2016 中国互联网安全大会
China Internet Security Conference

协同联动 共建安全+命运共同体

虚拟化平台的安全研究现状与趋势

栾尚聪 阿里云资深安全工程师

栾尚聪（好风）：
阿里云 - 云平台安全团队

- 虚拟化逃逸漏洞：XSA-148、XSA-182 …
- HITB AMS 2016 “Xen平台逃逸攻击”
- Blackhat USA 2016 “XSA-182撕裂虚拟化平台”

目录

·虚拟化安全与威胁

·虚拟化逃逸攻击

·侧信道攻击

·Q&A

一 虚拟化安全与威胁

虚拟化技术

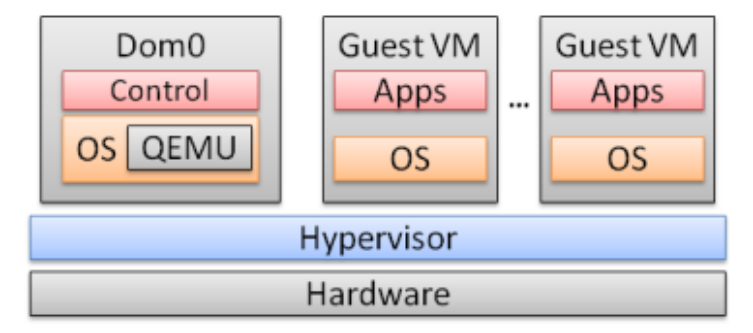
通过使用控制程序隐藏特定计算平台的实际物理特性，为用户提供抽象的、统一的、模拟的计算环境。

虚拟化技术

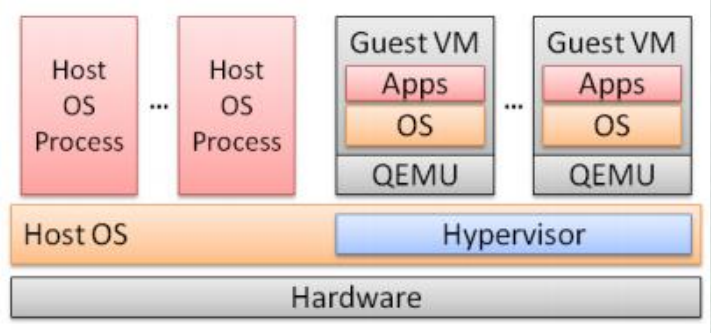
通过使用控制程序隐藏特定计算平台的实际物理特性，为用户提供抽象的、统一的、模拟的计算环境。

