> Virtual Machine Monitor，VMM，也称 Hypervisor

type1

Type-1 型虚拟机监控器运行在最高特权级，直接控制物理资源，并负责实现调度和资源管理等功能（可理解为一种特殊的操作系统）

type2

需要依托一个宿主操作系统（如 linux，windows），Type-2 型 VMM 可以复用操作系统的调度和资源管理等功能

可虚拟化架构

下陷（trap）： 指 CPU 特权级从低特权级（如 EL0）切换到高特权级 （如 EL1）

特权指令（prilileged instruction）： 指在用户态执行时会触发下陷的指令。（如软中断，和不允许在用户态执行的指令等）

敏感指令（sensitive instruction）： 指管理系统物理资源或更换 CPU 状态的指令。（如 I/O 等）

可虚拟化架构： 所有敏感指令都是特权指令。（如果有些敏感指令在用户态执行不会触发下陷，那这些指令就不会被虚拟机监控器捕捉到）

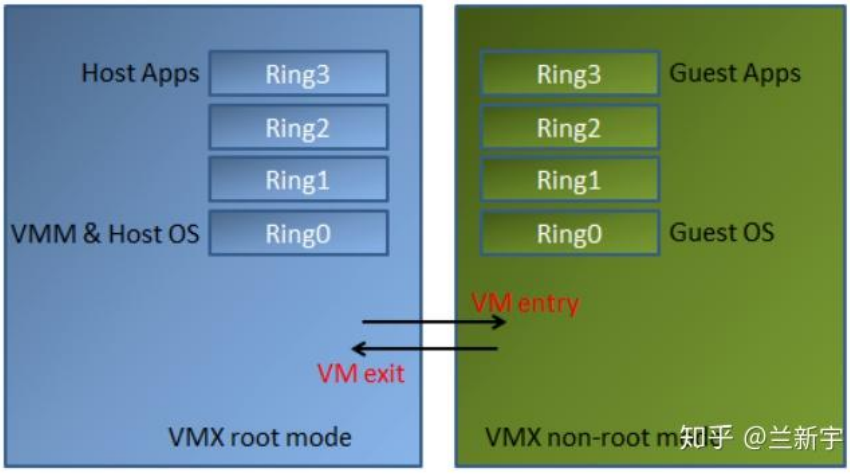
由于客户操作系统运行在 host 的用户态，而虚拟机监控器运行在 host 的内核态。但同时 guest 在自己的虚拟内核态执行敏感指令的时候，事实上是在 host 的用户态，如果不能下陷的话，VMM 就无法感知指令的执行。

硬件虚拟化技术

> guest OS 中的各个线程/进程分时复用了 vCPU，而各个 vCPU 又在 VMM 控制下分时复用了 pCPU

Intel VT-x（Virtualization Technology for x86）

> 从硬件上提供对 vCPU 调度和切换的支持



VM exit

> 从 guest VM 进入 VMM；non-root mode → root mode

- 执行 VMCALL 指令，被称为 hyper call （类似 syscall 的实现）
- 发生了硬件中断或软件异常
- guest VM 执行敏感指令

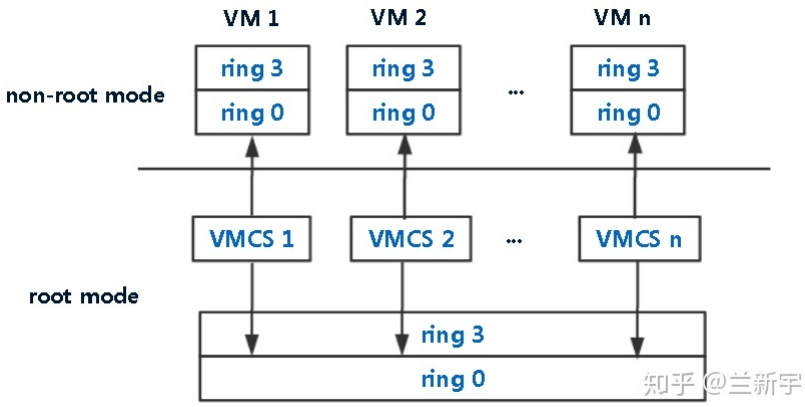
VM entry

> 从 VMM 进入 guest VM；root mode → non-root mode

mode 的切换导致上下文的保存和恢复

VMCS（Virtual Machine Control data Structures）

> 负责保存 vCPU 需要的相关状态和上下文信息



pCPU 只能运行一个 vCPU，所以同一时间只能与一个 VMCS 绑定