

Whats up at the virtualization/ emulation front?

Christian Horn chorn@fluxcoil.net

January 4, 2010

...Christian Horn?

- OpenSource enthusiast, Linux Engineer, Sysadmin
- current main topics cobbler, kerberos, ldap, kvm
- Japan fan, cycling, reading
- first computer was a KC85/3 build in hometown Muehlhausen
- RedHat Admin: using Redhat since 5.2, certified RHCE in 2002, RHCA in 2009
- mostly redhat/suse at work, debian on private boxes

on the content

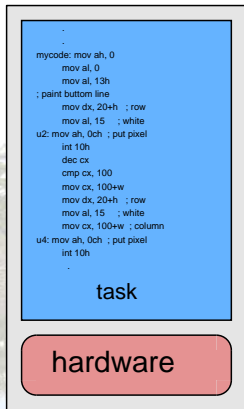
notice

- This is virtualization in a nutshell as i see it.
- I will skim through the technology in chronological order, basic mechanisms will be explained in between.
- Will just name the most important software from my point of view, will simplify things to some level.
- If you have questions, don't wait to ask them! If the question will take too long to answer or take us too far off topic, we'll move on and you can talk to me afterward.

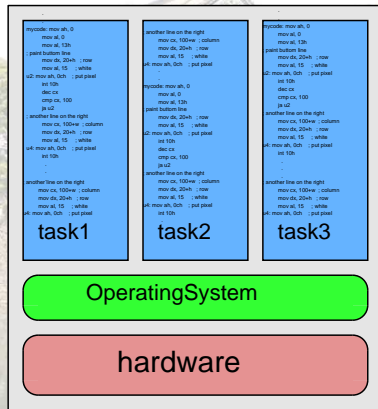
IBM System/360 Model 67

- 24bit mainframe shipped since 1966
- CP/CMS works as native hypervisor (so sits directly on metal) comparable to xen-kernel or vmware ESX
- features like memory protection appearing, time-sharing appear (in past single big programs), operating systems managing multiple tasks
- funny note: first versions of CP/CMS without copyright-notes so free for publice use :)

jump single tasks -> OSs with multiple tasks



Code running directly on classic computers (think calculators)

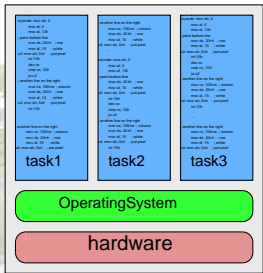


Code running as tasks in an OS which runs on the hardware

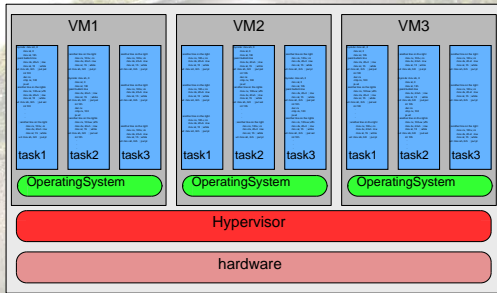
Also starting.. virtualization!

- Why virtualization? To keep old software running, the new hardware allows this now
- accomplishing full virtualization of hardware: all features used by software like i/o, cpu functions, memory etc. are provided in virtual machines -> multiple OSs can now run virtualized

