Nested Virtualization on Xen

Nov. 2009

Qing He <qing.he@intel.com>

Xen Summit Asia 2009

Agenda

- Overview
- Architecture
- Principles and operations
- Status

Background

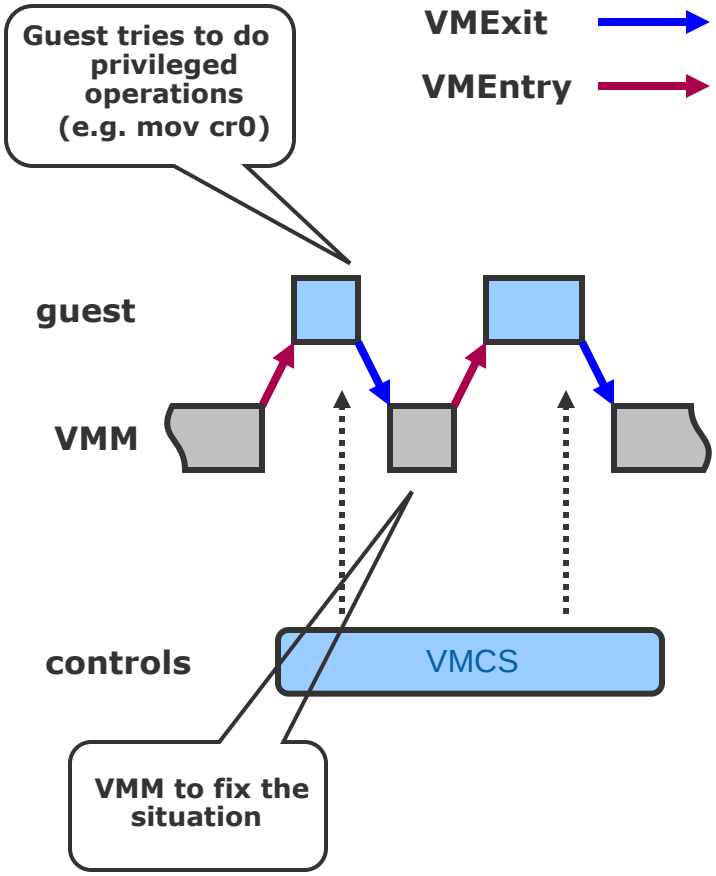
- What is nested virtualization?
  - Virtual machines inside virtual machine
    - Running a VMM inside a guest
  - Specifically, hardware-based: e.g. VMX
- Why nested virtualization?
  - Virtualization becoming ubiquitous
    - Clouds, Xen Client
  - Use of hardware virtualization in ordinary OS
    - Windows 7, XP compatibility mode
  - Facility for investigating VMM behavior

The fundamental idea

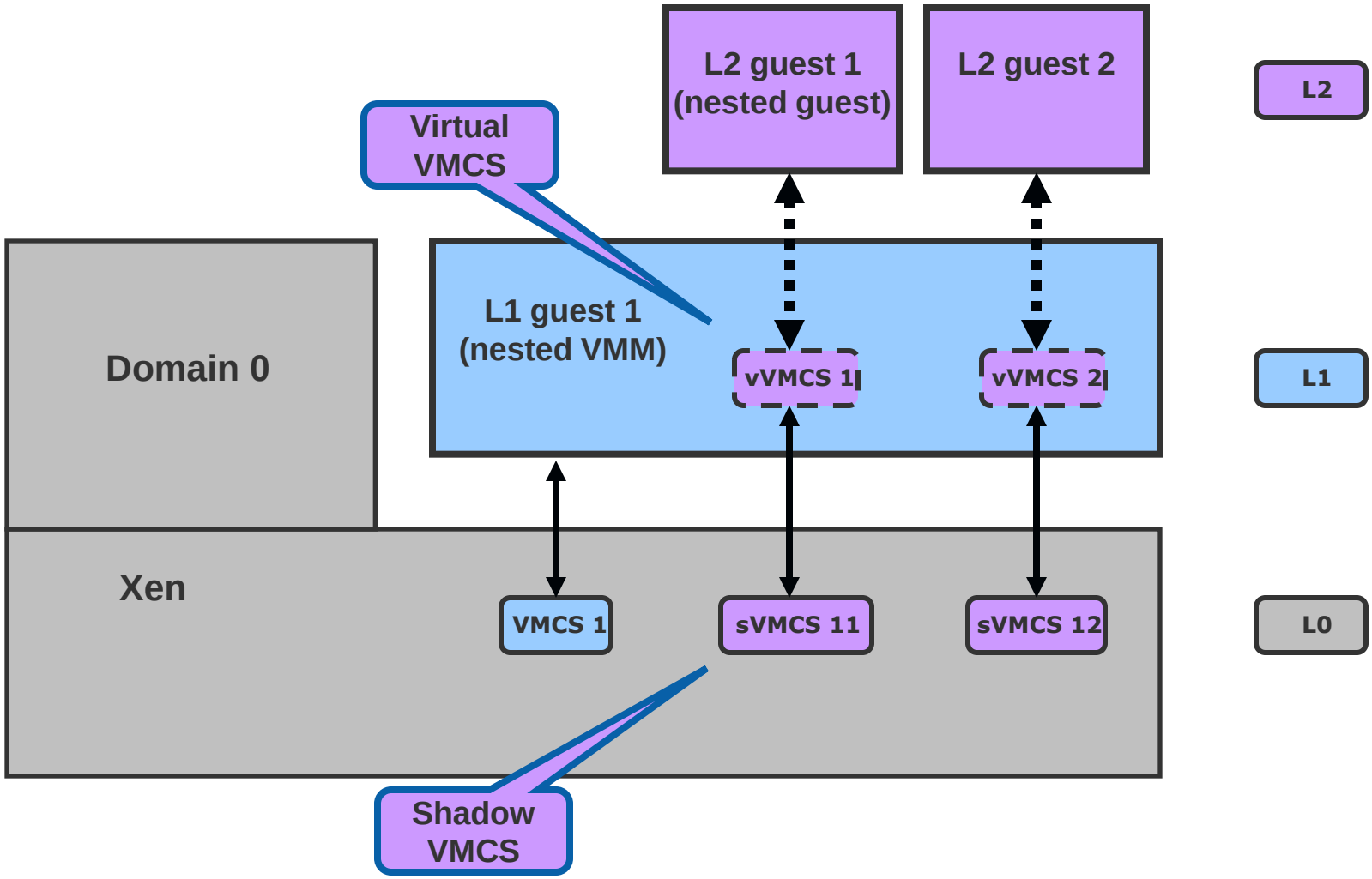
- Target: virtualization of VMX
  - Present a virtualized VMX to guest
    - VMX data structure
    - VMX instructions
    - VMX execution flow