控制程序：VMM（虚拟机监视器）或 Hypervisor
计算环境：VM（虚拟机）或 Domain（域）

虚拟化架构



Xen Architecture

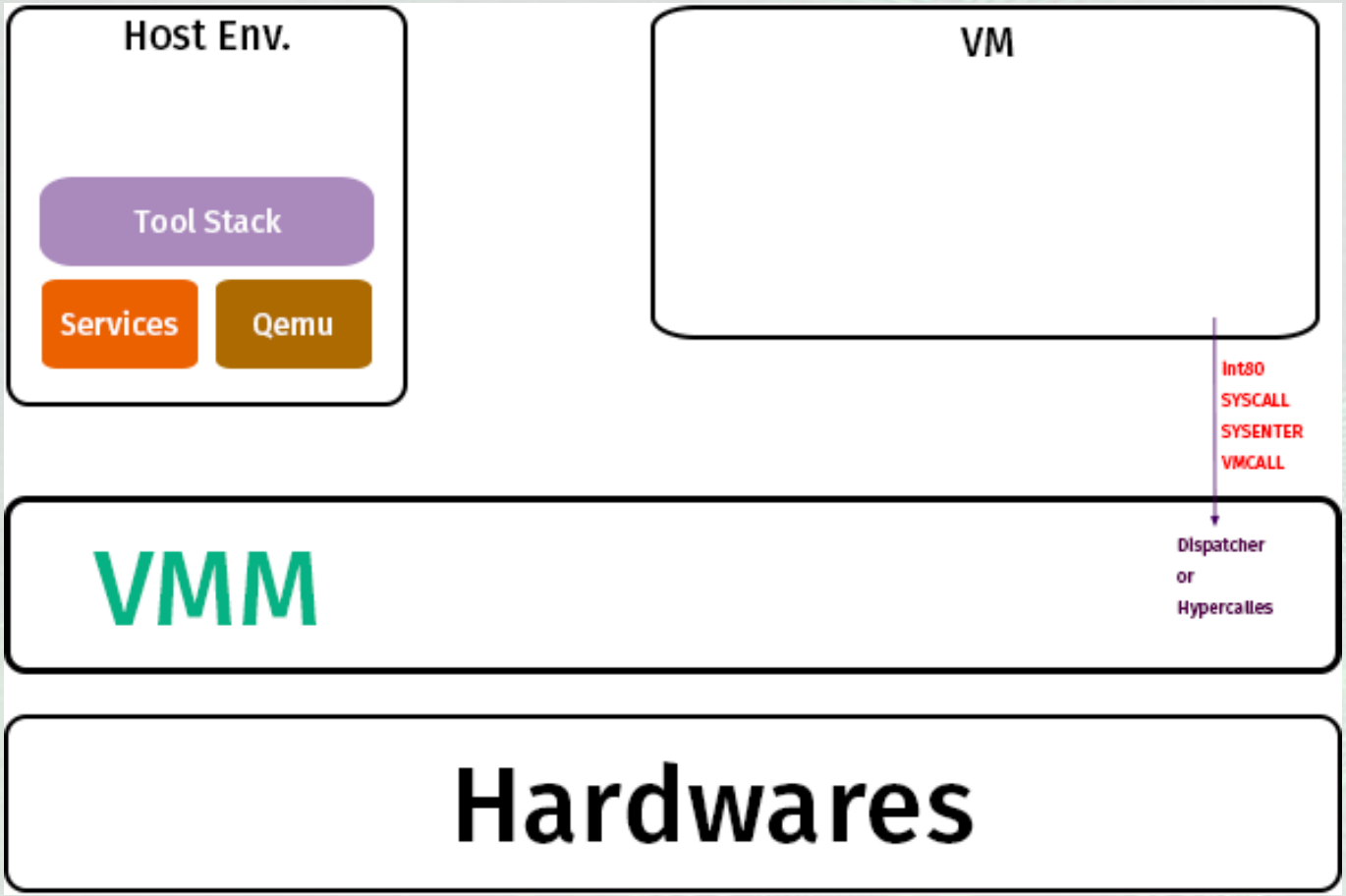


KVM Architecture

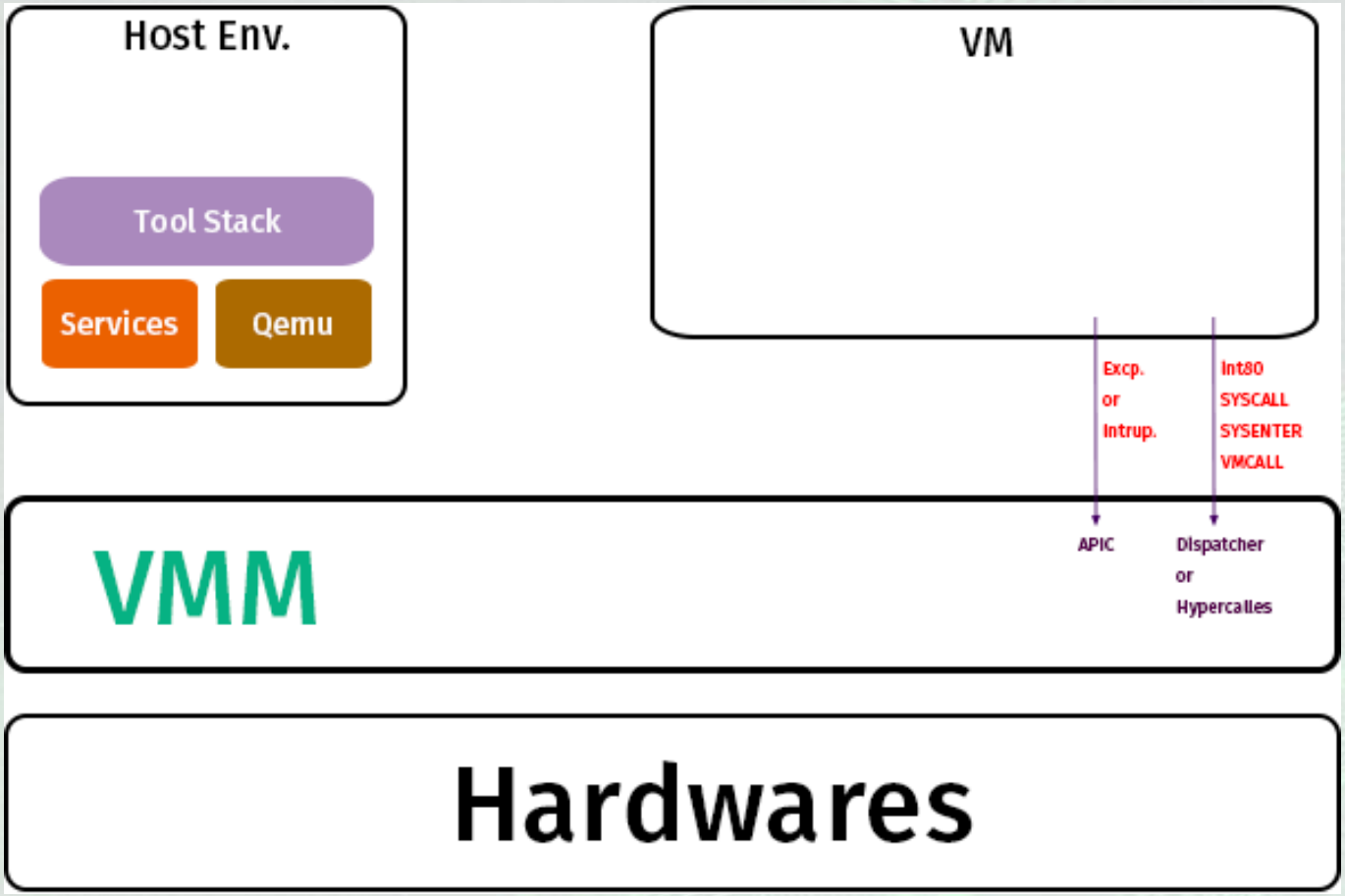
虚拟化基础设施

- Hypervisor
- Qemu
- Services
- Tool Stack

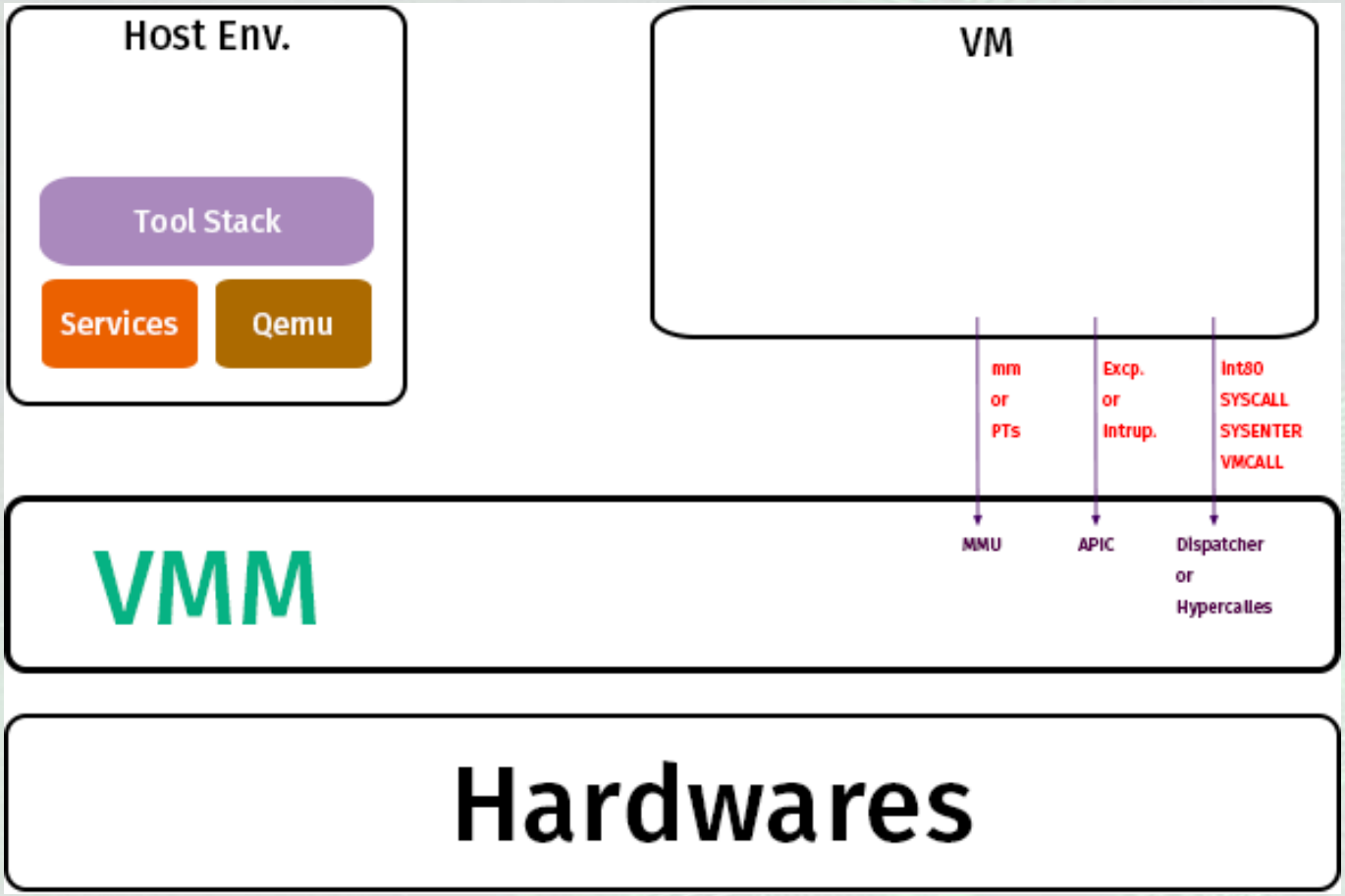
基本攻击界面


Host Env.
Tool Stack
Services
Qemu
VM
Int80
SYSCALL
SYSENTER
VMCALL
VMM
Dispatcher
or
Hypercalles
Hardwares

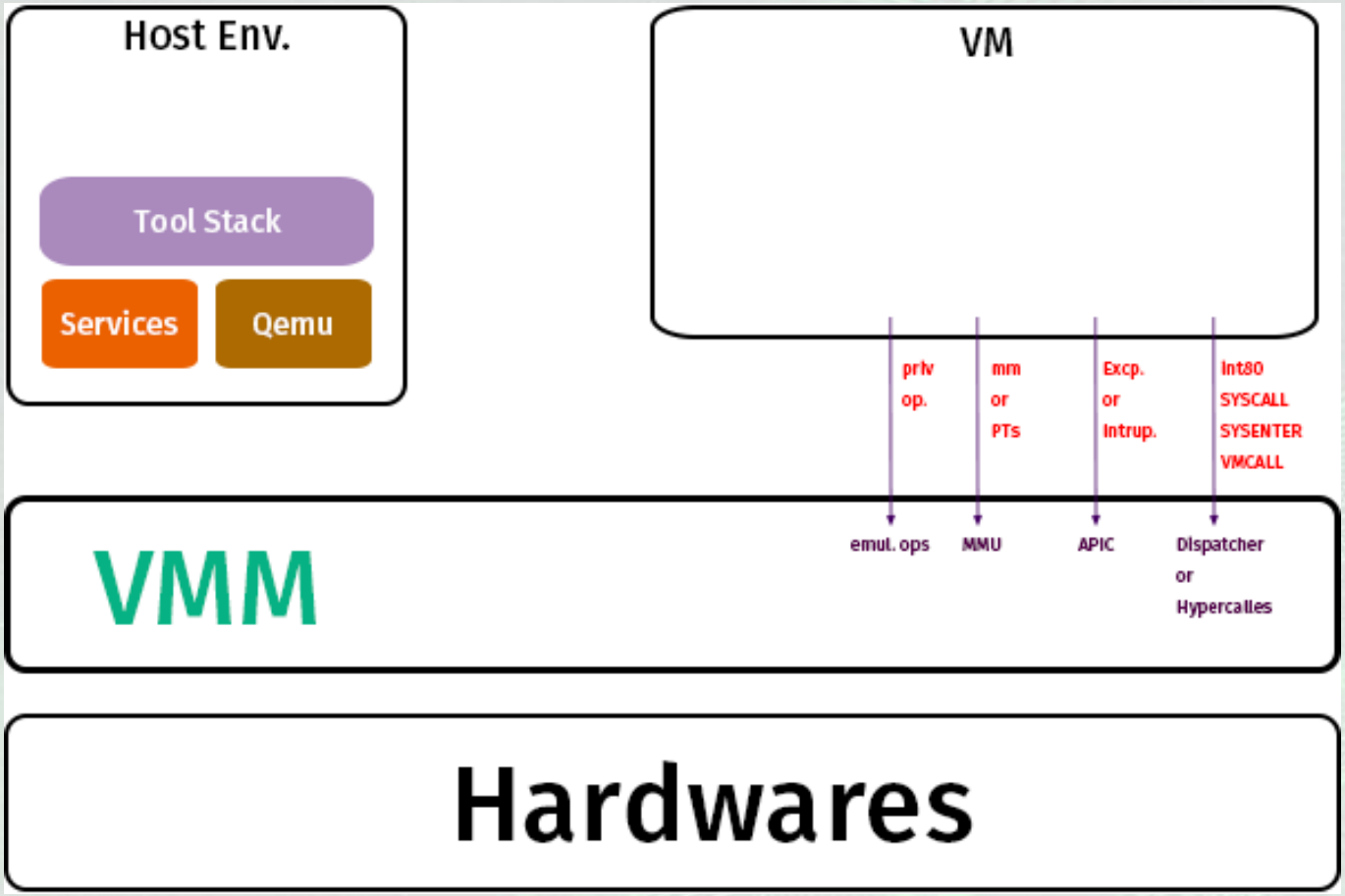
基本攻击界面


Host Env.
Tool Stack
Services
Qemu
VM
Excp.
or
Intrup.
Int80
SYSCALL
SYSENTER
VMCALL
APIC
Dispatcher
or
Hypercalles
VMM
Hardwares

