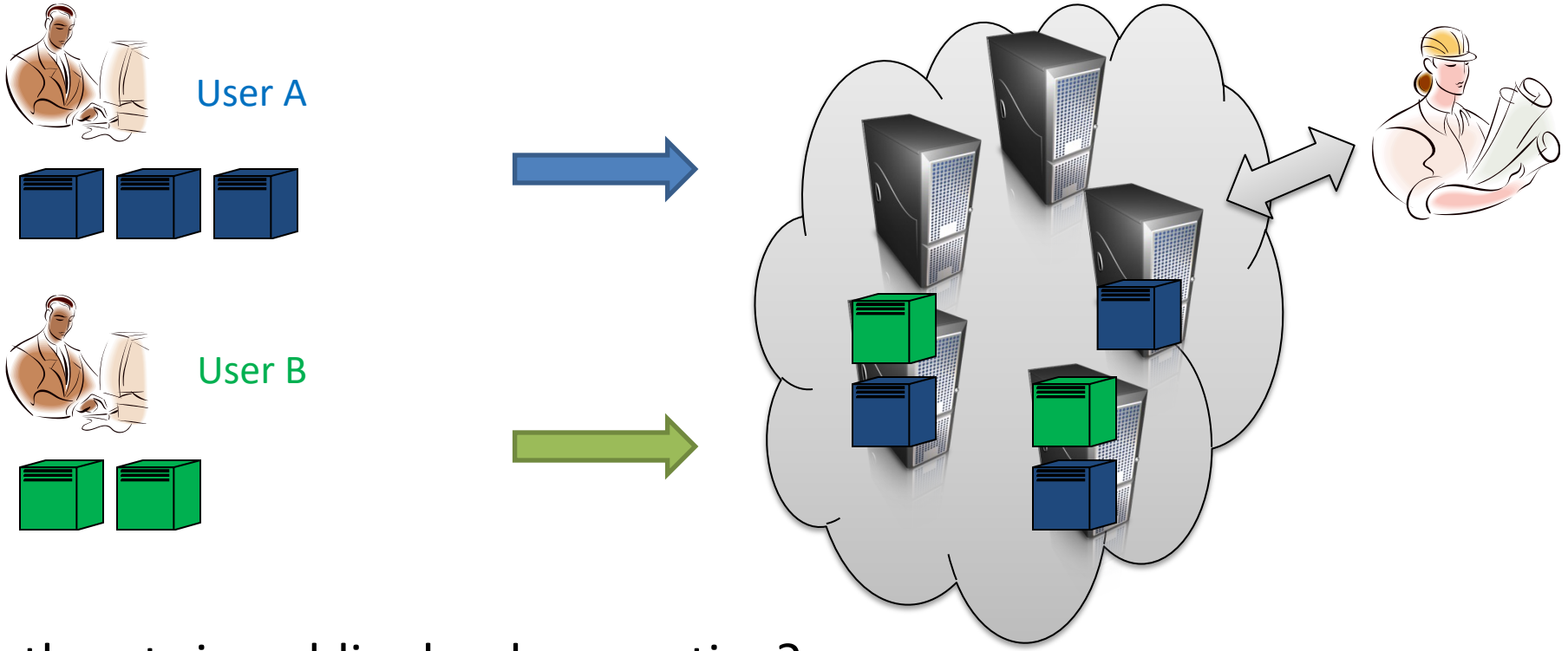


Multitenancy (users share physical resources)

Virtual Machine Manager (VMM)
manages physical server resources for VMs

To the VM should look like dedicated server

Trust models in public cloud computing



Security threats in public cloud computing?