development of hypervisor virtualization



Single OS running multiple tasks

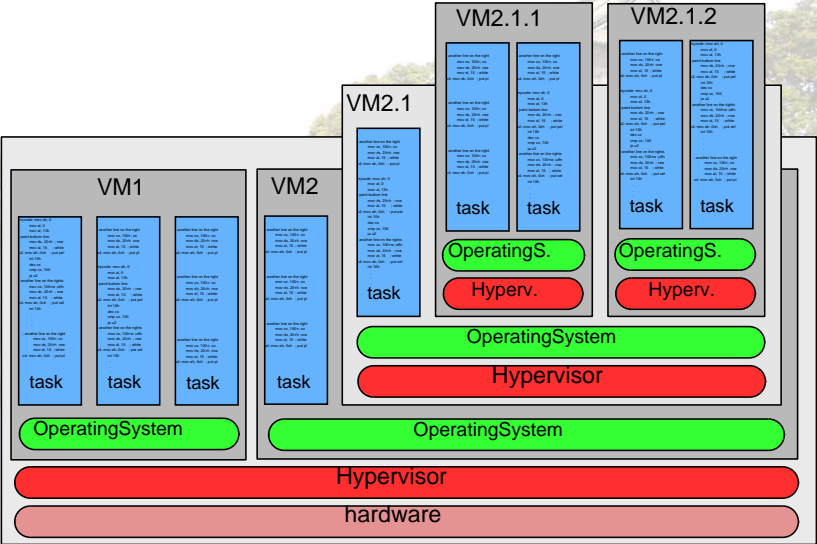


Hypervisor running multiple OS with multiple tasks


IBM System/360 Model 67 cont'd

- paravirtualization supported, aka “DIAGNOSE code” in IBM-terms
- System/360 later evolved into System/370, System/390, zSeries, System z9 and current System z10
- current OSs are z/VM or linux. A z/VM acting as hypervisor can run other z/VMs virtualized, even cascading possible
- nowadays hercules emulates z-architecture, even architectures that dont exist in reality

cascading virtual machines



the void

- 
- ...and then during the 1980s and 1990s virtualization is abandoned, the trend is to spread apps over many small lowcost computers
 - x86 architecture rises as industry standard, but lacks functions to support virtualization, grows fast and stays backward compatible

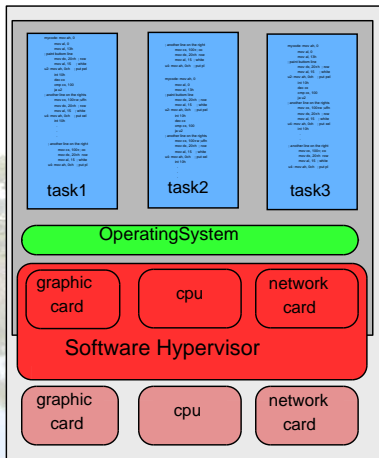
Bochs (spoken: box)

- created 1998, bought by MandrakeSoft and released under LGPL in 2000
- no virtualization, pure emulation of x86-hardware => relatively slow
- often used for OS-development when accurate emulation is required and speed doesnt count
- parts reused in other projects like qemu and kvm
- runs on: linux, macosx, bsd, win
- hosts: most x86-systems like linux, bsd, dos, win

Vmware Workstation/Fusion

- 1999 vmware publishes its 'workstation 1.0': focusing on desktops
- among fastest emulators, closed source, HVM required
- features: usb, smp (x16), complex snapshots management, seamless mode integrates guests apps in hosts windowmanager (aka unity), Directx8/9 for winguests
- VMI paradrivers interface about to be removed: cpu-functions efficient enough