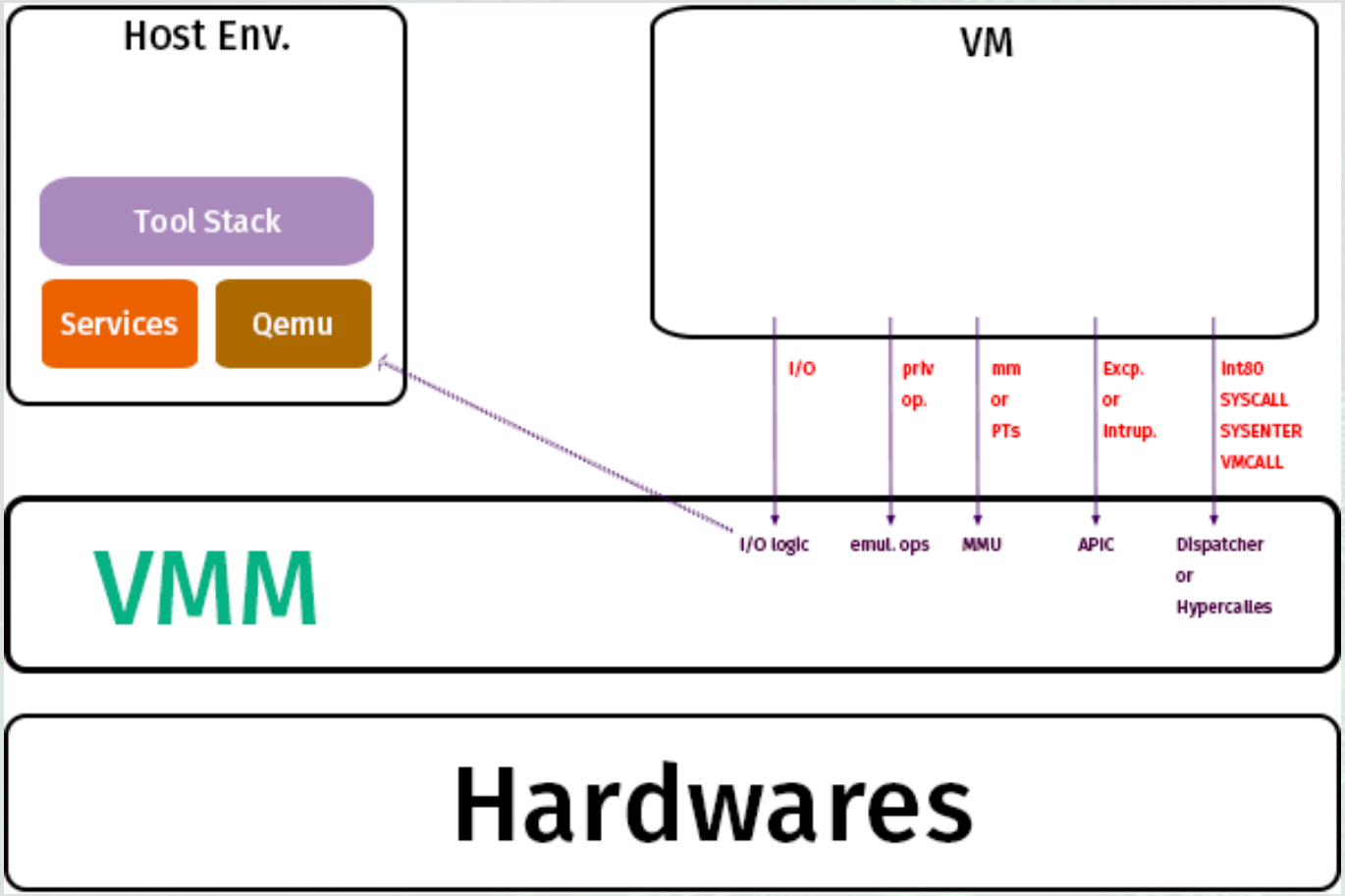
基本攻击界面


Host Env.
Tool Stack
Services
Qemu
VM
mm
or
PTs
Excp.
or
Intrup.
Int80
SYSCALL
SYSENTER
VMCALL
VMM
MMU
APIC
Dispatcher
or
Hypercalles
Hardwares

基本攻击界面


Host Env.
Tool Stack
Services
Qemu
VM
priv op.
mm or PTs
Excp. or Intrup.
Int80 SYSCALL SYSENTER VMCALL
emul. ops
MMU
APIC
Dispatcher or Hypercalles
VMM
Hardwares

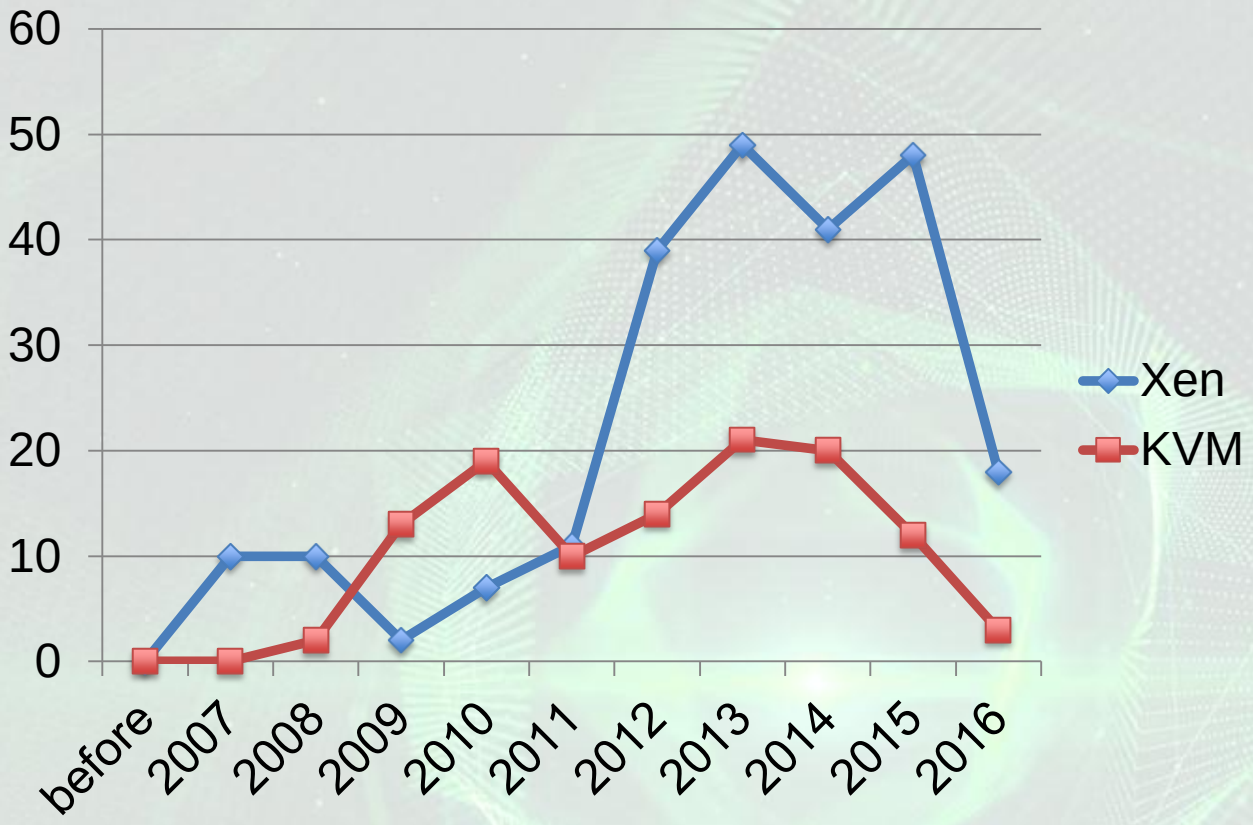
基本攻击界面


Host Env.
Tool Stack
Services
Qemu
VM
I/O
priv op.
mm or PTs
Excp. or Intrup.
Int80 SYSCALL SYSENTER VMCALL
I/O logic
emul. ops
MMU
APIC
Dispatcher or Hypercalles
VMM
Hardwares