Provider spying on running VMs / data

External attacks against infrastructure

Cross-user attacks

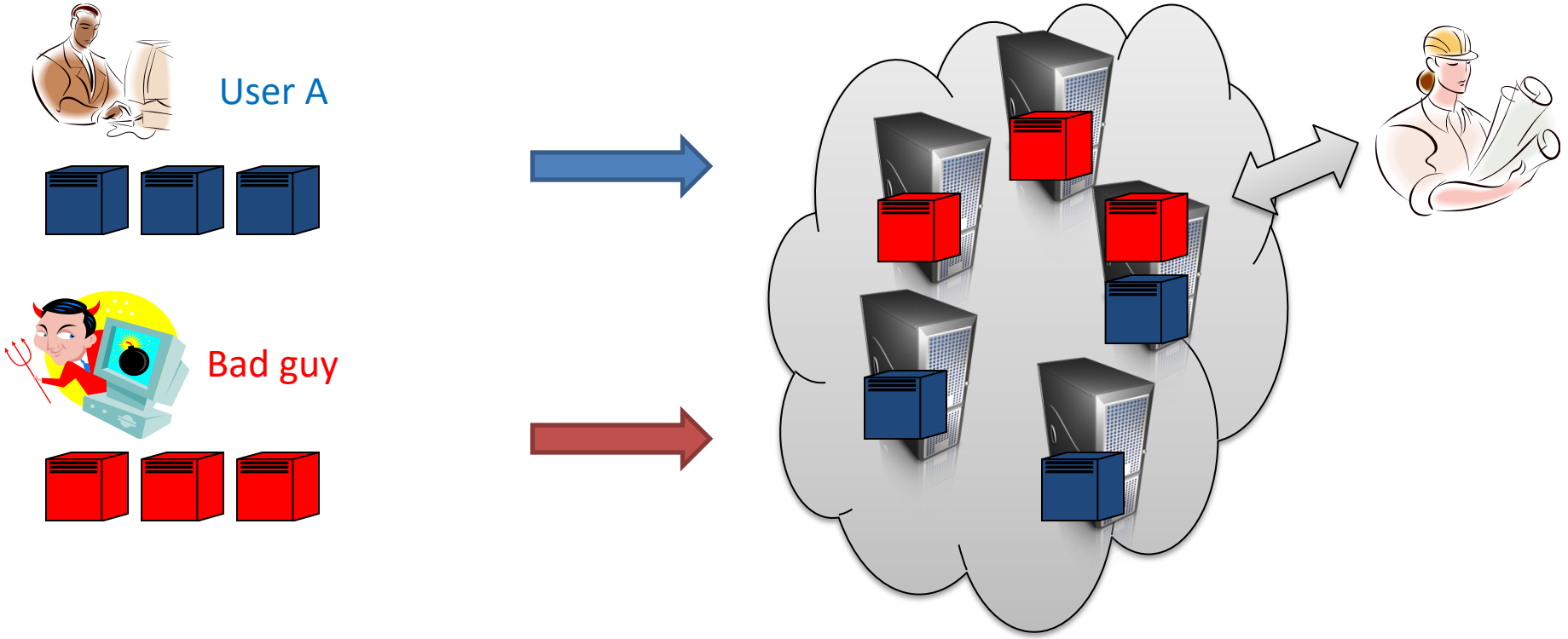
Escape-from-VM

Side-channels attacks

Degradation of service attacks

Resource-stealing attacks

Trust models in public cloud computing



Attacker identifies one or more victims VMs in cloud

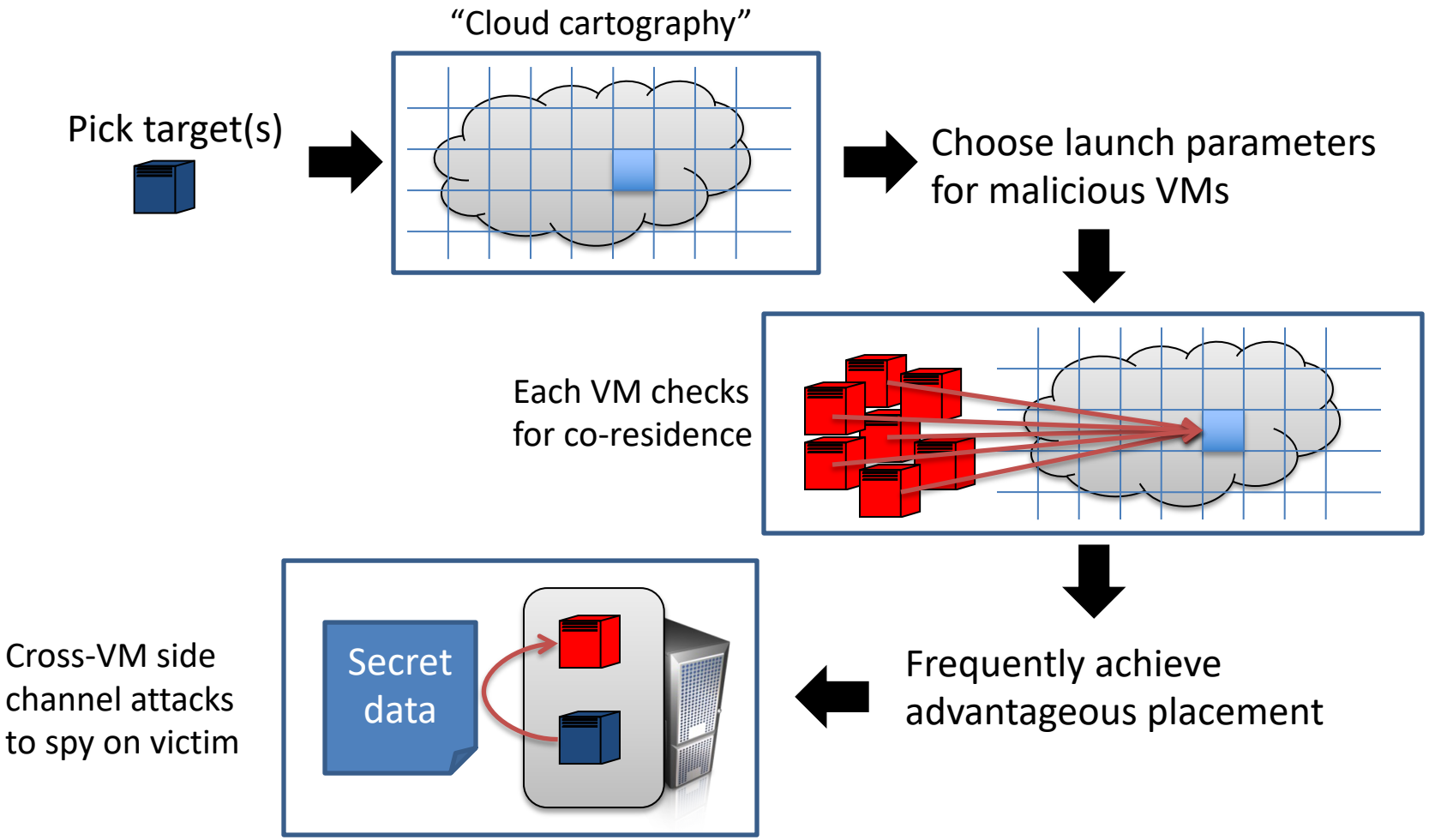
1) Achieve advantageous placement via launching of VM instances

2) Launch attacks using physical proximity

Exploit VMM vulnerability DoS Side-channel attack

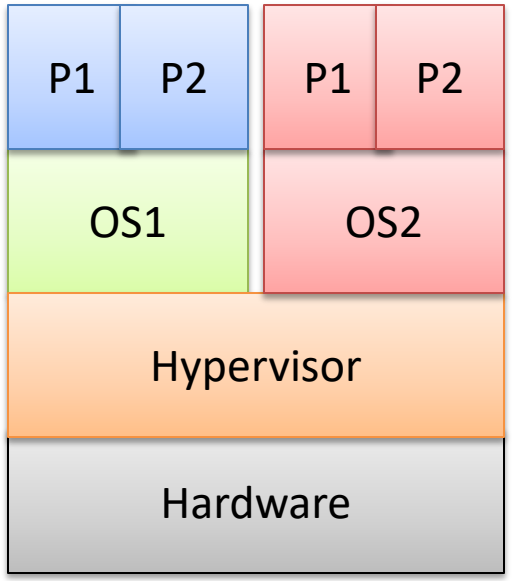
2009 case study with Amazon's EC2

Adversary able to:



Violating containment

- Escape-from-VM
  - Vulnerability in VMM or host OS (e.g., Dom0)
- Memory management flaws in VMM



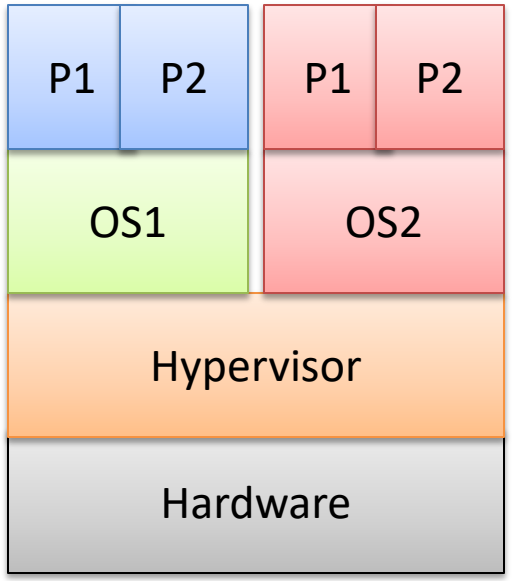
Zero-Day Exploit Published for VM Escape Flaw in VirtualBox

by Lucian Constantin on November 8, 2018

A security researcher disclosed a yet unpatched zero-day vulnerability in the popular VirtualBox virtualization software that can be exploited from a guest operating system to break out of the virtual machine and gain access to the host OS.

