Xen 3.0 - What is it?

By Luke St.Clair
Overview

- Memory Management
- Interrupts
- Device I/O
- Starvation
- HVM
- And so much more...
Memory

- Start of Day - Initial page, and address of the translation
- Domain builder takes over
- Different types of pages - like pt, rw
- Different types, different restrictions
- Read only pt, can’t be rw / pt, etc.
- Two lookups - physical to machine, and machine to physical
General

- SMP Guests
- HVM
- x86_64
- Driver Domains?
- Migration
- XenStore
Page Tables, Default

- Read only
- explicit hypervisor requests
- much like standard linux
- ref counts
- final result - a lot of the updates are "batched", but this requires a lot of guest OS modifications
Page Tables, writing

- Domain tries to write
- Disconnected Page
- Reconnected
  - pre-emption
  - explicit calls
Page Tables, Shadow Mode

• DomU keeps its own tables - separate from HW
• Xen then takes changes and puts them in HW
• Basically, a copy in domU and a copy in dom0
Shared Info Pages

- Pages shared between VMM and guest OS
  - timing
  - interrupt vectors
- Basically, communicates VCPU state
Interrupts

- Xen “event” == “hardware interrupt”
- `event_channel_op` - controls event channels from domU
- domU registers PIRQs, just like real IRQs, and waits for notifications from them
- domUs can talk to each other
Devices

• Frontend - everyone can use
• Backend - direct access to hardware. Can be dom0 or driver domain
  • Serious, serious safety concerns - see the safe hardware paper, ironically
• Block - thru VBD driver
Networking Devices

- As mentioned before...
- 2 rings - one transmit, one receive
- domU gives up memory, to get back memory
• Memory - Not an issue
• CPU - Not an issue
  • scheduler can now be changed, via hypercall - DOM0_SCHEDCTL
• I/O - Still a problem
  • Can this be fixed?
CPU Scheduler

- BVT
- Atropos - Atropos is a soft real time scheduler. It provides guarantees about absolute shares of the CPU, with a facility for sharing slack CPU time on a best-effort basis. It can provide timeliness guarantees for latency-sensitive domains.
- Round Robin
- Edf
HVM

- Unified interface to VT-X and Pacifica
- Support unmodified VMs
  - If you can install on 386, it probably works
- Right now, a little slow
More HVM

• Adds a “ring -1”
• Adds VM_ENTER, VM_EXIT
• hw support for switching VMs
• If you want to see the details, go to Intel’s http://www.intel.com/technology/computing/vtech
Migration

- Let the other talk speak for itself
Future Work

- Better HVM
- VMWare killer?
- Coalitions
- LVM snapshotting that isn’t terrible
Didn’t Mention

- Grant Tables
- Pausing/Unpausing Domains
- Networking Testing