威胁与分类

- 设计与实现问题
  - 拒绝服务（>90%）
  - 信息泄露
  - 客户机安全
  - 逃逸攻击
- 侧信道问题
  - 信息获取
  - 行为影响

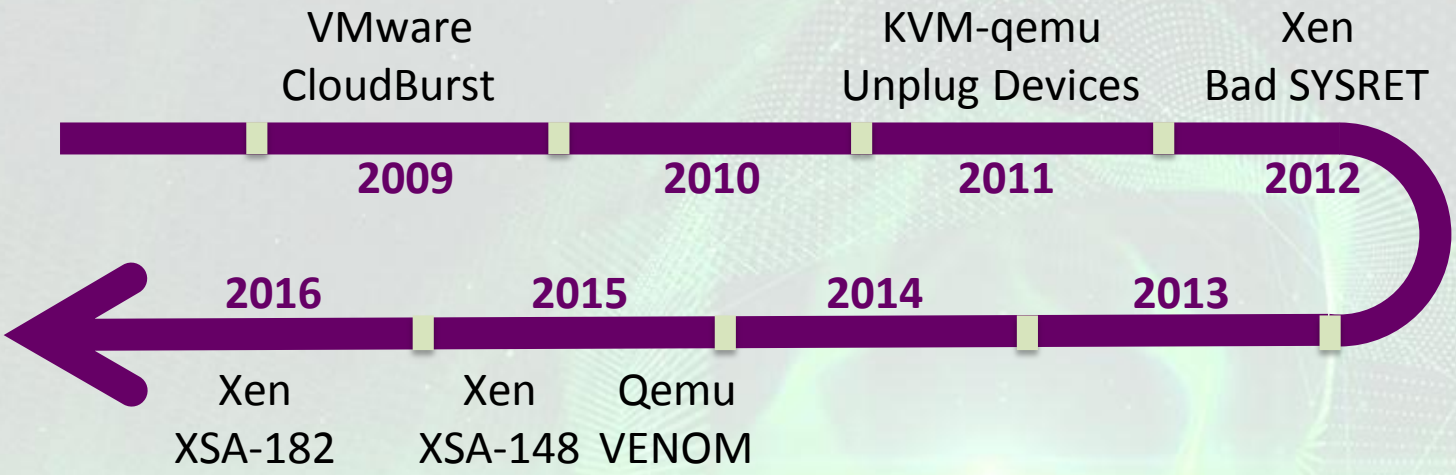
Xen/KVM漏洞统计



注：
1. 原始数据采集自CVE漏洞库和XSA公告
2. 相应的Qemu漏洞算入各自的漏洞数据中

二 虚拟化逃逸攻击

Escape Researches时间线



1）VMware CloudBurst - 2009

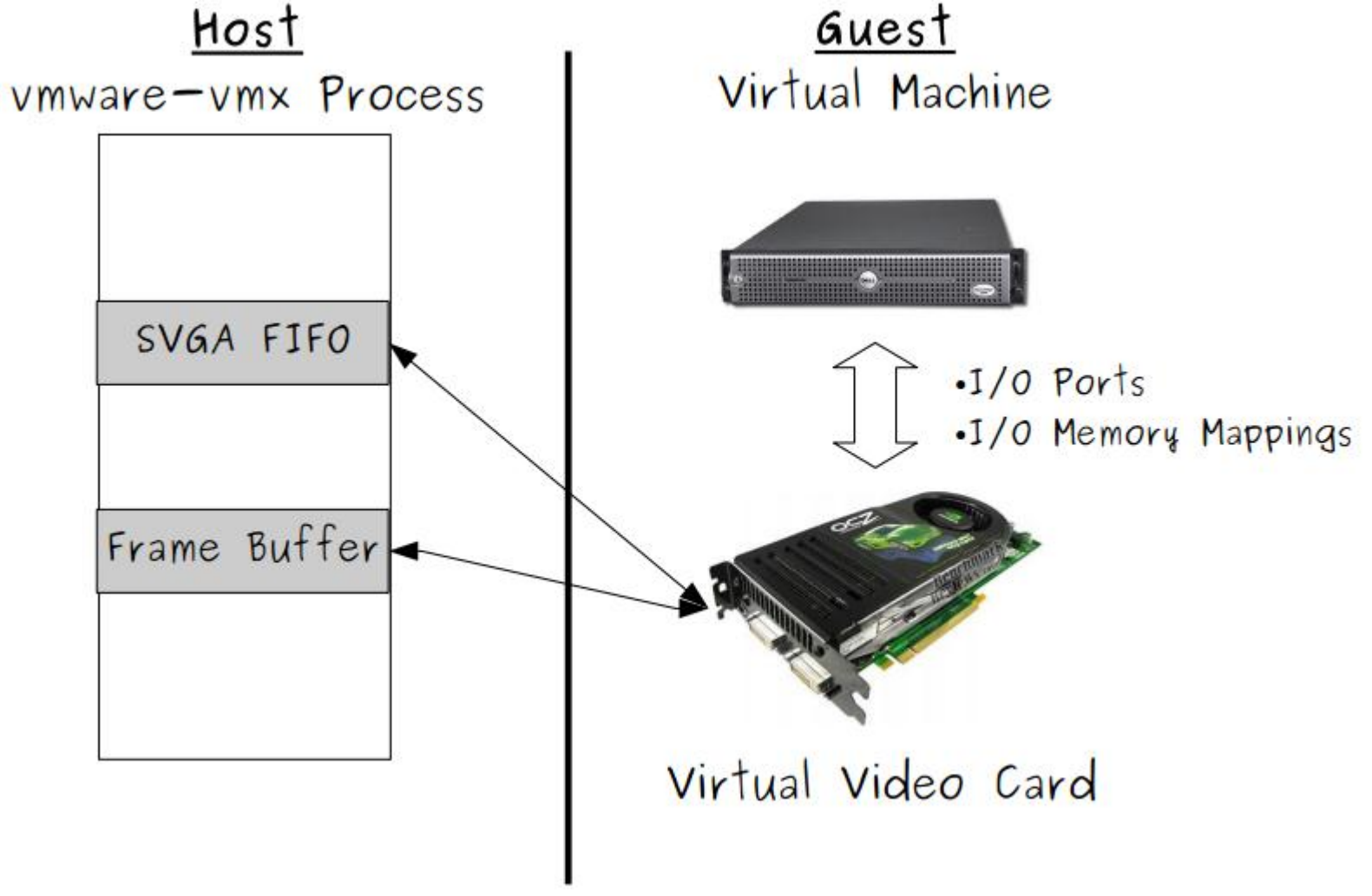
SVGA 2D

- SVGA_CMD_RECT_COPY
- SVGA_CMD_DRAW_GLYPH

SVGA 3D

- SETRENDERSTATE
- SETLIGHTENABLED
- SETRENDERTARGET
- SETCLIPPLANE
- SETTRANSFORM

VMware SVGA Guest <-> Host



2) Bad SYSRET - 2012

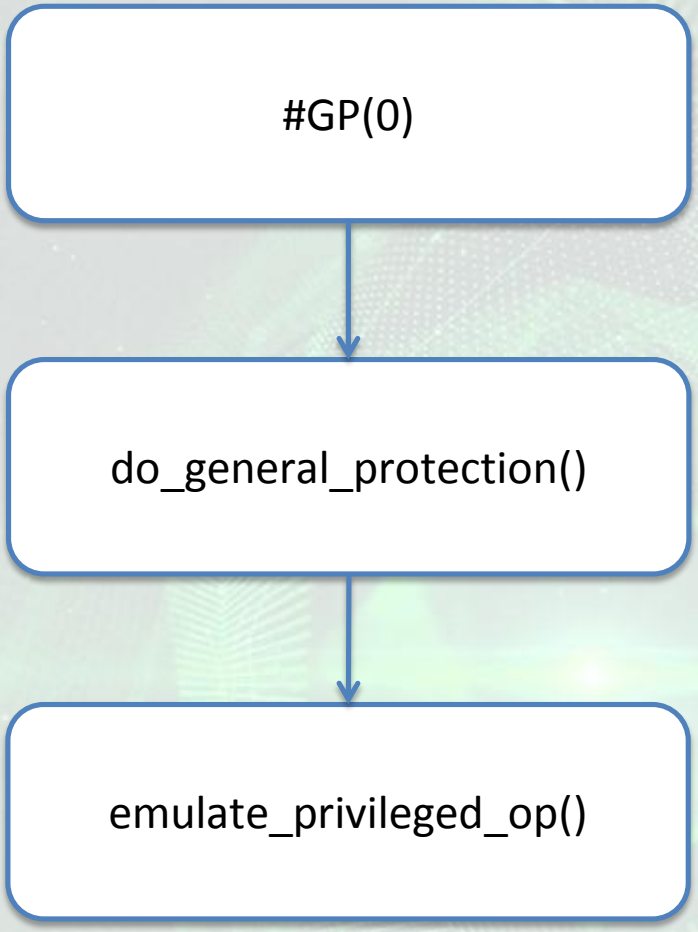
CVE-2012-0217

The x86-64 kernel system-call functionality in Xen 4.1.2 and earlier,
…Solaris…illumos…SmartOS …FreeBSD…
NetBSD…Microsoft Windows…and possibly other operating systems,
when running on an Intel processor,
incorrectly uses the sysret path in cases where a certain address is not a canonical address,
which allows local users to gain privileges via a crafted application.

Bad SYSRET - 2012

SYSRET—Return From Fast System Call
IF (CS.L ≠ 1) or (IA32_EFER.LMA ≠ 1) or (IA32_EFER.SCE ≠ 1)
(* Not in 64-Bit Mode or SYSCALL/SYSRET not enabled in IA32_EFER *)
THEN #UD; FI;
IF (CPL ≠ 0) OR (RCX is not canonical) THEN #GP(0); FI;
IF (operand size is 64-bit)
THEN (* Return to 64-Bit Mode *)
RIP ← RCX;
ELSE (* Return to Compatibility Mode *)
RIP ← ECX;
FI;
RFLAGS ← (R11 & 3C7FD7H) | 2; (* Clear RF, VM, reserved bits; set bit 2 *)
... (* deal with CS)
... (* desl with SS)

VUPEN Team Exploitation Tech.