software emulator components

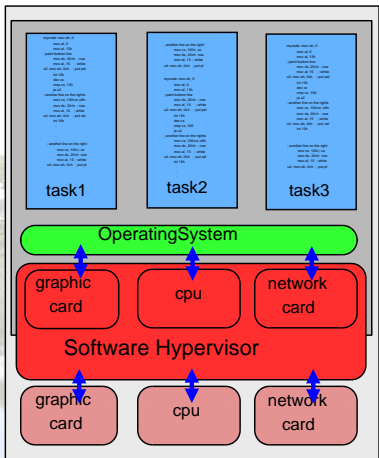


pure software emulator

i.e. Bochs



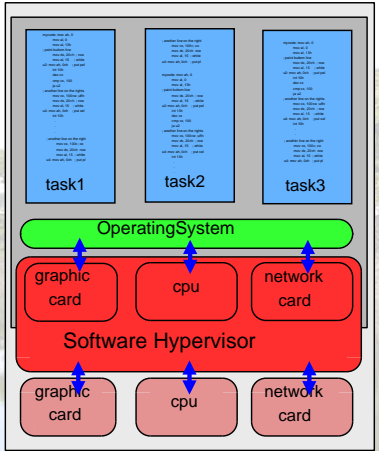
software emulator access



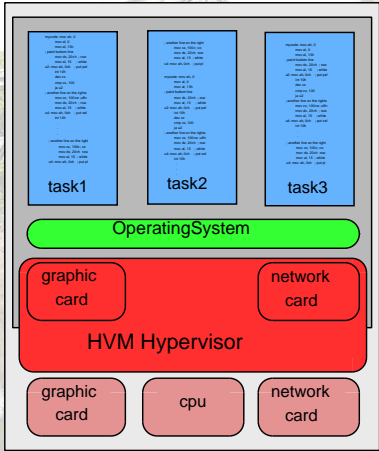
pure software emulator
i.e. Bochs



software-emu vs. virtualizer components

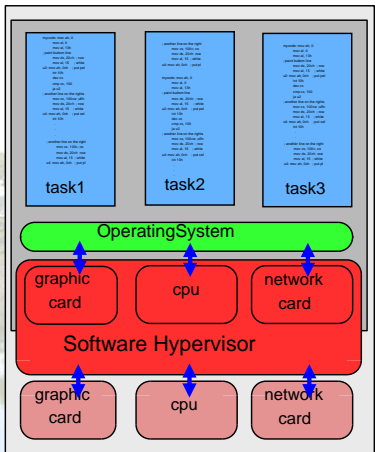


pure software emulator
i.e. Bochs

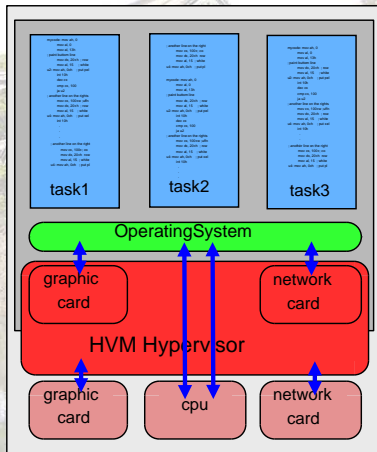


CPU virtualized
i.e. XEN / AMD-V

software-emu vs. virtualizer access



pure software emulator
i.e. Bochs



CPU virtualized
i.e. XEN / AMD-V

Vmware Server/ESX

- in 2001 introducing Vmware Server (lives on host-OS just as another process) and ESX-server (running pure ESX-kernel on hardware)
- later Vmware Player/Fusion/ESXi
- nowadays well established in industry, famous for conversion-tools, HA-functions, management GUIs, virtualizing complex network-setups

The Qemu generation

- on Mar/23/2003 Fabrice Bellard (ffmpeg, tcc) announces an “x86 emulator” for x86 and PowerPC Linux hosts: “Its main goal is to be able to run the Wine project on non-x86 architectures.”
 - mode linux-user runs linux-binaries across different archs
 - mode system-emu provides a full system
- emulated architectures: x86, ARM, SPARC, MIPS, m68k (and some others like alpha in less usable states)
- snapshots of virtual machines possible

