

Sample Audio From Mark Minasi's 15-CD "Installing, Troubleshooting and Maintaining Server 2008"

from www.minasi.com/2008class/audio

Hardware Hurdles

but if both the OS and the VMM are in ring 0...

- It *should* be impossible, then, to create virtual 32-bit machine managers securely atop an x86 processor, as both the OSes and the VMM run with equal power
- And if a virtual machine's OS runs in ring 0 – alongside the VMM in ring 0 – then why don't errant OSes on virtual machines tend to crash virtual machine managers?

Hardware Hurdles

answer: "ring compression"

- Through a bit of deception called "ring compression"
- The virtual machine manager runs in ring 0, as expected...
- ... and it fools the VM's kernels into running in ring 1!

Hardware Hurdles

does that *work*?

- Yes, most of the time, and fairly well given the seeming impossibility of the whole thing
- But 10 opcodes have the potential of blowing the whole thing up (as they assume that they live in ring 0 and may just shove their way into there if necessary), and so VM managers have to watch their VMs very carefully
- This requires a lot of shifting between rings, and that costs time
- If *only* we had some hardware help...

Hardware Hurdles

sidebar: back to Intel and AMD

- The appearance of the Pentium in March 1993 began about 13 years of slow but non-revolutionary progress in CPUs ... interesting when you realize that the first 32-bit x86 CPU appeared in 1985 (that's 23 years of largely unchanged CPU structure!)
- That stabilization enabled AMD the time to clone the Pentium, creating a faster, cheaper, *compatible* alternative

Hardware Hurdles

sidebar: back to Intel and AMD

- Meanwhile, Intel turned its attention to creating a new, cleaner 64 bit platform, the Itanium...
- ... that flopped
- AMD exploited this by developing a 64-bit chip that was more Pentium-compatible and cheaper than the Itanium
- All of sudden, it was silly to buy a 32-bit Pentium when AMD sold a faster 64-bit alternative

Hardware Hurdles

what's this got to do with virtualization?

- AMD, employing the time-honored tactic of "kick 'em when they're down," announced that it was high time that we had a Pentium-compatible chip that gave VMMs a better shot at virtualizing entire Pentiums
- The beta name for the "virtual-friendly" chips was "Pacifica," and once ready, you could identify a virtual-friendly chip because it had the suffix "-V" in its name

Hardware Hurdles

surely Intel didn't take this lying down?

- Intel, meanwhile, hastened to announce that they, too, were working on chips like that, the code name was "Vanderpool," and once shipped, the chips had a "-VT" suffix
- All shipped around 2006
- All are 64 bit
- Intel's now building chips that are faster and cooler temperature-wise than AMD
- (Don't you just love when this happens?)

Hardware Hurdles

virtual-friendly chips appear

- Result: new line of 64-bit chips that include a new "ring -1" privilege
- Result: VMMs can run in ring -1 and the virtual OSes are largely unaware that they're being virtualized
- AMD's implementation is a superset of Intel's
- Not compatible, so VMMs must recognize and handle AMD and Intel changes, or just limit itself to the common capabilities

Hardware Hurdles

more on V/VT

- Bottom line: we finally have Pentiums that can fully virtualize Pentiums!
- Intel manages VM memory in software, AMD does it in hardware
- Adding "ring -1" makes building VMWare, Virtual Server and XenServer much easier, and makes the VMMs more reliable

Hardware Hurdles

security side-effect and BIOSes

- Modern (most post-2006) VMMs can use ring -1
- But ring -1 offers a big possibility of bad news... a virtual rootkit, like Blue Pill
- Imagine a rootkit that you could not see under any circumstances, as it lives outside of ring 0!
- So most BIOSes disable ring -1 by default

Summarizing so far:

so this means...

- Hyper-V needs 64 bit "V" or "VT" chips
- You must be able to enable ring -1 in your BIOS, which is probably called simply "virtualization"
- BTW, you also need hardware data execution prevention enabled ("DEP"), again in the BIOS – it's probably already on, but check