3.3）QEMU 漏洞

- KVM-Qemu Unplug Device, 2012
- Qemu VENOM, 2015
- …

3.4) Xen XSA-148 - 2015

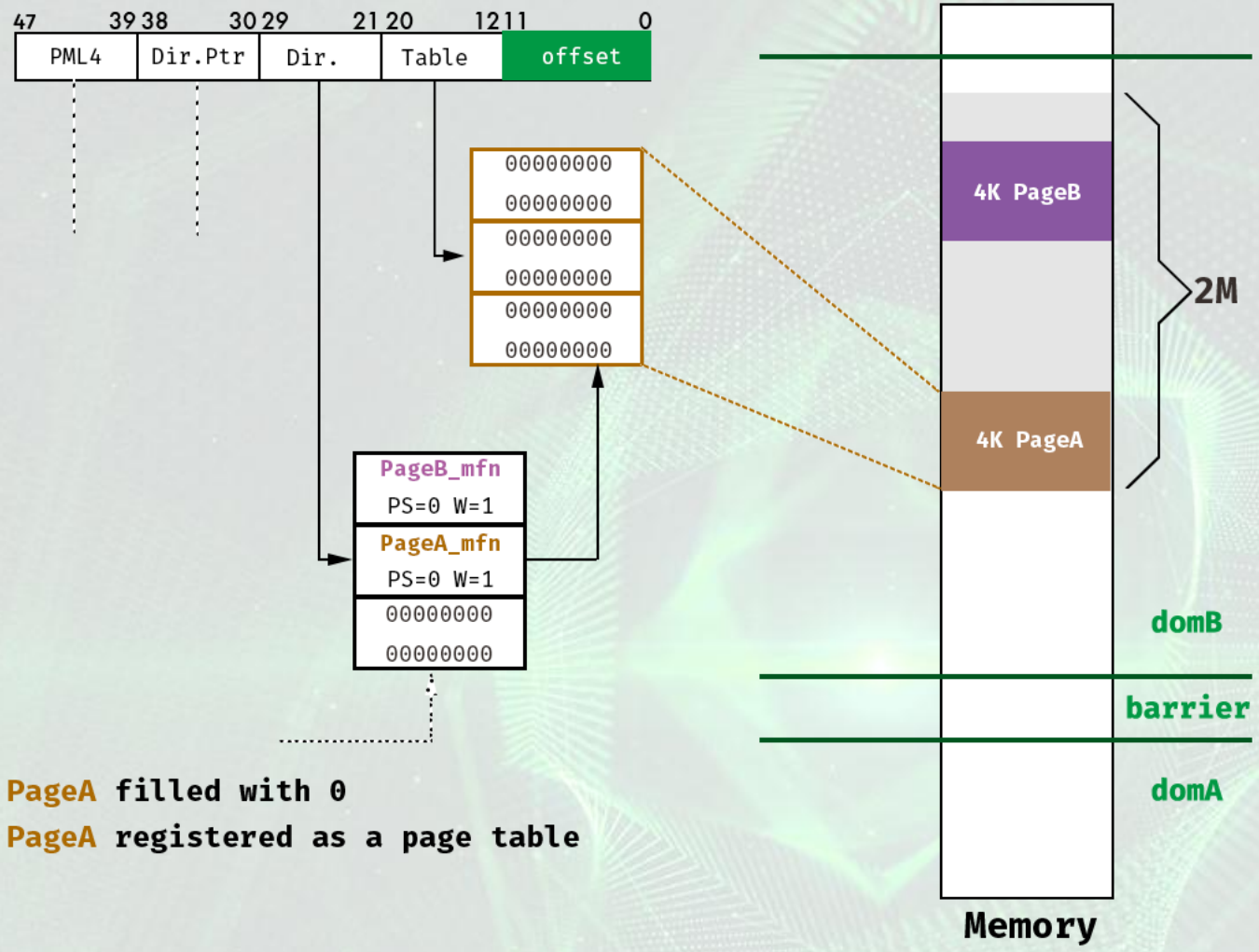
Information

Advisory	XSA-148
Public release	2015-10-29 11:59
Updated	2015-10-29 11:59
Version	4
CVE(s)	CVE-2015-7835
Title	x86: Uncontrolled creation of large page mappings by PV guests

Files

advisory-148.txt (signed advisory file)
xsa148.patch
xsa148-4.4.patch
xsa148-4.5.patch

