qemu kvm study

https://lwn.net/Articles/810033/

1. EPT guest paging structures → Host physical address guest-physical address: Only 47:0 will be used for 64 bit address; high bit should be zero,

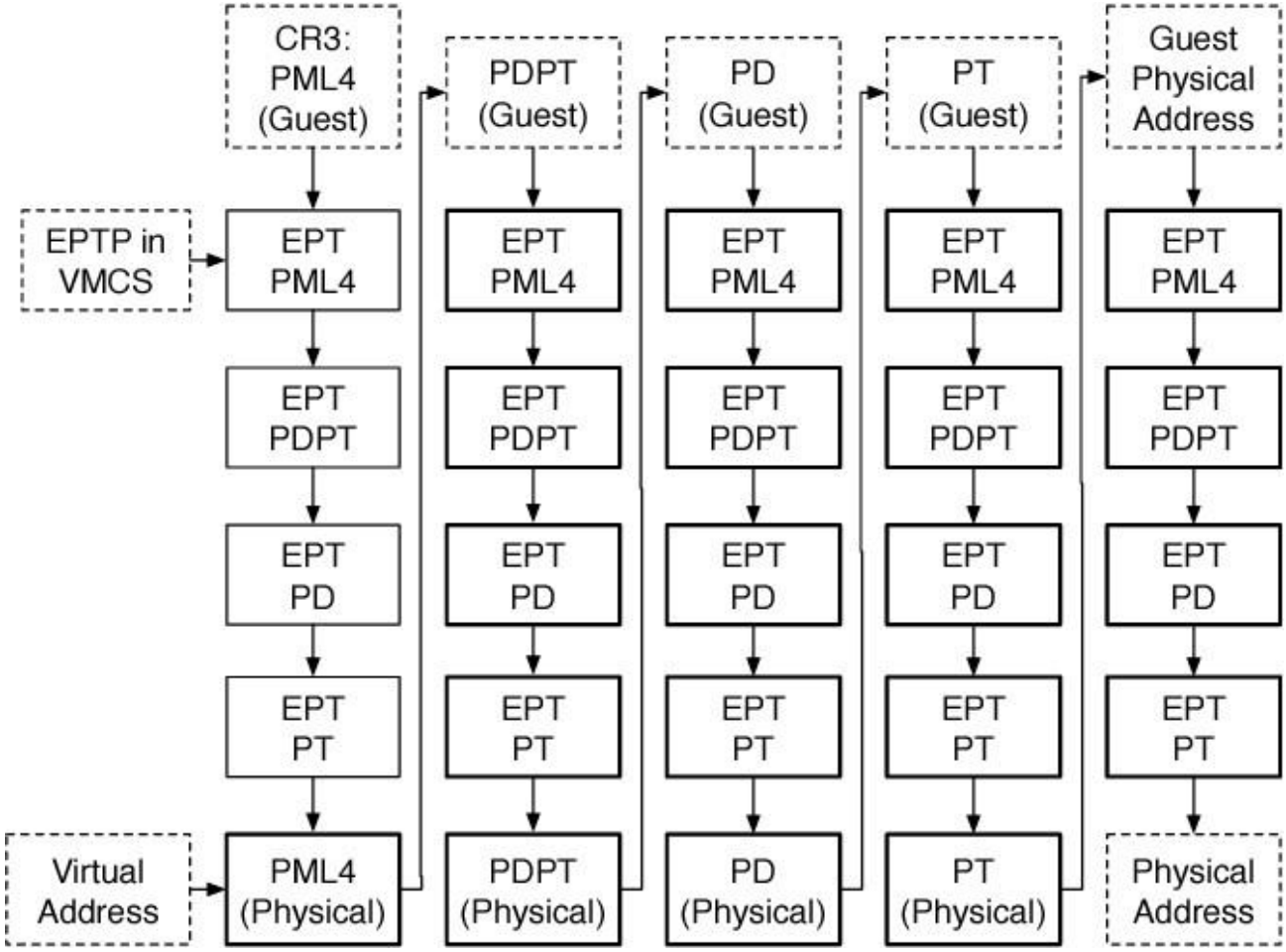
   4 级页表:

47:0	47：39	[38:30]	[29:21]	[20:12]	low 12bit
最大 256T	512G/entry	1G/entry	2M/entry	4k/entry	4K
EPTP	EPT PML4 Entry[512]	EPT Page-DirectoryPointer[512]	EPT page table[512]	EPT pte[512]	one host page

63	62	61	60	59	58	57	56	55	54	53	52	51	M[1]	M-1	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	
Reserved														Address of EPT PML4 table																								Rsvd.				SSS[2]	A/D	EPT PWL-1		EPT PS MT			EPTP
Ignored													Rsvd.	Address of EPT page-directory-pointer table																								Ign.	XU[4]	Ign.	A	Reserved				X	W	R	PML4 prese
SVE[7]	Ignored																																													0	0	0	PML4 not prese
SVE	Ign.	SSS[8]	Ignored											Rsvd.	Physical address of 1GB page																								Reserved							Ign.	XU	D	A	1	PAT	EPT MT		X	W	R	PDPT 1GE pag
Ignored													Rsvd.	Address of EPT page directory																								Ign.	XU	Ign.	A	0	Rsvd.			X	W	R	PDPT pag direct
SVE	Ignored																																													0	0	0	PDTP not prese
SVE	Ign.	SSS	Ignored											Rsvd.	Physical address of 2MB page																								Reserved							Ign.	XU	D	A	1	PAT	EPT MT		X	W	R	PDE 2ME pag
Ignored													Rsvd.	Address of EPT page table																								Ign.	XU	Ign.	A	0	Rsvd.			X	W	R	PDE pag tabl
SVE	Ignored																																													0	0	0	PDE not prese
SVE	Ign.	SPPg	SSS	Ignored										Rsvd.	Physical address of 4KB page																								Ign.	XU	D	A	Ign	PAT	EPT MT		X	W	R	PTE 4KE pag
SVE	Ignored																																													0	0	0	PTE not prese

页表 not present 的定义：X W R 均为 0， 即所指向页表即不可读，也不可写，也不可执行；

One long way：a Guest Virtual Address → Host Physical Address we need one guest physical address, and CR3, and EPTP in VMCS and all configured:



Two method to accerate the translation：

1. Huge page：1G、2M for host， 1G、2M for guest 2.

TLB（bigger） and Paging-Structure Caches

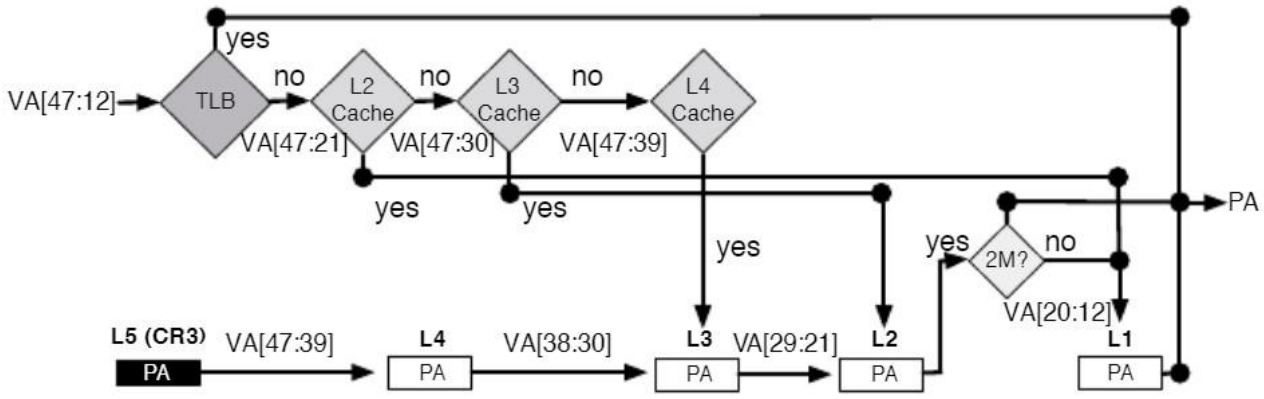


Figure 1. The x86 native page walk.

