Midterm-Report

11911609 葛兆宁

Part1 Terminologies of virtualization

(1) The virtualization is something like illusion, just make you think you own it, but actually maybe not. Just like the process virtualize the CPU resource or the address space virtualize the memory resource, even the VMM technology. For example, the VMM technology makes the OSes on it "think" they are directly in control of those physical resources, however only the VMM is directly in control of those physical resources.

(2)

- workload isolation:

```
workload isolation means isolate the software stacks in their own OSes to
get more reliability and security.
```

- workload consolidation:

```
workload consolidation means run multiple OS(especially with single
workload) with one single physical platform.
```

- workload migration:

```
workload migration means use virtualization to encapsulating a running OS
to migrate to other platform.
```

(3)Microsoft Virtual PC, VMware Workstation, Sun Virtual Box

(4)

- paravirtualization:

```
paravirtualization means the VMM is not fully virtualize all the physical
resource for the OS, so the OS can only run on a paravirtualization VMM
after being modified.
```

- full-virtualization:

```
Full virtualization means the VMM provide all the physical resources
virtualization for the OSes.
```

- binary translation:

```
It's a technique that it'll interpret the instructions in guest OS to the
instructions in VMM, in one aspect , it's used to avoid some unexpected
trap in physical CPU, in another aspect, it's used to prevent some
privileged states appearing in physical CPU, and reserved it only in
virtual CPU.
```

- hardware-assisted virtualization:

```
It means the VMM need some hardwares and maybe the software in that to
assist virtualization.
```

- hybrid virtualization

```
It's a technique that combine other techniques together to get a better
result, like combine different techniques' parts together to reap the
length, or design a heuristic function to which technique should be used.
```

Part2

(1) 4, user processes' instructions will always in low level privilege(only use system call), OS will always in high level privilege. Examples: address space accessible, instructions schedule, interrupt.

(2) Ring compression: get OS's privilege down to lower than VMM.

(3) For segmentation, it will see level 1,2,3 if fit, for paging, it will only in level3.

(4) For 64 , it will see level 1,2,3 if fit

(5) ring aliasing means the process find it's privilege level is different to the level it should be.

(6) VMX root and VMX none-root are two modes of operations in VT-x, while the root mode is for Xen processes and the none-root mode is for the OSes processes. Each mode both have 4 privilege level.

(7) VT-x allows Xen processes to run in the VMX root mode, and OSes processes to run in VMX none-root mode, because each mode has four privilege levels, thus the process will run in its original privilege level so that it will not get the ring aliasing. On the other hand, the none-root mode's privilege level is always lower than the root mode's, so it just ensures the ring compression.

Part3

(1) The system call is used to call kernel service by user process.
The function call is realised in user mode with a low privilege, on the contrary, the system call is realised in kernel mode with a high privilege.
The way to pass the parameter is firstly store the parameter in the specific registers, then call a system trap, the trap handler will change to the kernel mode, and get parameters from those registers.

(2) The hypercall is a special function call that make from the guest OS to request the service of Xen's hypervisor.

(3) It virtualizes a table with each exception registered with Xen for validation. When a exception in guest OS happens, it will make a copy of exception stack and return to the right exception handler.
The sole modification is page fault handler, which is written in extended stack frame.

(4)External interrupts unrelated to guests are intercepted when VM exits, and virtualized interrupts are being injected when VM entry to the guests

(5) It is virtualized by a general-purpose mechanism called event channels.
It makes OS can share some interrupt.

(6)A virtual-machine control structure (VMCS) is used to manage VM entries and exits, and it controls the instructions in a non-root operation.
The VM-entry is the transition from the VMX root operation to VMX non-root operation, the VM exit is the transition from the VMX none-root operation to VMX root operation.
VMCS is logically divided into sections, the guest-state area and the host-state area. These sections contain fields related to different components of processor state. The processor state is loaded by the VM entries from the guest-state area. VM exits will firstly save processor state in the guest-state area and then load processor state in the host-state area.

(7) The Xen leverage Intel VT-x to virtualize interrupts by external-interrupt exiting VM-execution control and an interrupt-window exiting VM-execution control

(8) The VT-x use the exception bitmap to support exceptions.

Part4

(1) Firstly, the VMM get the guest linear address, and then it check it with the address and offset from the IA-32 Page Table, it get the next EPT page table, with the second offset and EPT page table , the VMM get the host physical address at last.

(2) Firstly, the VMM get the guest linear address, and then it check it with the address and offset from the Intel 64 Page Table, it get the next EPT page table, with the second offset and EPT page table , the VMM get the host physical address at last.

(3)The guest virtual address can be translated into the guest physical address by page table provided by the guest OS, the guest physical address can be translated into the machine address by page table supported by the VMM.

(4) Use OS itself to manage its process page table. Use VHPT to manage OS page table.

(5)The term address-space compression means that the difficulties of protecting these parts of the virtual-address space and supporting the guest to access to them.

(6) The linear-address space can be changed by any transition between guest software and the VMM , allowing the guest software can fully use its own address space.

(7) EPT's full name is Extended page tables, it's a page table to mitigate some of the page table shadowing related overheads.
The MMU will mantain and update EPT.

(8) Firstly it will allocate with domain, and memory will be divided into domains static. However, if memory pressure is within a domain increases, the VMM will enlarge the domain capacity, otherwise the VMM will shride the domain's capacity.