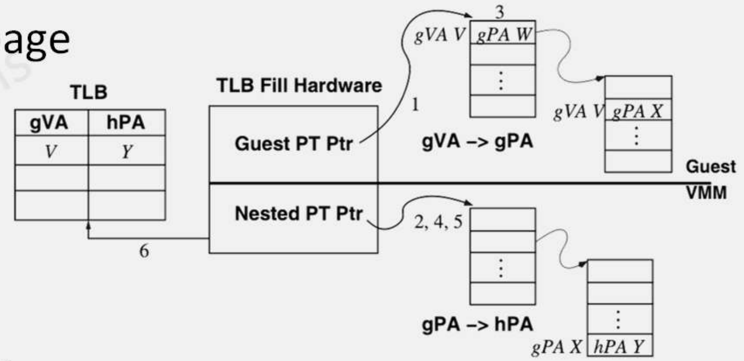
CLOUD
COMPUTING
APPLICATIONS

Virtualization: 2nd & 3rd Gen Hardware Virtualization

Prof. Reza Farivar

Second Generation Hardware Virtualization

- AMD’s RVI and Intel’s EPT (Extended Page Tables)
- The VMM maintains a hardware-walked “nested page table” that translates gPAs to hPAs
  - eliminating the need for VMM interposition
- Many issues of the first-gen are resolved
  - No trace-induced `exits`
  - no context-switch `exits`
  - no hidden/true fault `exits`
  - The VMM does not have to allocate memory for shadow page tables, reducing memory usage
- The cost to service a TLB miss will be higher with nested paging than without
  - TLB Caching helps a lot
  - Large Memory pages (1 GB vs 2 MB)



I/O Virtualization

- Most hypervisors “emulate” I/O devices
  - Generic display
  - Generic network
  - Generic storage
- Trap and Emulate idea
  - Or paravirtualization
- Cloud Data Center requirements necessitate optimal performance
  - Hardware-based I/O Virtualization

Third Generation Hardware Virtualization

- Since the Haswell microarchitecture (announced in 2013), Intel started to include VMCS shadowing as a technology that accelerates **nested virtualization** of VMMs
- Interrupt Virtualization (AMD AVIC and Intel APICv) 2012
- I/O MMU virtualization (AMD-Vi and Intel VT-d)
  - An input/output memory management unit (IOMMU) allows guest virtual machines to directly use peripheral devices, such as Ethernet, accelerated graphics cards, and hard-drive controllers, through DMA and interrupt remapping.
  - This is sometimes called PCI passthrough
- PCI-SIG Single Root I/O Virtualization (SR-IOV)