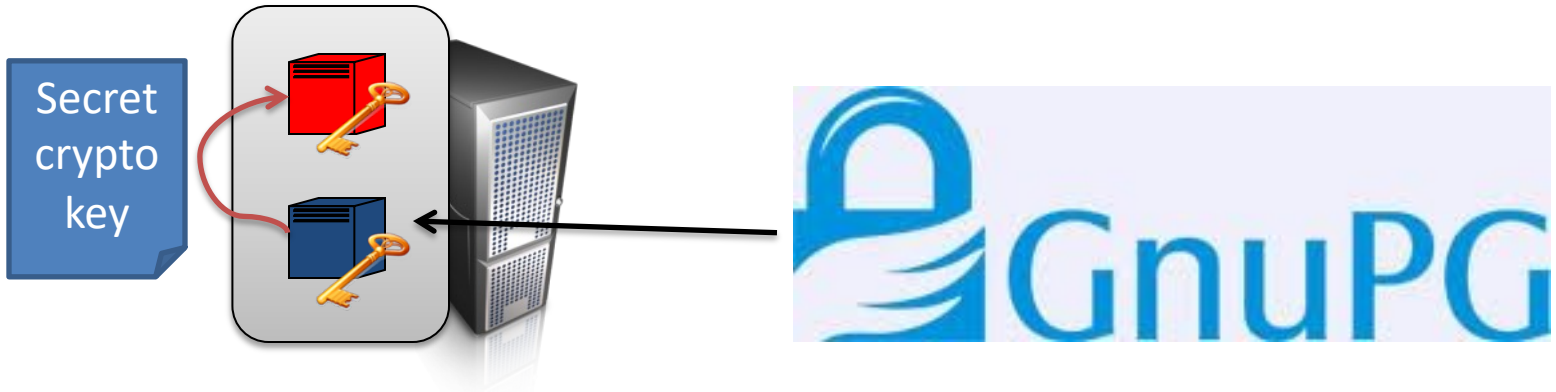
Violating isolation

- Covert channels between VMs circumvent access controls
  - Bugs in VMM
  - Side-effects of resource usage
- Degradation-of-Service attacks
  - Guests might maliciously contend for resources
  - Xen scheduler vulnerability
- Side channels
  - Spy on other guest via shared resources



Cross-VM cryptographic side-channel attacks

[Zhang, Juels, Reiter, R. – CCS `12]



Target is 4096-bit ElGamal secret key e

Modular Exponentiation (x, e, N):

let $e_n \ldots e_1$ be the bits of e

$y \leftarrow 1$

for e_i in $\{e_n \ldots e_1\}$

- $y \leftarrow$ **Square**(y) **(S)**
- $y \leftarrow$ **Reduce**(y, N) **(R)**
- if $e_i = 1$ then
  - $y \leftarrow$ **Multi**(y, x) **(M)**
  - $y \leftarrow$ **Reduce**(y, N) **(R)**

return y // $y = x^e \bmod N$

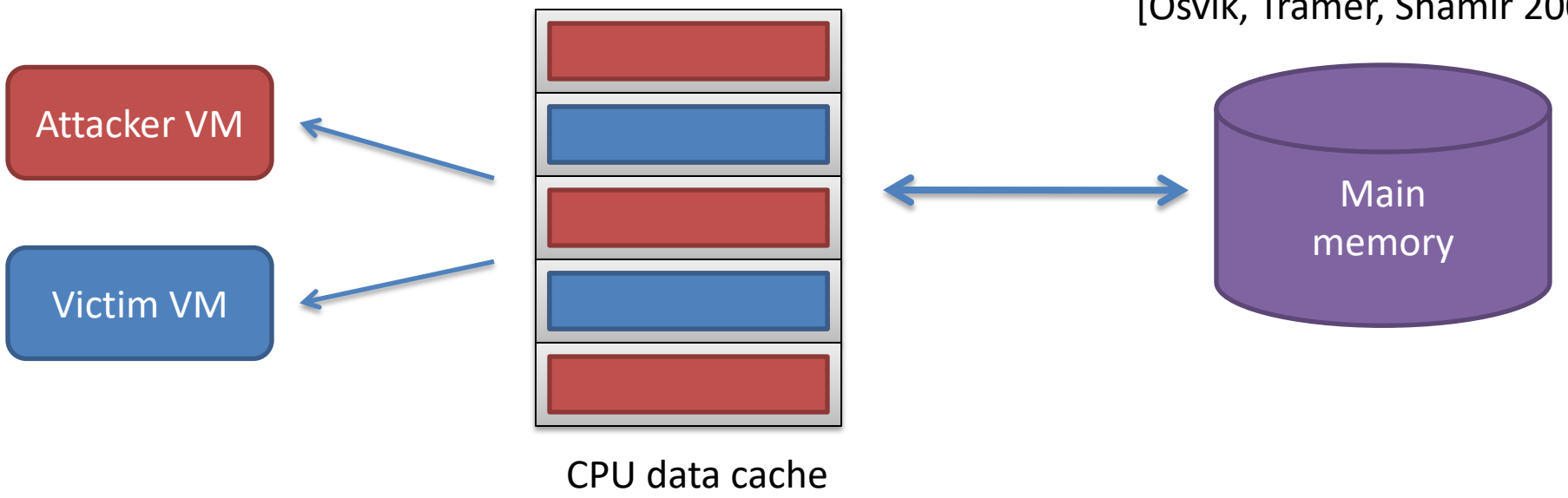
$e_i = 1 \rightarrow$ "SRMR"

$e_i = 0 \rightarrow$ "SR"

Sequence of function calls reveals secret key

Cross-VM side channels using CPU cache contention

Introduced in cross-process setting by [Osvik, Tramer, Shamir 2006]



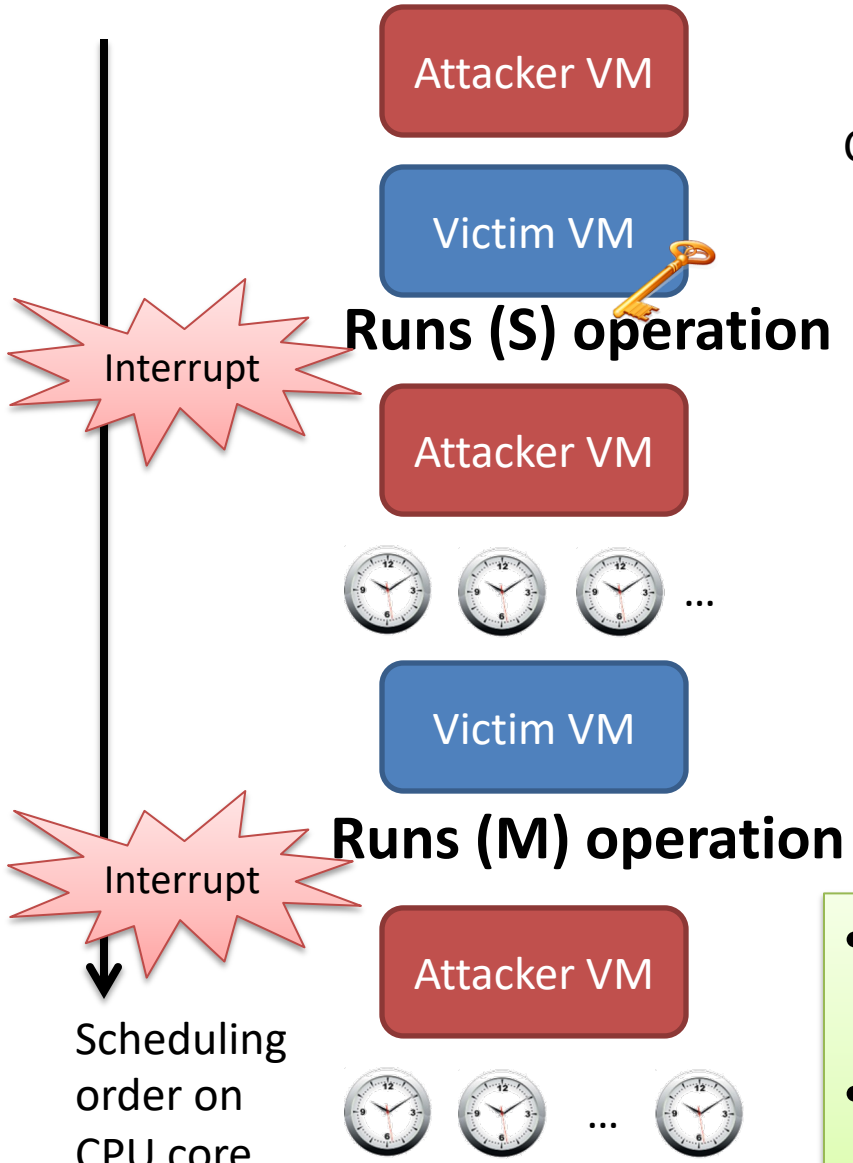
1) Read in a large array (fill CPU cache with attacker data)

