Kernel-based Virtual Machine (KVM)
Memory Management Unit (MMU)
Virtualization

Mohammad H. Mofrad
February 11, 2016
Contents

• Kernel-based Virtual Machine
• Linux memory management unit
• memory management unit virtualization
KVM as a Virtual Machine Monitor (KVM)

```c
open("/dev/kvm")
ioctl(KVM_CREATE_VM)
ioctl(KVM_CREATE_VCPU)
ioctl(KVM_MEMORY_REGION)
for (;;) {
    ioctl(KVM_RUN)
    switch (exit_reason) {
        case KVM_EXIT_IO: /* ... */
        case KVM_EXIT_HLT: /* ... */
    }
}
```
A Memory Management Unit (MMU) translates physical memory addresses to virtual address used by the operating system.

It requests a page fault interrupt

MMU uses a Translation Lookaside Buffer (TLB) to improve virtual address translation speed.

Source: The Linux Kernel's Memory Management Unit API, William Gatliiff
MMU in Virtualization

Host Virtual Address (HVA)
Host Physical Address (HPA)
Guest Virtual Address (GVA)
Guest Physical Address (GPA)
Memory virtualization problems

• 3 levels of indirection
• But MMU can translate only one
• But MMU virtualization translates GVA to HPA.
Memory Virtualization

• Shadow paging
• Hardware assisted virtualization
• MMU paravirtualization
Shadow Paging

Guest process page table

-----------------------

|                      |
-----------------------

Page fault >> VM EXIT

Shadow page table

-----------------------

|                      |
-----------------------

HVA

1|2|3|4|5|6|7|8|9|10|...

HPA

A|B|C|D|E|F|G|H|I|J|...

GPA

1|2|3|4|5|6|...

GVA

1|2|3|4|...

KVM/Qemu GPA >> HPA

GVA >> GPA

Shadow page table

---

Page fault >> VM EXIT

Page table

-----------------------

|                      |
-----------------------

VM GPA >> GPA

---

Shadow page table

---

Guest process page table

---

Page fault >> VM EXIT

---

Shadow page table

---
Shadow Paging pros and cons

+ + Software solution
+ + No need to change guest OS
+ + No special hardware is required
-
-
- - Every guest page needs a shadow version
- - Shadow page table must be consistent with guest and host
Hardware Assisted Virtualization (NTP/EPT)

- Eliminate the need for shadow paging (software approach)
- Provide architectural support
- Nested Page Table (NPT) by AMD
- Extended Page Table (EPT) by Intel
Hardware Assisted Virtualization (NTP/EPT)

KVM/Qemu

GPA >> HPA

N/ETP table

Page fault >> NTP/EPT violation

Guest process page table

-----------------------

<table>
<thead>
<tr>
<th>N/ETP page table</th>
</tr>
</thead>
</table>

-----------------------

<table>
<thead>
<tr>
<th>VM</th>
</tr>
</thead>
</table>

GVA

GPA

HVA

HPA

1|2|3|4|5|6|

Page fault >> NTP/EPT violation

1|2|3|4|5|6|7|8|9|10|

1|2|3|4|5|6|

1|2|3|4|5|6|7|8|9|10|

A|B|C|D|E|F|G|H|I|J|...
NTP/EPT Pros and Cons

+ + OS independent
+ + No need to change guest OS
+ + NTP/EPT table is proportional to guest memory
- - Need hardware support
- - Considerable overhead for TLB misses
- - More memory access + longer latency
MMU Paravirtualization

• Paravirtualization involves modifying OS kernel to replace nonvirtualize instructions with virtualization hypercalls.

• MMU paravirtualization is one of the original innovations of Xen VMM.

• So, guest is aware of how to map GVA to HPA!!!
MMU Paravirtualization

MMU Paravirtualization Pros and Cons

+ + Near native solution
+ + Low memory overhead while taking full advantage of hardware
- - Compatibility cost [modifying the guest kernel]
- - All updates/changes need to be validated by VMM
- - More memory access + longer latency
Useful References