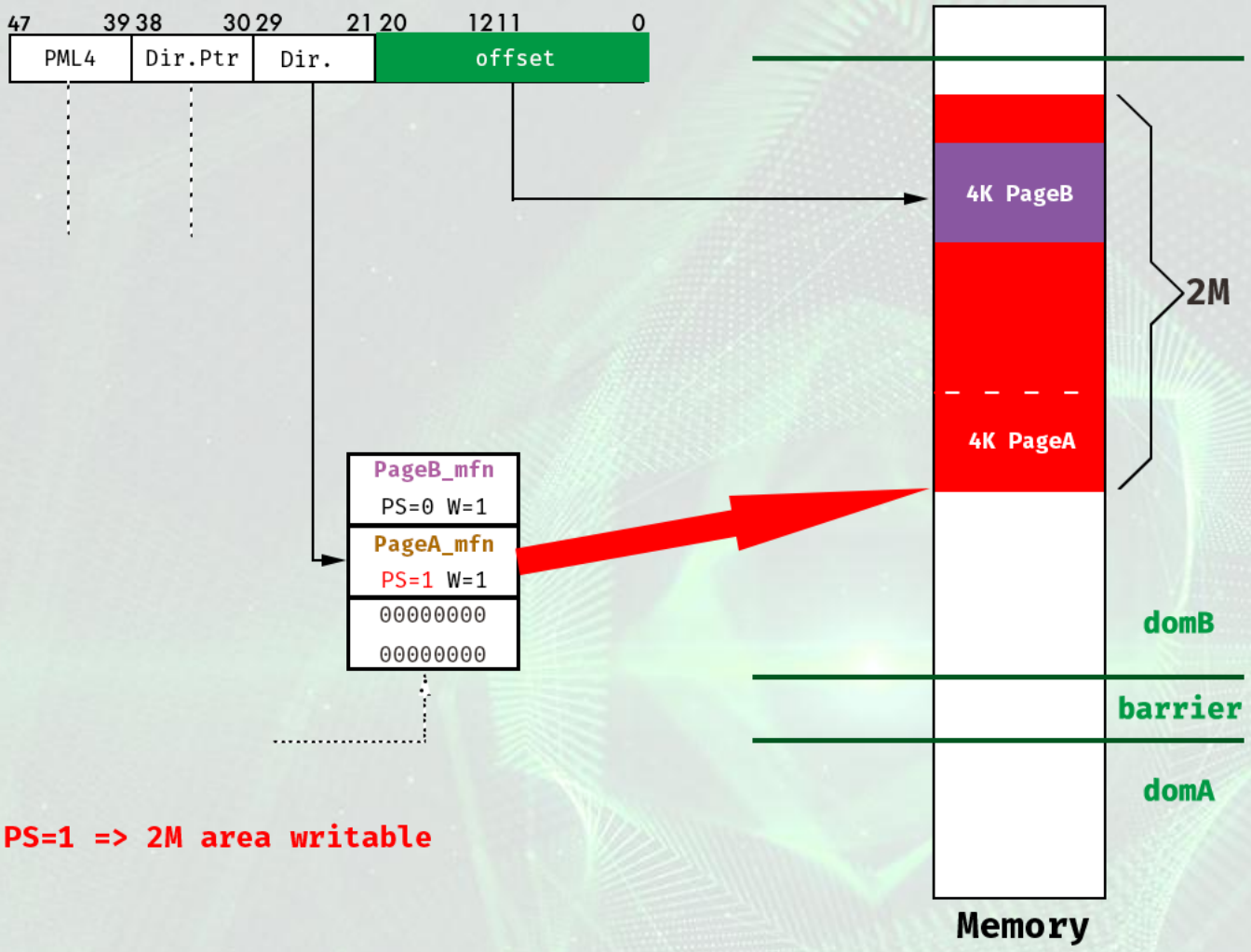
XSA-148：构造特殊页表



PageA filled with 0

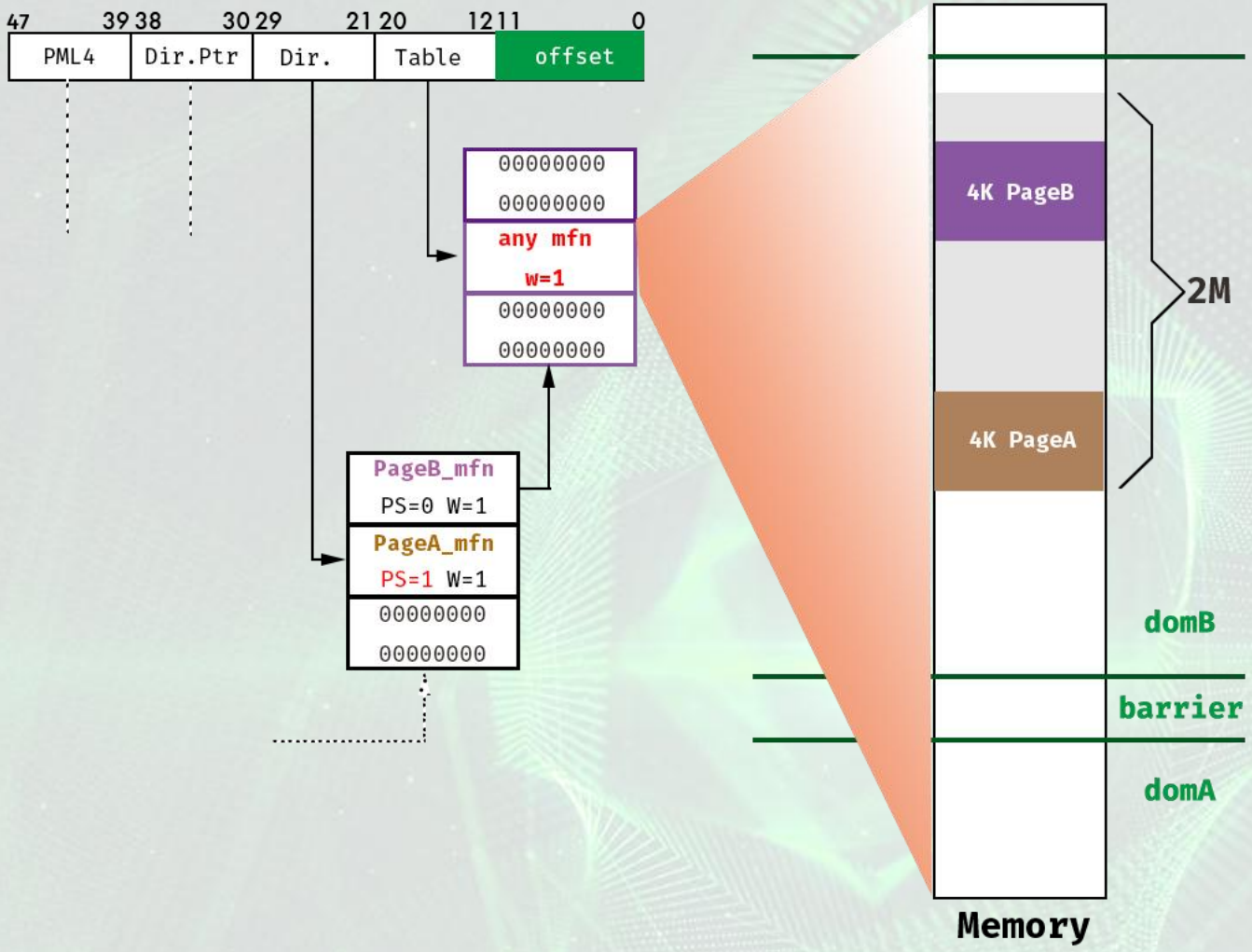
PageA registered as a page table

XSA-148：漏洞触发



PS=1 => 2M area writable

XSA-148：任意机器内存读写



XSA-148：代码执行

执行流劫持：

- VDSO/vsyscall Page (Linux)
- SharedUserData Page (Windows)
- Hypercall Page
- …

3.5) Xen XSA-182 - 2015

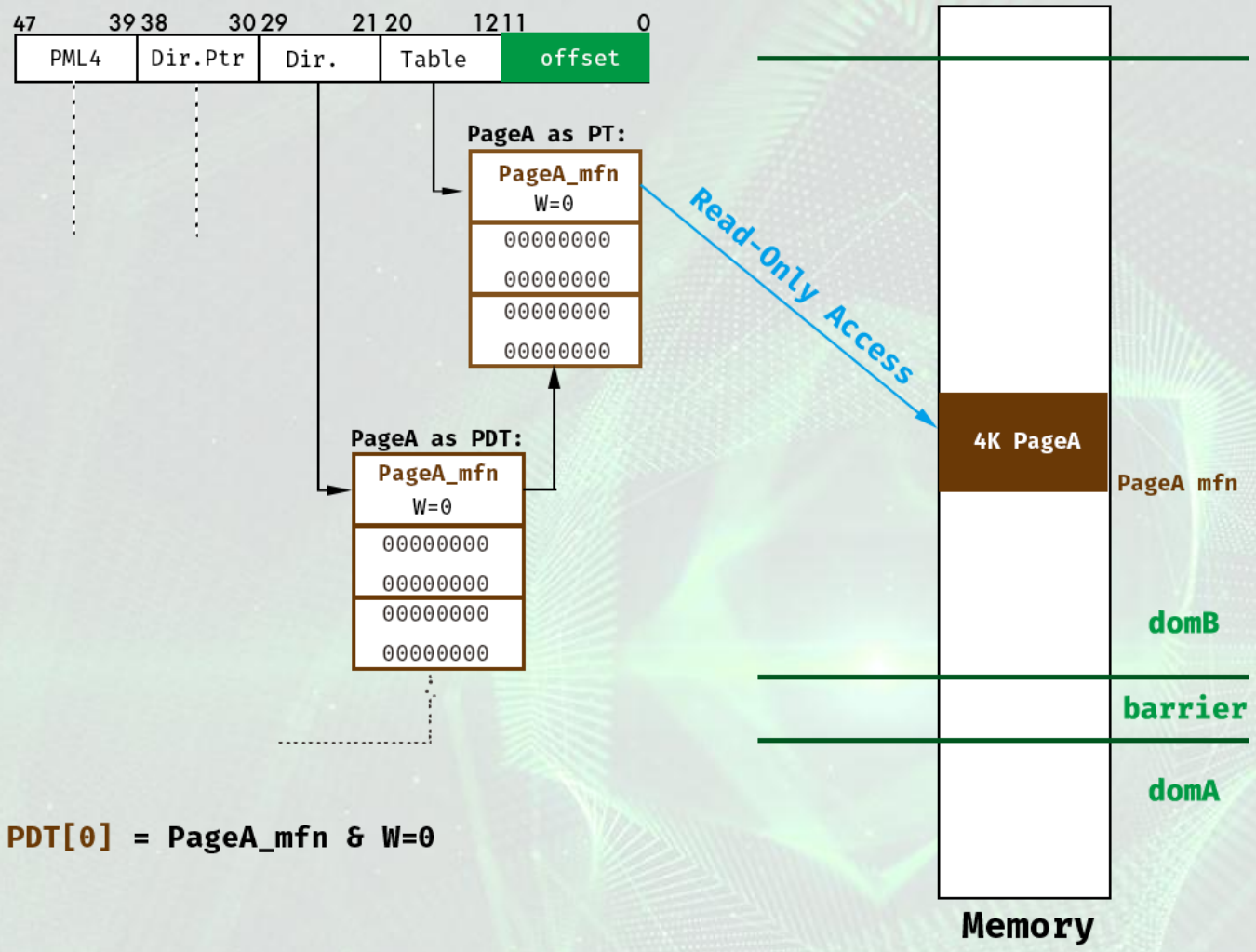
Information

Advisory	XSA-182
Public release	2016-07-26 11:32
Updated	2016-07-26 11:32
Version	3
CVE(s)	CVE-2016-6258
Title	x86: Privilege escalation in PV guests