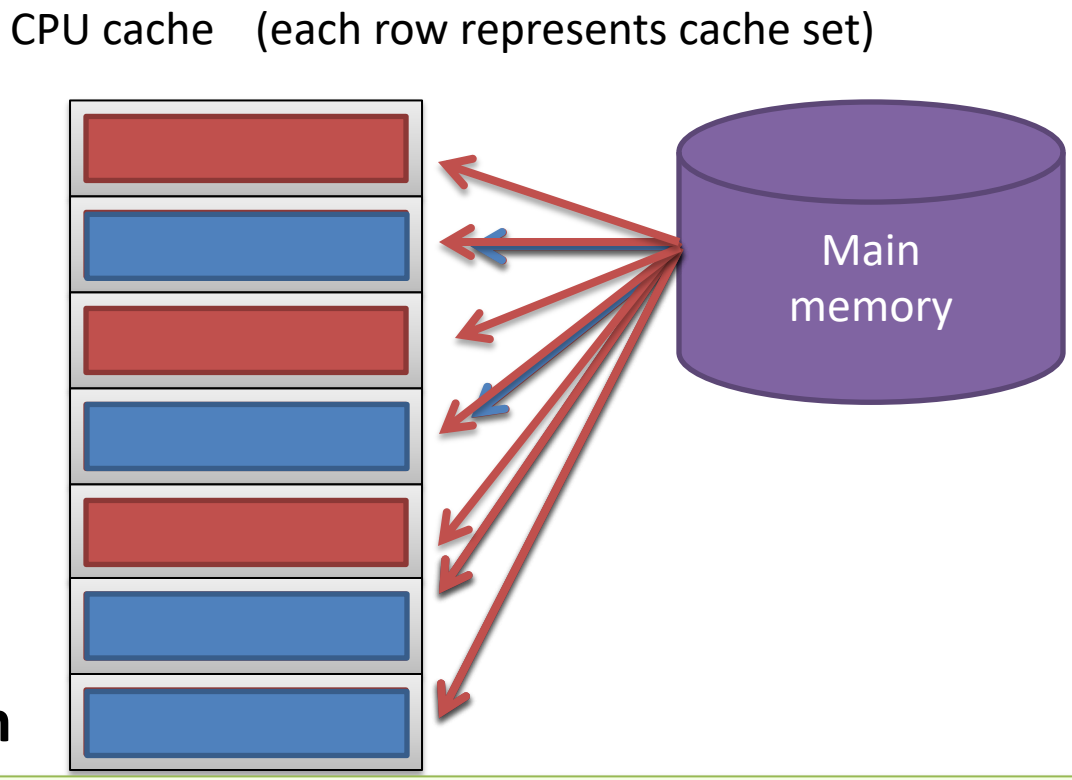
2) Busy loop (allow victim to run)

3) Measure time to read large array (the load measurement)

Locations in cache occupied by victim will take longer to load → Information about victim's use of cache revealed to attacker