VMX revisit

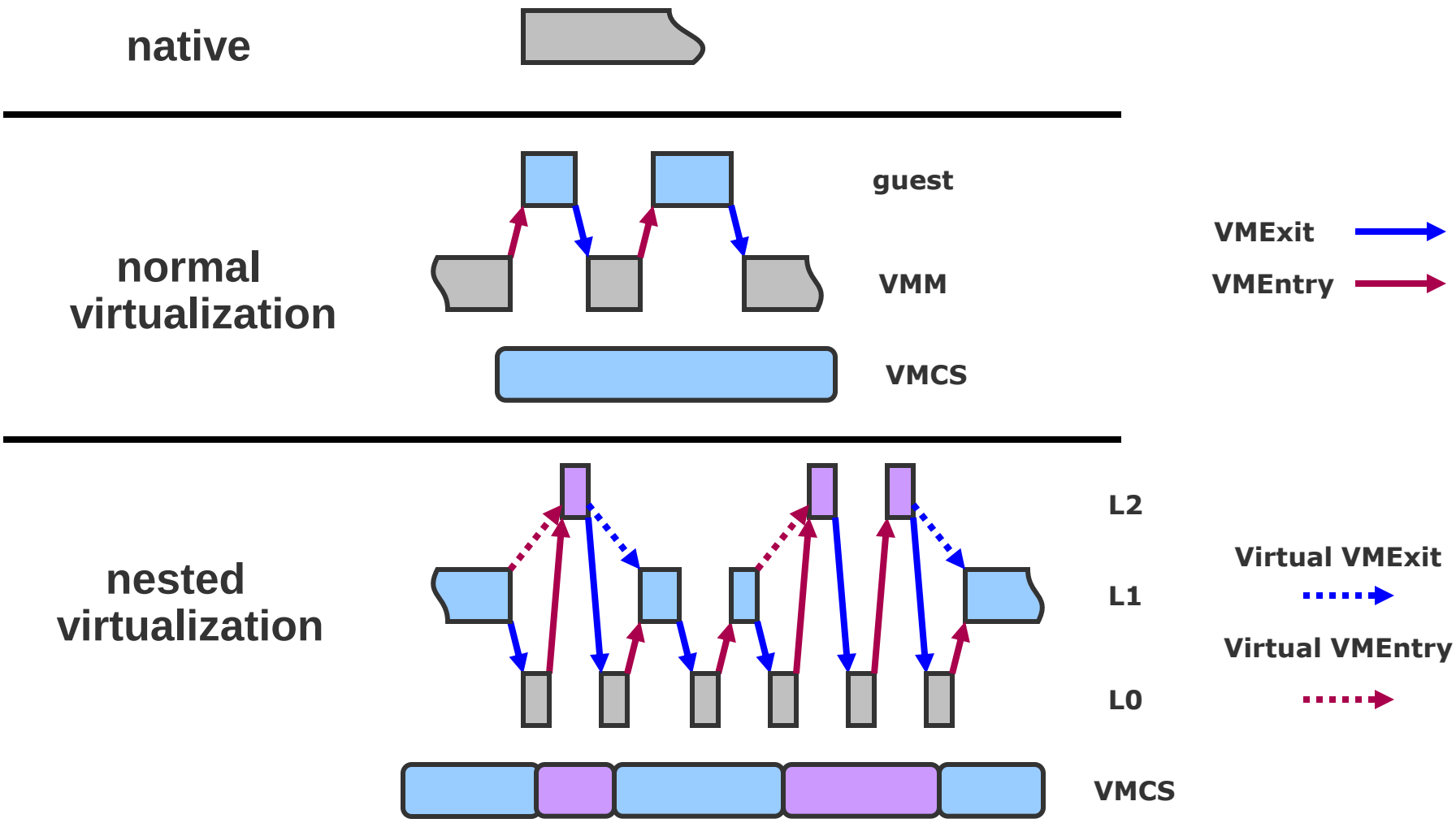
- VMX key concepts
  - — Control structure: VMCS
  - — Execution flow, VMM to guest: VMEntry
  - — Execution flow, guest to VMM: VMExit
- VMM to fix guest exits
- VMCS controls the VM
  - Guest running context
  - When the guest exits
  - Information exchange