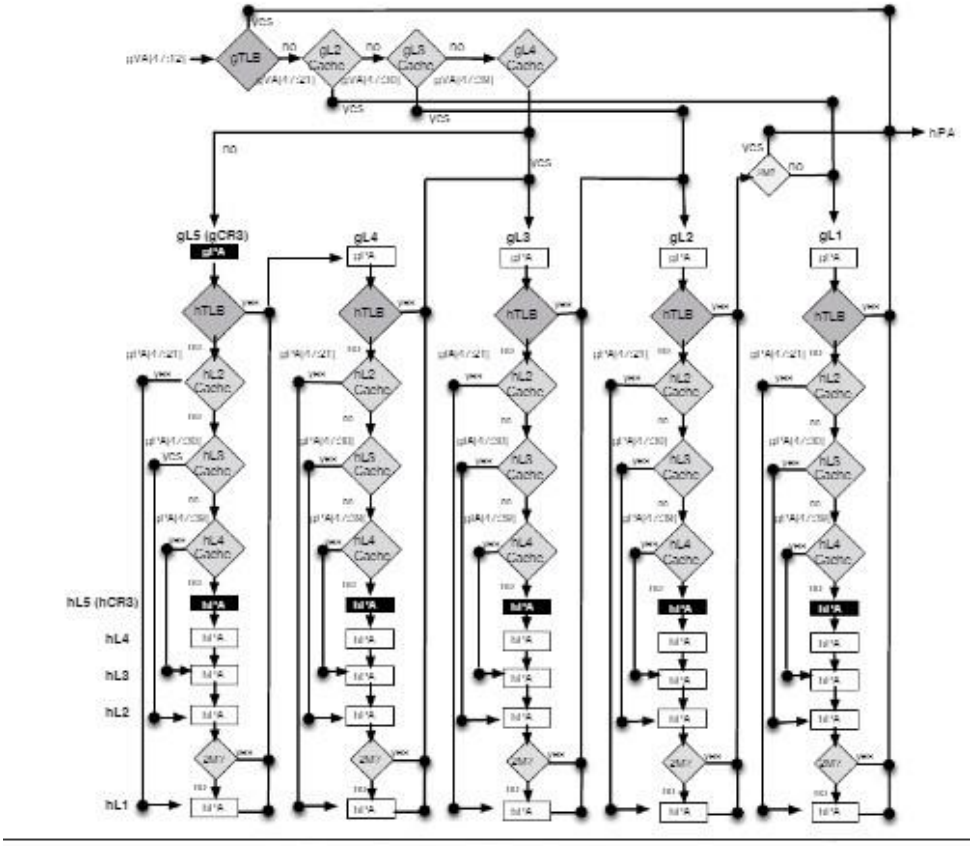


Figure 2. The x86 2-D (virtual) page walk.

EPT induced VM exit:

EPT misconfigurations	EPT paging-structure entry that contains an unsupported value	Bit 0 is clear bit 1 is set	
EPT violations	EPT pagingstructure entries disallow an access using the guest-physical address	an EPT paging-structure entry that is not present	
pagemodification log-full event	the logical processor determines a need to create a pagemodification log entry and the current log is full		

Accessed and Dirty Flags for EPT

Whenever the processor uses an EPT paging-structure entry as part of guest-physical-address translation, it sets the accessed flag in that entry		
Whenever there is a write to a guest-physical address, the processor sets the dirty flag		
These flags are “sticky,” meaning that, once set, the processor does not clear them; only software can clear them.		

Page-Modification Logging Memory type:

EPT 页表的 memory type：

If CR0.CD = 0, ept 页表的 memory type 取决于 extended-page-table pointer (EPTP) 2:0 ， 0--uncacheable type (UC)，6write-back type
(WB)

If CR0.CD = 0，ept 页表的 memory type，uncacheable (UC)；

Memory type for Translated Guest-Physical Addresses 三个因素：

1.CR0 CD；

2.last EPT paging-structure entry；

3. PAT memory type；

无标签