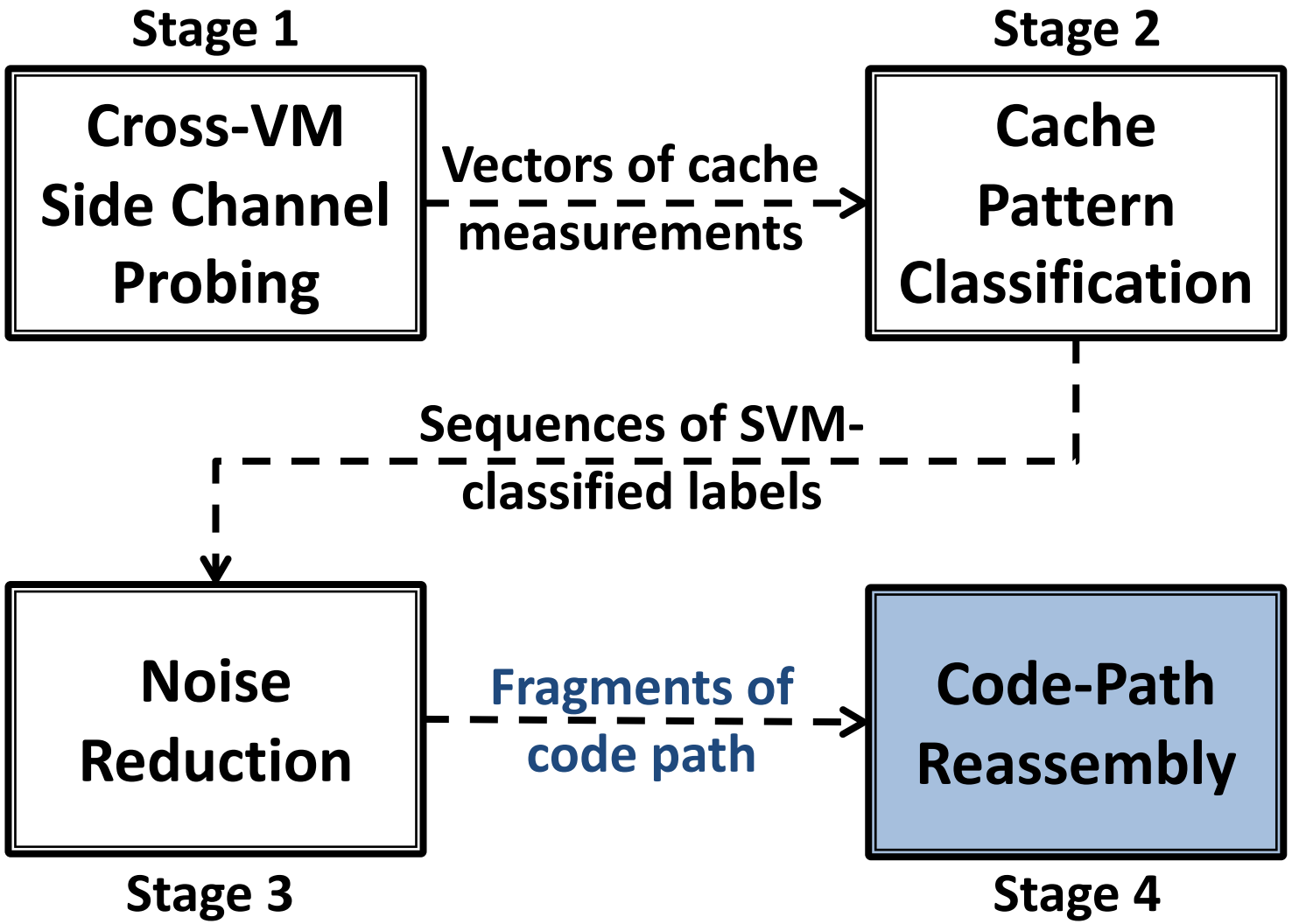
Prime+Probe protocol





- Timings correlated to (distinct) cache usage patterns of S, M operations
- Can spy frequently (every ~16 µs) by exploiting scheduling

Stages of the prototype attack



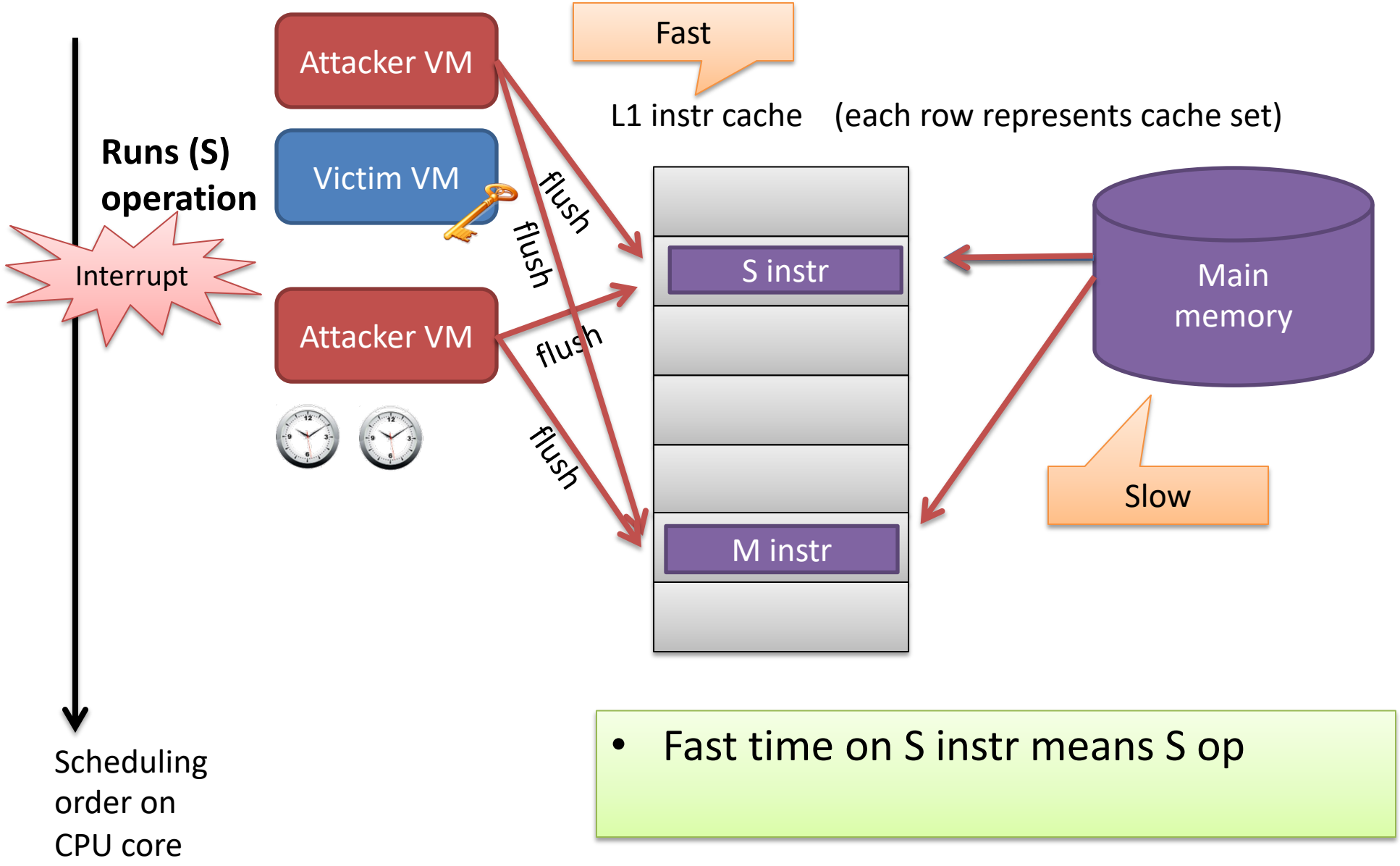
Results

- Setup for in-lab experimentation:
  - Intel Yorkfield processor (4 cores, 32KB L1 instruction cache)
  - Xen + linux + GnuPG + libgcrypt
- **Best result:**
  - 300,000,000 prime-probe results (6 hours)
  - Over 300 key fragments
  - Brute force the secret key in ~9800 guesses