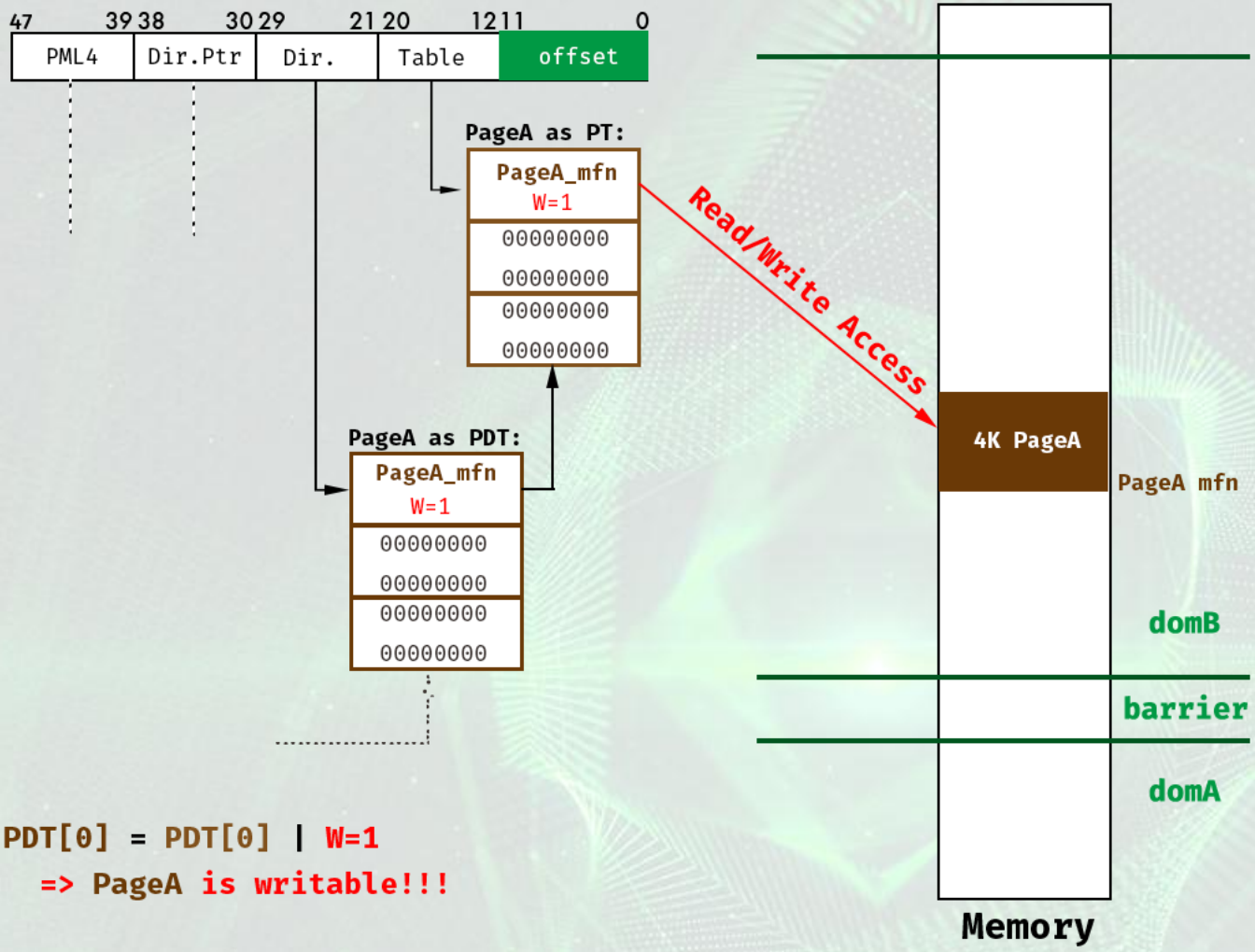
Advisory

```
            Xen Security Advisory CVE-2016-6258 / XSA-182
                                version 3

                x86: Privilege escalation in PV guests

ISSUE DESCRIPTION
=================

The PV pagetable code has fast-paths for making updates to pre-existing
pagetable entries, to skip expensive re-validation in safe cases
(e.g. clearing only Access/Dirty bits).  The bits considered safe were too
broad, and not actually safe.

IMPACT
======

A malicous PV guest administrator can escalate their privilege to that
of the host.

VULNERABLE SYSTEMS
==================

All versions of Xen are vulnerable.

The vulnerability is only exposed to PV guests on x86 hardware.

The vulnerability is not exposed to x86 HVM guests, or ARM guests.
```

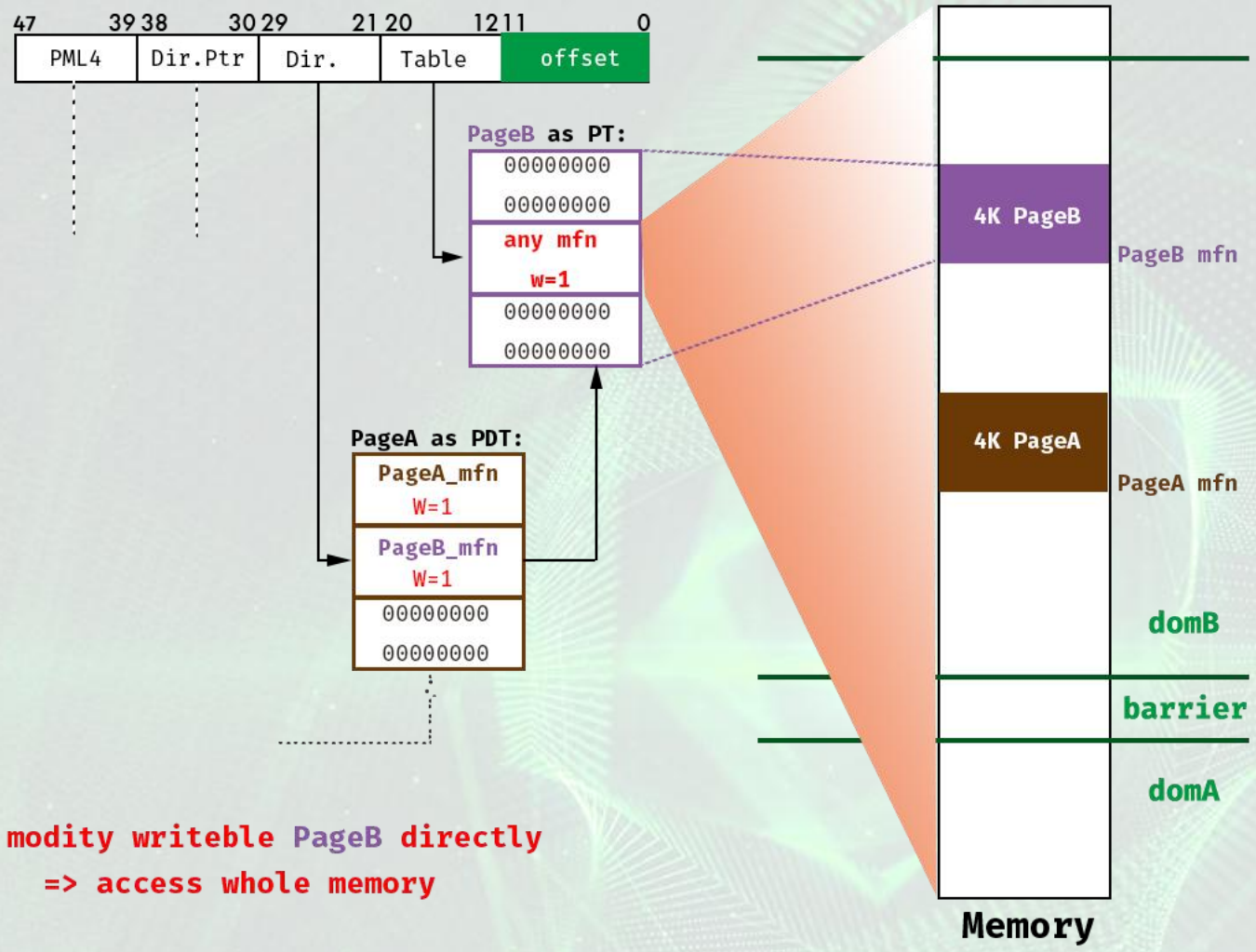
XSA-182：2级页表自引用



XSA-182：漏洞触发


47
39 38
30 29
21 20
12 11
0
PML4
Dir.Ptr
Dir.
Table
offset
PageA as PT:
PageA_mfn
W=1
00000000
00000000
00000000
00000000
PageA as PDT:
PageA_mfn
W=1
00000000
00000000
00000000
00000000
Read/Write Access
4K PageA
PageA mfn
domB
barrier
domA
Memory
PDT[0] = PDT[0] | W=1
=> PageA is writable!!!

XSA-182：任意机器内存读写



modity writeble PageB directly

=> access whole memory

XSA-148/XSA-182 White Paper

Exploit Two Xen Hypervisor Vulnerabilities

Shangcong Luan

Cloud Platform Security Team of Alibaba Cloud

shangcong.lsc@alibaba-inc.com

Abstract

The Xen Project is a widely used virtualization platform powering some of the largest clouds in production today. In the process of virtualization security research on it, our team has discovered two critical vulnerabilities in the PV mode Memory Management of Xen Hypervisor. This paper aims to present a comprehensive study of Xen Hypervisor PV Guest Memory Management and detail our two critical vulnerabilities. Furthermore, full exploitation technologies will be discussed.

Keywords: Xen Security, XSA-148, Dome Breaking, XSA-182, Ouroboros, hypervisor exploitation, VM Escape

1. Introduction

Xen is an open source project providing virtualization services that allow multiple computer operating systems to execute on the same computer hardware concurrently. It originated as a research project at the University of Cambridge and the first public release was made in 2003. Since then the project has attracted extensive attention from virtualization and security researchers. In the past decades, virtual machine escape was considered as an unreal story because of the complex and effective isolation supported by virtualization technologies although some security vulnerabilities had been found. But unfortunately, a SVGA emulation bug was reported on VMware in 2008 and at the next year's Blackhat conference, researchers from Immunity Team disclosured a fully working exploit which proved virtual machine escape isnt a joke. In 2012, the unbelievable SYSRET vulnerability was disclosured and Xen was affected at this time. In 2015, the infamous VENOM vulnerability in QEMU evoked worldwide repercussions although no one could exploit it in the real scene.

