VM exit相关问题

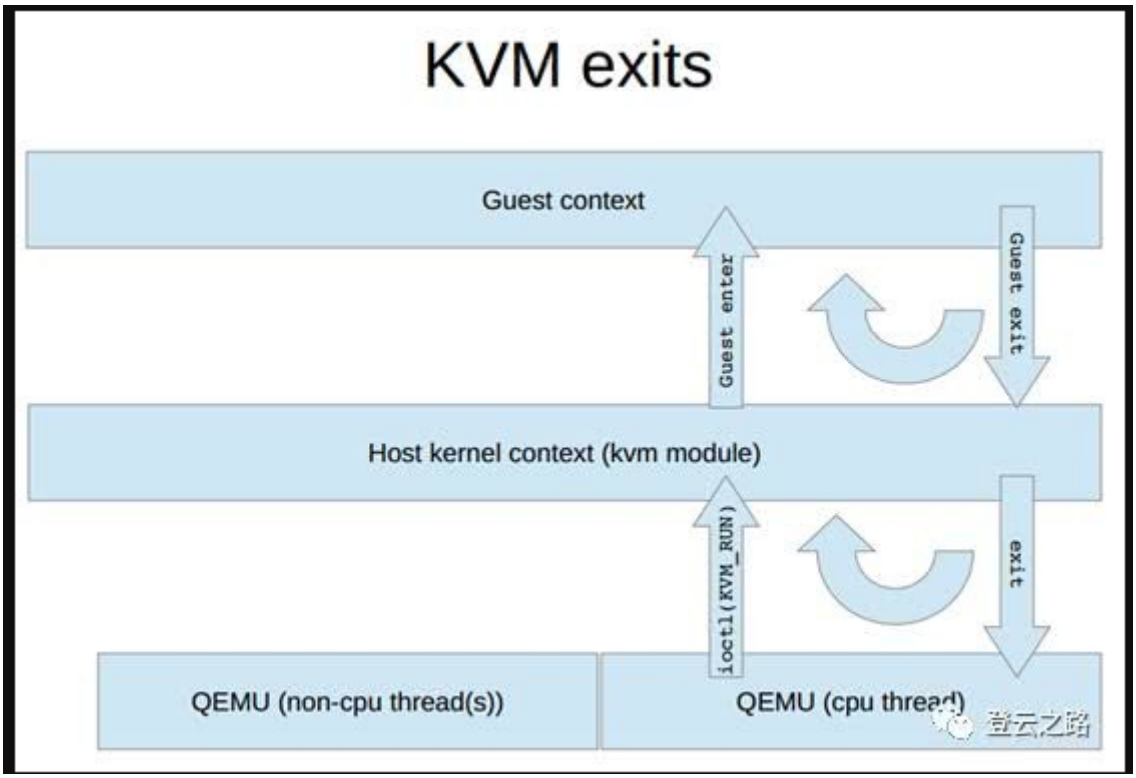
汇报人：奚宇博 贺征

VM exit类别

VM-EIXT大致分为：

- 由外部中断引起的
- 由敏感指令引起的
  - 特权指令
  - 导致错误的指令

VM-EXIT退出流程



VM-EXIT退出场景

退出原因	详细描述	解决方法
FAILED_VMENTRY	进入虚机失败	
EXCEPTION_NMI	不可屏蔽的中断异常	
EXTERNAL_INTERRUPT	外部中断	
TRIPLE_FAULT	连续错误	
PENDING_INTERRUPT	等待中断	
NMI_WINDOW	不可屏蔽中断窗口	
TASK_SWITCH	任务切换	
CPUID	返回处理器标识和功能信息	
HLT	特权指令。会暂停中央处理单元（CPU），直到触发下一个外部中断为止	
INVD	特权指令。当处理器以保护模式运行时，程序或过程的CPL必须为0才能执行此指令，使内部缓存无效	
INVLPG	特权指令。使tlb条目无效	
RDPMC	读取性能监控计数器	
RDTSC	读取时间戳计数器	
VMCALL	指令允许guest软件可以向基础VM监视器发出服务呼叫	
VMCLEAR	将VMCS数据复制到内存中的VMCS区域。	
VMLAUNCH	启动由当前VMCS管理的虚拟机	
VMPTRLD	从内存中加载当前的VMCS指针	
VMPTRST	将当前的VMCS指针存储到内存中	
VMREAD	读取指定的VMCS字段	
VMRESUME	恢复由当前VMCS管理的虚拟机	
VMWRITE	写入指定的VMCS字段	
VMOFF	退出VMX操作	
VMON	输入VMX root操作	
CR_ACCESS		
DR_ACCESS		
IO_INSTRUCTION	IO指令	
MSR_READ	读取专用寄存器	

退出原因	详细描述	解决方法
MSR_WRITE	写专用寄存器	
INVALID_STATE	无效状态	
MSR_LOAD_FAIL		
MWAIT_INSTRUCTION	监视等待指令	
MONITOR_TRAP_FLAG	监视器陷阱标志	
MONITOR_INSTRUCTION	设置监视器地址	
PAUSE_INSTRUCTION	通知CPU这是一个自旋锁等待循环，因此可以优化内存和缓存访问	
MCE_DURING_VMENTRY		
TPR_BELOW_THRESHOLD		
APIC_ACCESS		
EOI_INDUCED		
GDTR_IDTR		
LDTR_TR		
EPT_VIOLATION		
EPT_MISCONFIG		
INVEPT		
RDTSCP		
PREEMPTION_TIMER	抢占计时器	
INVVPID		
WBINVD		
XSETBV		
APIC_WRITE		
INVPCID		
PML_FULL		
XSAVES		
XRSTORS		

操控VM exit的区域

VMCS（virtual-machine control structure）管理VM entry与VM exit，自动保存或恢复执行的上下文

包括：

- 客户机状态区（guest-state area）
- 主机状态区 （host-state area）
- 执行控制区 （VM-Execution Control Fields ），存放控制VM entry与VM exit的标志位

执行控制区 操控VM exit和VM entry的标志位：

- External-interrupt exiting：当设置了该标志位，所有外部中断都会导致VM exit，且客操作系统无法屏蔽这些中断
- interrupt-window exiting：当设置了该标志位，VM exit发生在客户软件准备接收中断时
- Use TPR shadow:通过CR8访问Task Priority Register(TPR)的时候,访问的是VMCS中的指针所引用的TPR影子;VMCS还包括TPR阈值:执行将TPR影子降低至TPR阈值以下的指令后,会发生VM exit

VMCS 还包括位图来提供对VM exit的选择性:

- Exception bitmap:选择哪些异常可以出发VM exit
- I/O bitmap:访问在I/O bitmap中设置的16位I/O端口后,会触发VM exit

目前大概就是这样的一些场景下会发生VM-EIXT事件，后续可以通过收集分析VM exit发生频次的实验，发现比较常见、出现频次较高或是用时较长的一些VM exit场景，做出针对性优化

关于解决方案大致有以下几种：

- 通过操控VMCS中的标志位，从而控制VM exit
- 将导致退出的主要中断绑定到固定的CPU核心上