- Not practical to run the attack in deployment settings

Microarchitectural side-channels

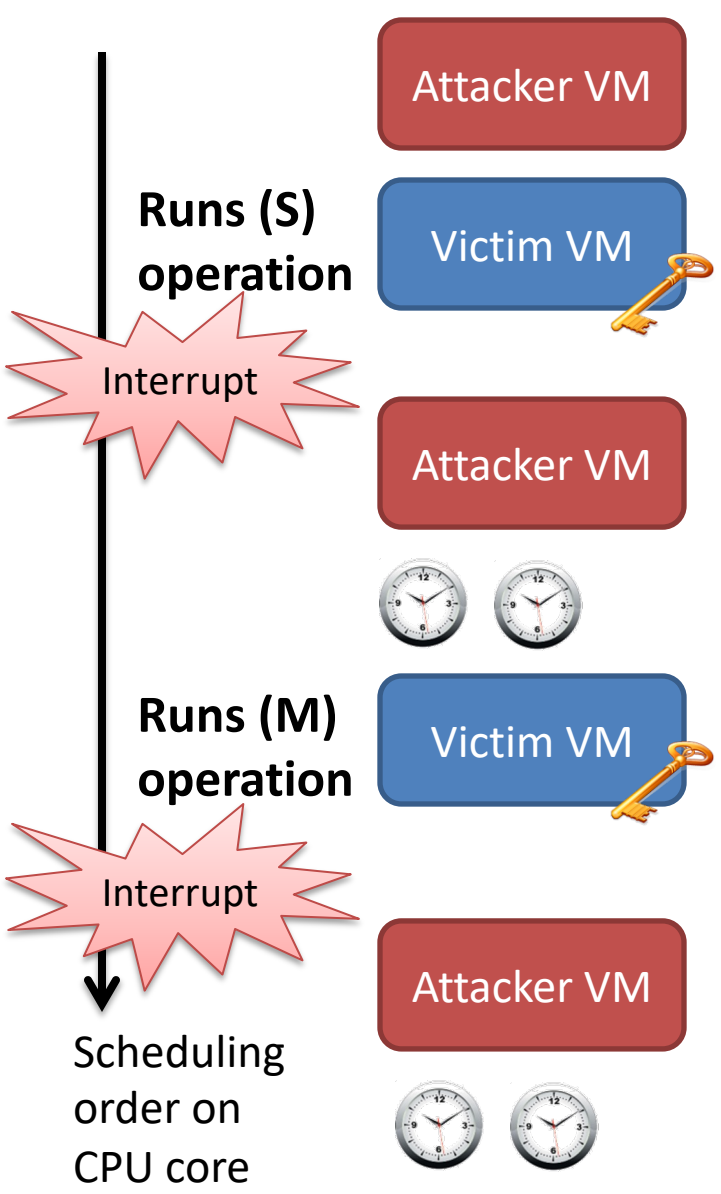
- Lots of research on exploiting microarchitectural side-channels
- State-of-the-art Prime+Probe attacks:
  - Sinan Inci et al. 2016 “Cache Attacks Enable Bulk Key Recovery on the Cloud”
- Flush+Reload (F+R) more robust side-channel in shared memory settings [Yarom, Falkner 2013]
  - Spy process flushes memory shared with victim from caches
  - Idles
  - Times how long it takes to read shared memory