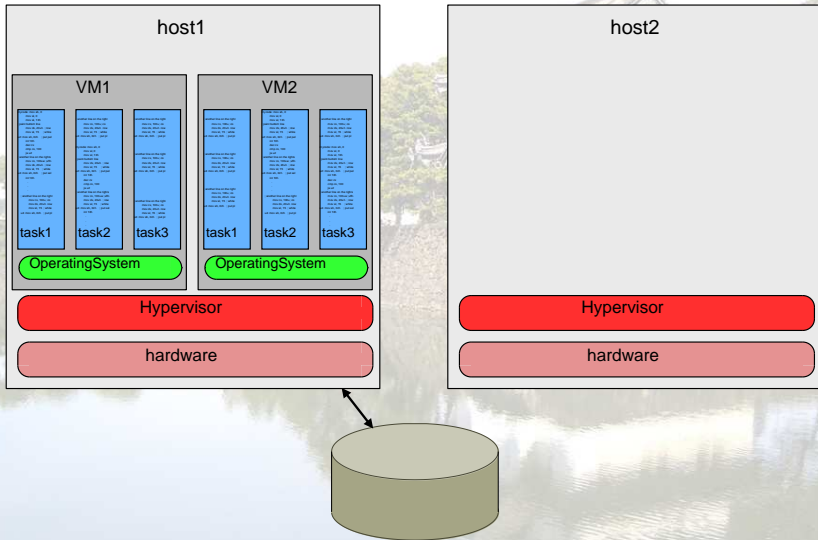
Qemu features

- much hardware emulated: haddisks, cdrom, different nics, sound chips, usb-devices, smp cpus, graphic cards
- also using hosts usb-devices possible
- qcow2 growing diskfiles (think of sparsefiles), http-blockdevices
- speedups possible by using kqemu, doesnt require HVM-cpu-technology!
- most important opensource-project in virtualization area, many reuses in xen, kvm, virtualbox, maemo
- >50% of coders also contribute to linux-kernel

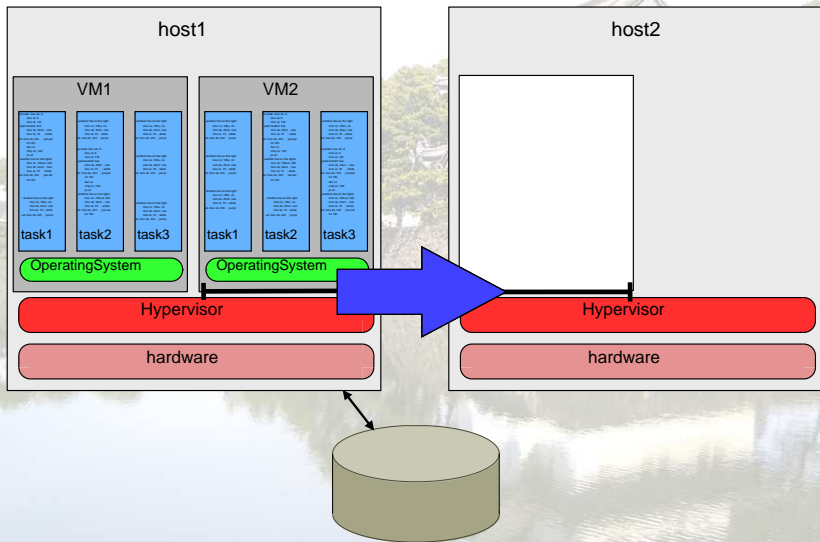
Xen

- in 2003 released by Univ. of Cambridge, nowadays owned by Citrix
- hypervisor, separated in vm 'dom0' for control and real VMs
- 2 modes: paravirt. (OS modified for operation) and fullvirt. (runs unmodified OSs)
- using QEMU-code for emulation of devices
- features: fastest soft for paravirtualization, live migration, nowadays also puts newer cpus into sleepstates to save power, GPL paradrivers, pcidevice delegation, vista++ have para-extensions for hyper-v, vmgl
- problems: fewer features than linuxkernel, code not in vanilla kernel, code heavy for adoption