三 侧信道威胁

侧信道攻击

侧信道攻击：

攻击者通过非直接请求或控制被攻击目标的方式获取敏感信息或影响目标行为，该攻击方式建立在攻击者和被攻击目标的共享资源之上。

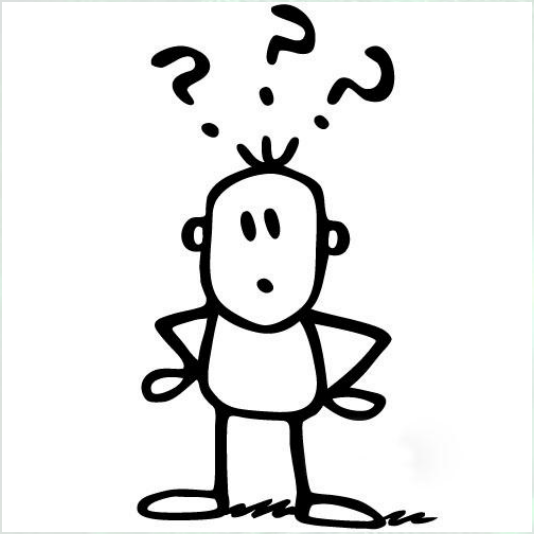
By hikerell

侧信道攻击

侧信道攻击：

攻击者通过非直接请求或控制被攻击目标的方式获取敏感信息或影响目标行为，该攻击方式建立在攻击者和被攻击目标的共享资源之上。

By hikerell



侧信道攻击

侧信道攻击：

攻击者通过非直接请求或控制被攻击目标的方式获取敏感信息或影响目标行为，该攻击方式建立在攻击者和被攻击目标的共享媒介之上。

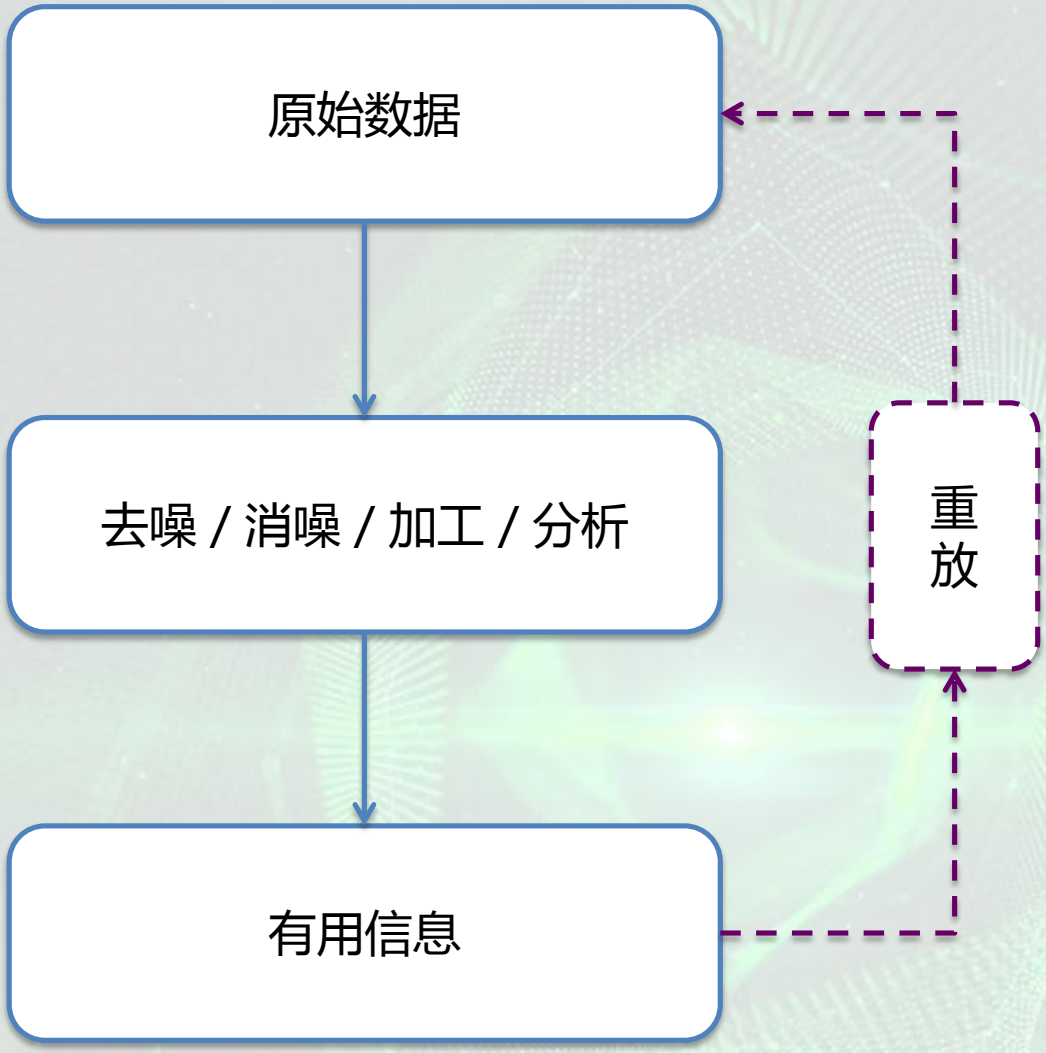
By hikerell

- 非直接请求或控制被攻击目标
- 获取敏感信息
- 影响目标行为
- 共享媒介

共享媒介——侧信道（Side Channel）

- 处理器（CPU/GPU）
- 内存
- Cache/TLB
- 系统总线
- 外部存储器
- …

获取敏感信息

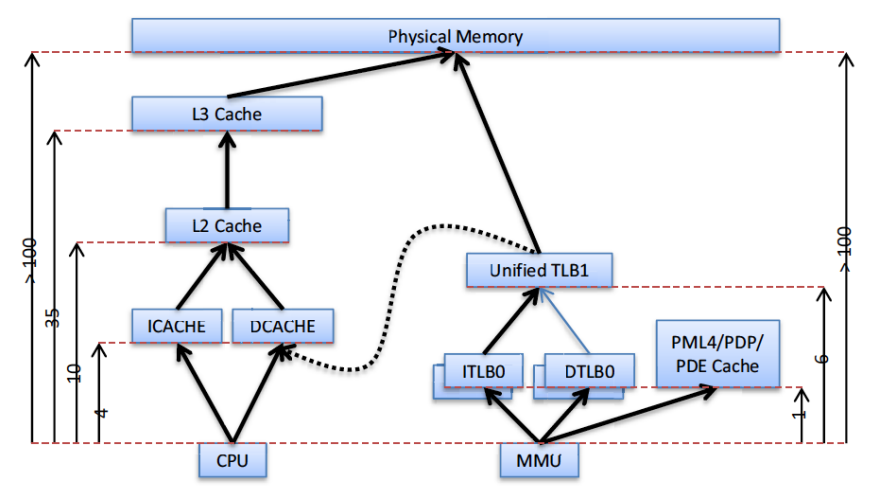


获取敏感信息

Practical Timing Side Channel Attacks Against Kernel Space ASLR

Ralf Hund, Carsten Willems, Thorsten Holz

2013 IEEE Symposium on Security and Privacy



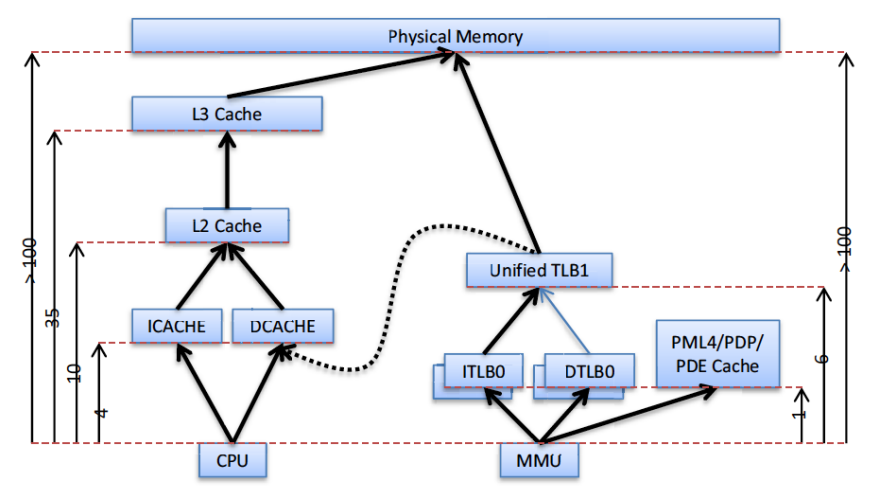
1. Prime + Probe
2. Double Page Fault
3. Address Translation Cache Preloading

获取敏感信息

Practical Timing Side Channel Attacks Against Kernel Space ASLR

Ralf Hund, Carsten Willems, Thorsten Holz

2013 IEEE Symposium on Security and Privacy



❶ Prime + Probe

❷ Double Page Fault

❸ Address Translation Cache Preloading

获取（部分）内核内存布局信息

ISC 2016 中国互联网安全大会
360互联网安全中心

从缓存中提取信息

侧信道：Cache
攻击思路：

- Evict + Time
- Prime + Probe
- Flush + Reload
- Flush + Flush

成熟的攻击套路

从缓存中提取信息

侧信道：Cache
攻击思路：

- Evict + Time
- Prime + Probe
- Flush + Reload
- Flush + Flush


ISC
2016
中国互联网安全大会
360互联网安全中心
谢 谢

附：侧信道影响目标行为？

Exploiting Processor Side Channels
To Enable Cross VM Malicious Code Executation

Sophia M. D'Antoine, RPI

April 2015, Thesis for the Degree of MASTER OF SCIENCE