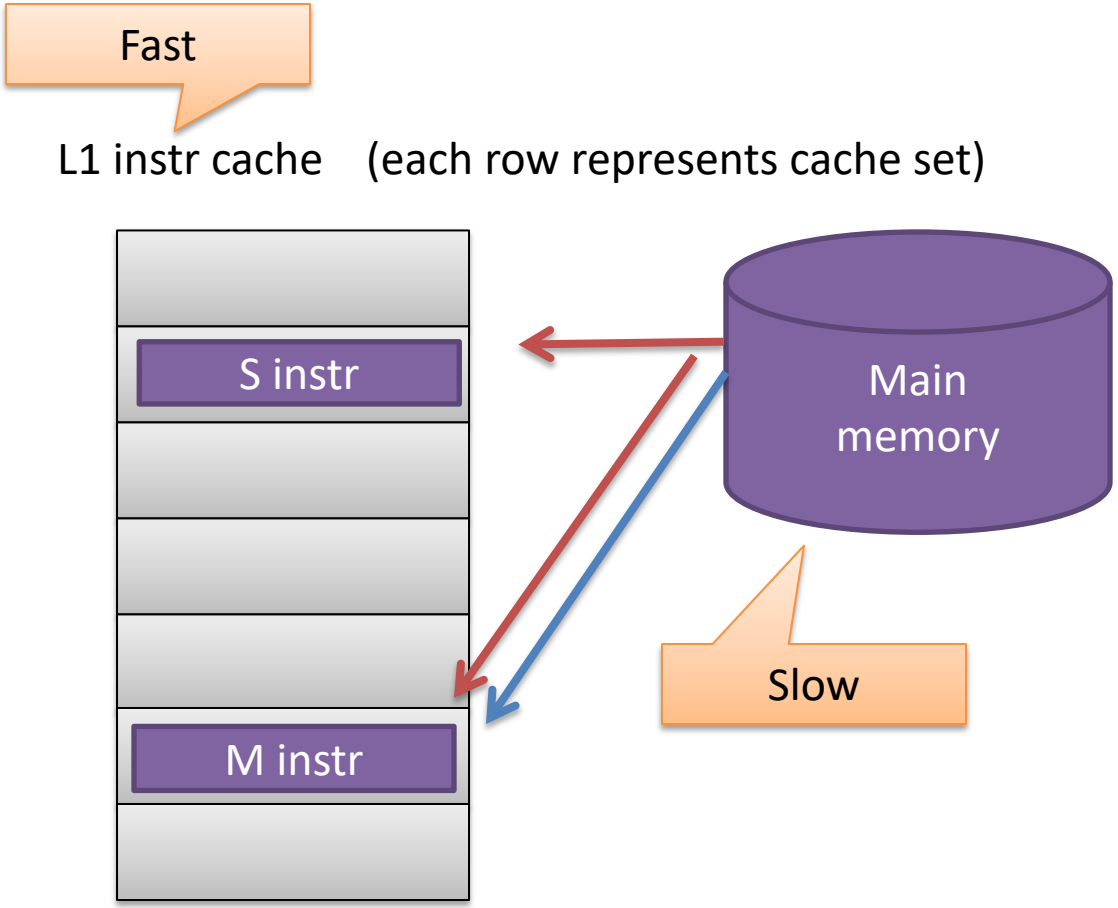
Flush+Reload protocol



- Fast time on S instr means S op

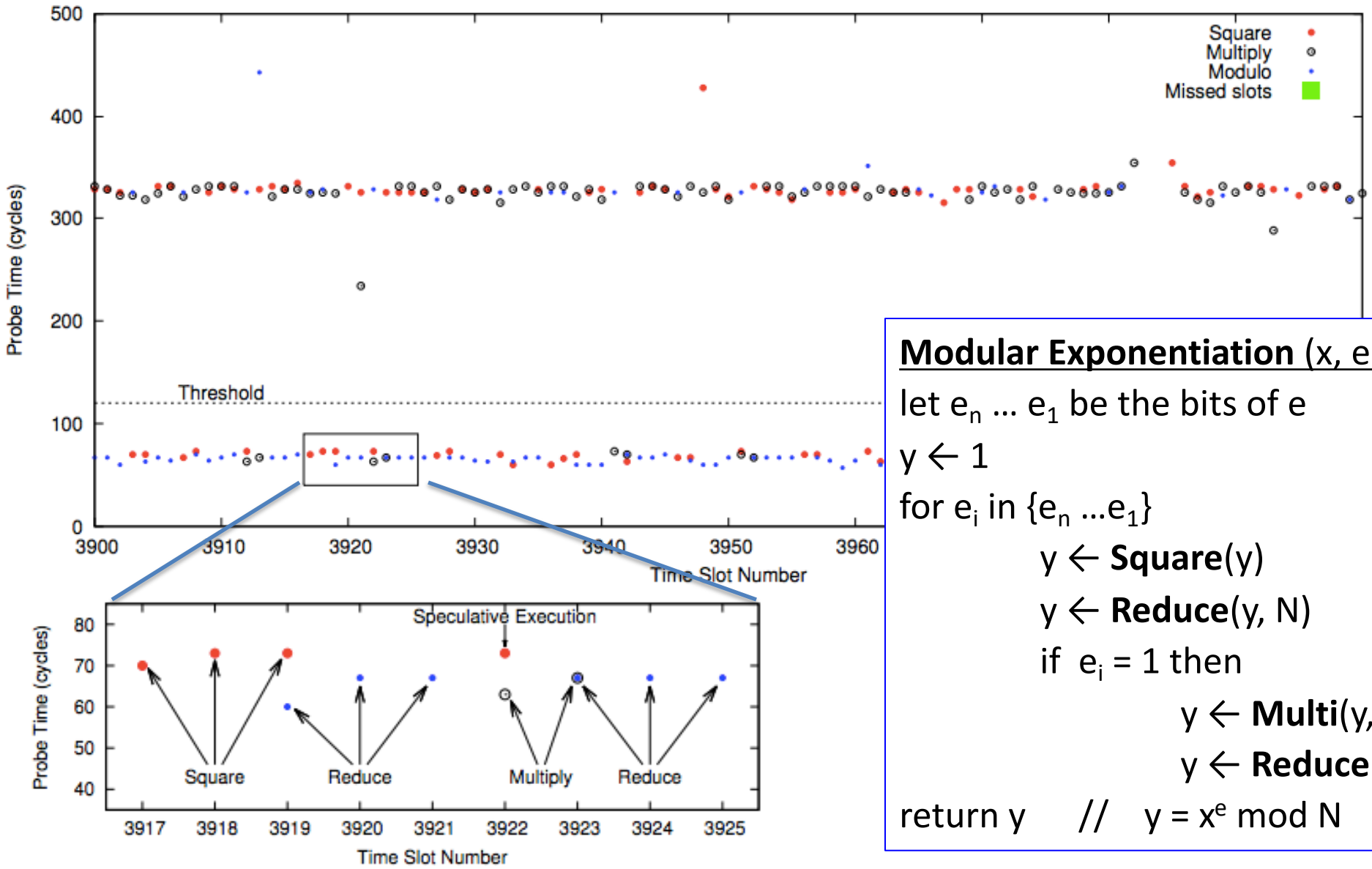
Flush+Reload protocol





- Fast time on S instr means S op
- Fast time on M instr means M op

Attacking Square and Multiply



Modular Exponentiation (x, e, N):

let $e_n \ldots e_1$ be the bits of e

$y \leftarrow 1$

for e_i in $\{e_n \ldots e_1\}$

- $y \leftarrow$ **Square**(y) **(S)**
- $y \leftarrow$ **Reduce**(y, N) **(R)**
- if $e_i = 1$ then
  - $y \leftarrow$ **Multi**(y, x) **(M)**
  - $y \leftarrow$ **Reduce**(y, N) **(R)**

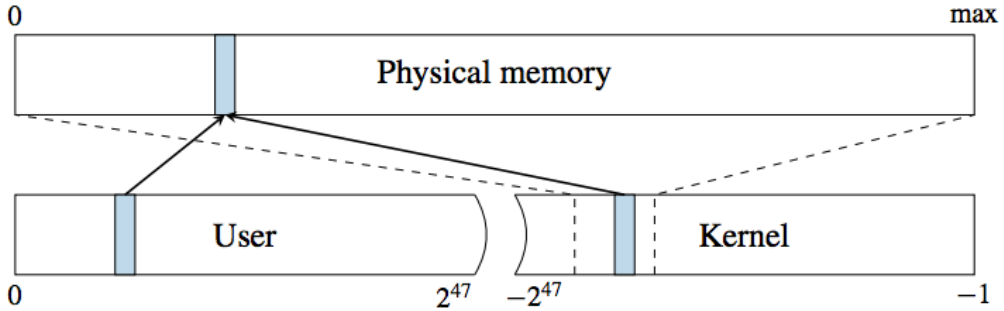
return y // $y = x^e \bmod N$

Flush+Reload widely applicable side-channel

- Useful anywhere code is shared across processes / VMs / containers
- Cross-tenant attacks in platform-as-a-service (PaaS) clouds
  - [Zhang, Juels, Reiter, R. 2014]
- Used as building block for Meltdown, Spectre vulnerabilities

Meltdown

- Speculative execution bugs in Intel x86, ARM, IBM processors + cache-based side-channels (I
  - Allows reading kernel (or hypervisor, other VM) memory