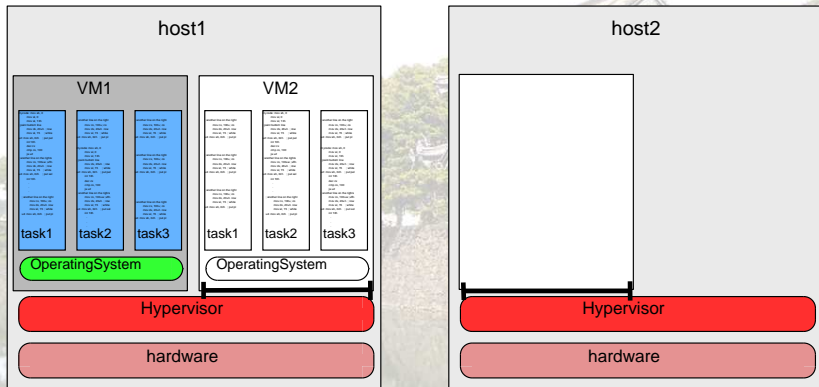
vm running on host1



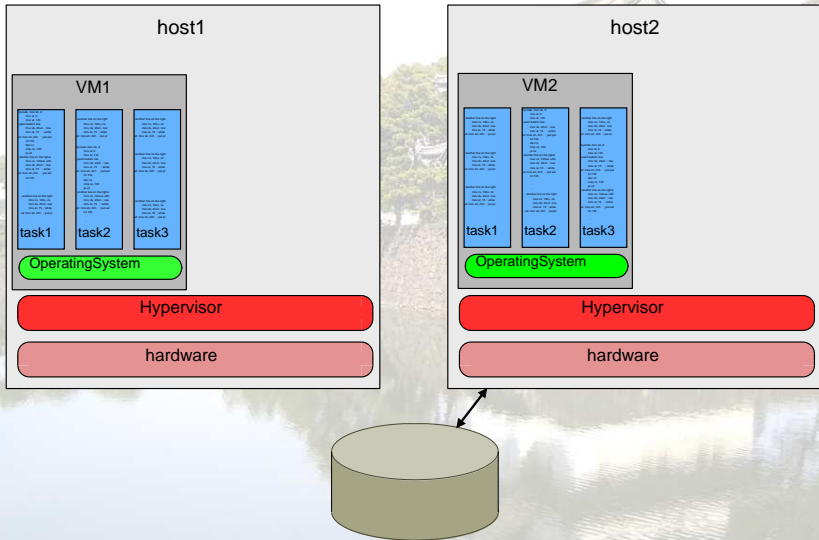
memorytransfer of running vm



freezing IO and copy of remaining memory



unfreezing IO, VM is moved



VirtualBox overview

- in 2007 Innotek (now part of SUN) releases VirtualBox OSE (OpenSource Edition) under GPLv2
- SUN sells support for the partly closed 'Personal use and Evaluation' (PUEL)-version, also adds features
- supports as hosts OS/2 and Solaris (besides usual Linux, MacOS X, Windows)
- fastest free emulator for computers without HVM-function in the CPU, comparable to vmware
- code partly from qemu, nowadays much faster development than qemu

VirtualBox features

- features: rdp-server, usb 1.1/2.0, snapshots, HVM-usage if provided, seamless mode integrates guests apps in hosts windowmanager, clipboard, opengl-drivers/Directx8/9 for winguest on linux-host, smp, livemigration to different host-OS/cputype (teleportation)
- emulates disks backed by files, iscsi, vmdk-files (vmware)
- on roadmap: paravirt windows drivers

KVM kernelbased virtual machine

- *2006 in kernel from Qumranet, later bought by RedHat
- OpenSource
- just a linux kernelmodule utilizing cpus HVM-functions => all features like powermanagement already in place, each virtual-cpu is a thread in the host
- runs unmodified guests, qemu-code emulates devices in userland
- paradrivers for windows & linux for better performance