Nested virtualization architecture



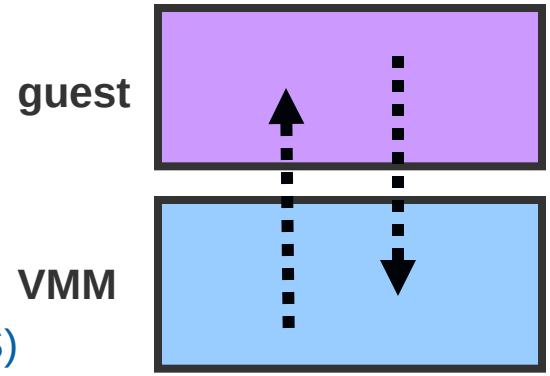
VMX execution flow



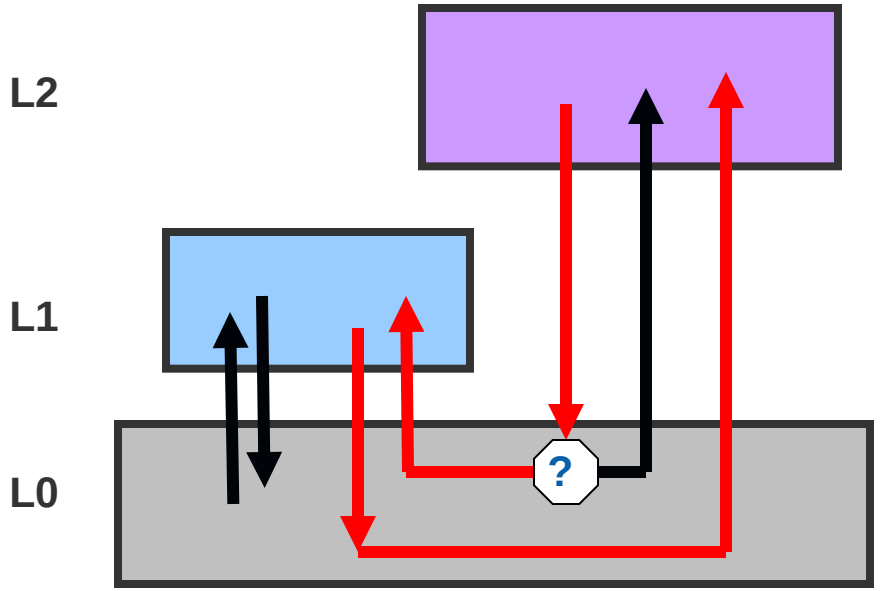
Execution flow as guest switch

- Consider nested guests also as guests
- Virtual VMEntry
  - L1->L0; guest switch;
  - L0->L2 (GUEST_RIP in virtual VMCS)
- Virtual VMExit
  - L2->L0;
  - Virtual VMExit? guest switch;
  - L0->L1 (HOST_RIP in virtual VMCS)
- Other VMExits

- Lightweight guest switch
  - In the same vcpu context



L1 guest point of view



Memory virtualization

- No special handing for shadow memory
  - Pure software
  - However, the performance is bad
    - Virtual VMExits is much longer than on hardware
- Nested EPT will be very helpful
  - Present EPT to guest
  - Significantly reduce number of virtual VMExits

Status

- POC for simple scenario
  - single cpu, one nested guest
  - Some VMX optimizations turned off
  - No suspend/resume/migration
- Nested guest can boot to an early stage
  - BIOS booting successfully on KVM as nested VMM
- Will stabilize it and refine it before send out for review

Questions?

Legal Information


intel®
Leap ahead™