Intel didn't warn US government about CPU security flaws until they were public

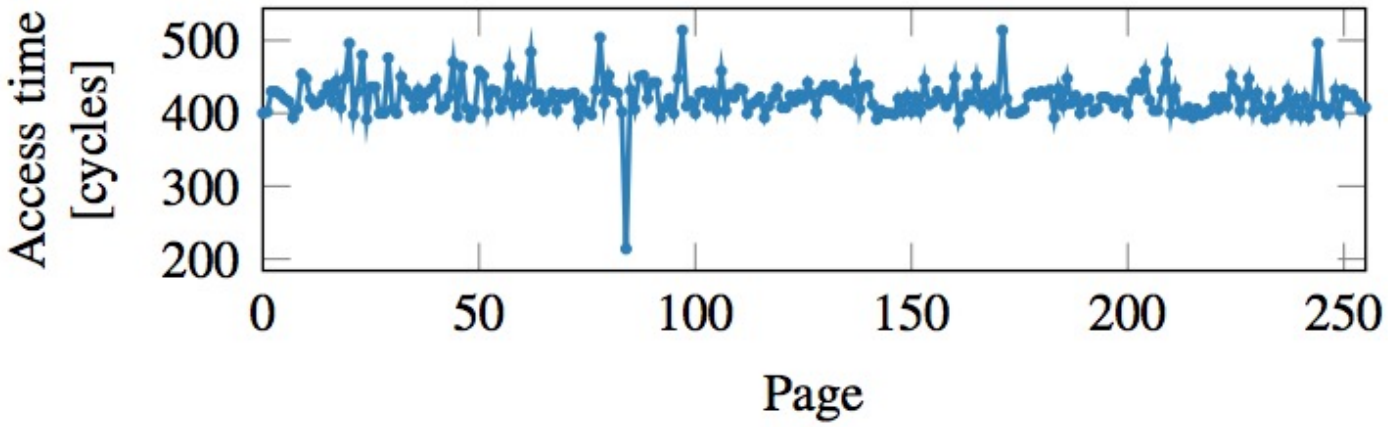
Meltdown and Spectre were kept secret

Researchers find malware samples that exploit Meltdown and Spectre

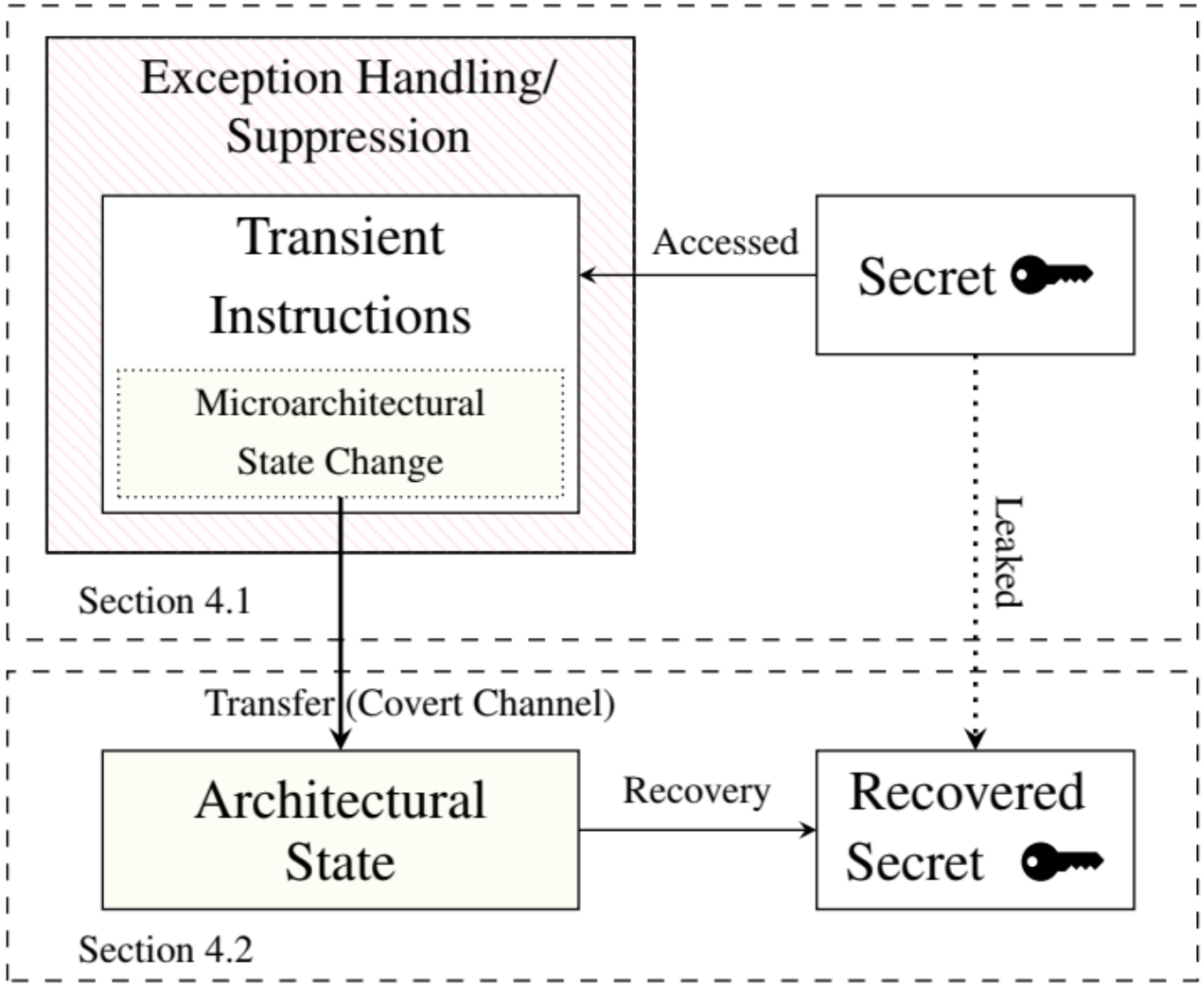
As of Feb. 1, antivirus testing firm AV-TEST had found 139 malware samples that exploit Meltdown and Spectre. Most are not very functional, but that could change.

Meltdown: intuition

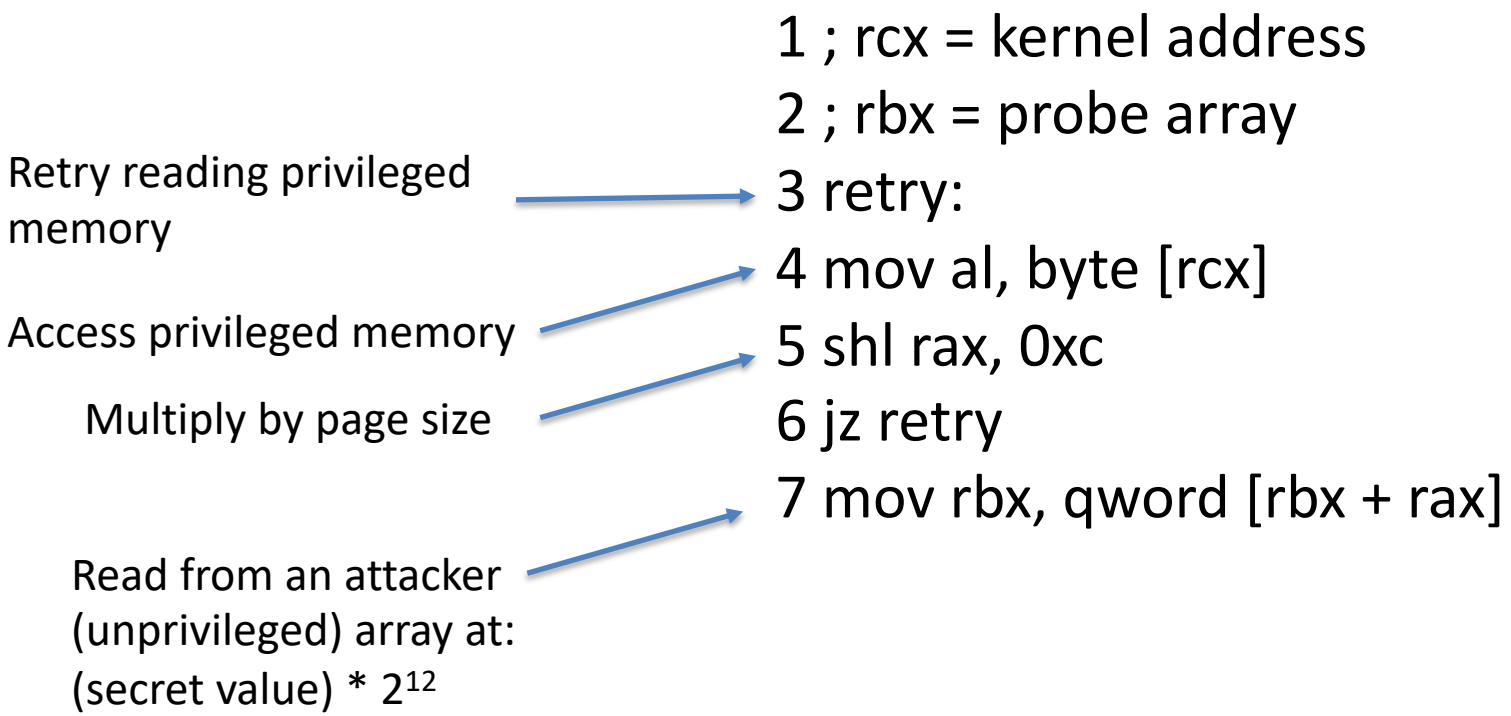
```
1  raise_exception();
2  // the line below is never reached
3  access(probe_array[data * 4096]);
```



Meltdown: design



Meltdown: core spy code



Attacker times accessing [rbx + rax] for different values of rax

When finds one that loads fast, learns sensitive byte

Lessons

- Isolation/containment useful for defense-in-depth
- Don’t rely *solely* on VMs for:
  - VMM transparency
  - Containment
  - Strong isolation (side channels exist)
  - Securing guest OS and host OS needed for defense-in-depth
- Side-channels are perennial problem
- Microarchitectural vulnerabilities like Meltdown new frontier
  - Spectre, ForeShadow, Fallout, Zombieload, …