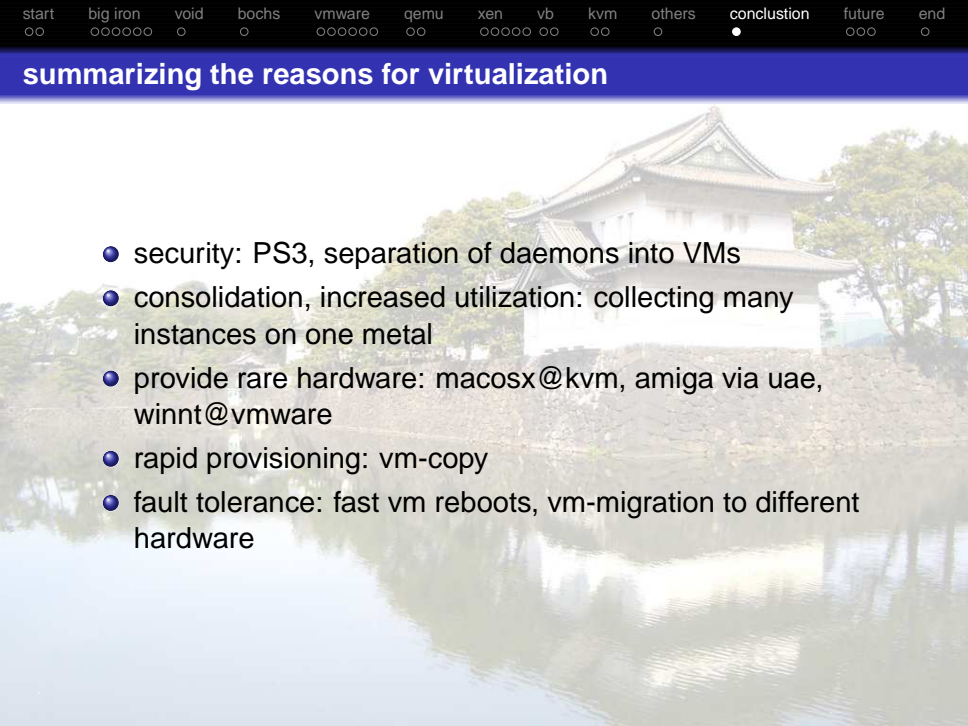
KVM features

- cpu/memory overcommit, live migration, hotplug of cpu/block/nic/pci
- SR-IOV (single root i/o virt.): native pci-e device sharing
- KSM (Kernel Same Page Merging): merges equal memory-pages into “shared pages”, think of filesystem hardlinks or Dedup => 600 VMs on a host with 48 cores and 256GB RAM
- pci dev. assignment: requires VT-d or IOMMU hardware (no memory overcommit, no migration, no graphic cards)
- Android & MacOSX virt. possible :)

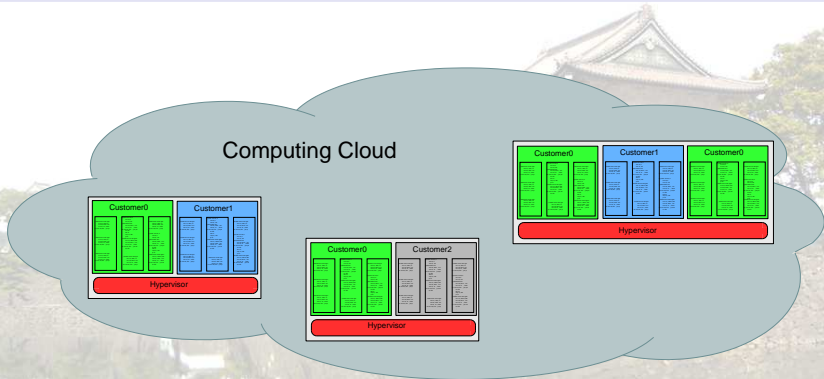
important others that didnt make it to own slides

- for one OS: wine, dosbox
- on one hostkernel: OpenVZ, bsdjails (hirarchical, networked), chroot, Solaris Zones
- sun xVM (think of xen or Hyper-V with ZFS and Gui)
- lpars on IBM power, design looks a lot like xen, also virtual scsi-traffic among lpars
- FAUmachine (Hardware Emu/Virt, also emulates faulty hardware)
- whole hardware emus: uae, hercules (Z-arch)

summarizing the reasons for virtualization

- 
- security: PS3, separation of daemons into VMs
 - consolidation, increased utilization: collecting many instances on one metal
 - provide rare hardware: macosx@kvm, amiga via uae, winnt@vmware
 - rapid provisioning: vm-copy
 - fault tolerance: fast vm reboots, vm-migration to different hardware

cloud overview



Customer0s Virtual machines



Customer1s Virtual machines



Customer2s Virtual machines

cloud topics/points of movement

- customers apps and their interaction has to be cloud-compatible, cloud is designed for horizontally scaling apps
- standards not yet there for secure data-exchange with the cloud
- cloud-provider trustworthy enough for customers data?
- protection of VMs against each other becomes more important
- upcoming arena: frontends to manage clouds/VMs, libvirt is nice here

also in a not too far future...

- virtualization on embedded/mobile for security
- linux process freezing/unfreezing
- vms get used like appliances (use while still downloading, boot from internet)
- live-migrations advancing, proxmox already does
kvm-migrations over ssh
- getting a common layer spanning all x86-virtualizers? So vms are movable/migratable

start
oo

big iron
oooooo

void
o

bochs
o

vmware
oooooo

qemu
oo

xen vb
oooo oo

kvm
oo

others
o

conclusion
o

future
ooo

end
●



Hoping you found the next toy to play with?

<http://fluxcoil.net>

chorn@fluxcoil.net