PRESENTED BY RED HAT

LEARN. NETWORK.
EXPERIENCE OPEN SOURCE.

www.theredhatsummit.com

KVM in Red Hat Enterprise Linux 6

Bill Burns
Engineering Manager, Red Hat Inc
June 23, 2010

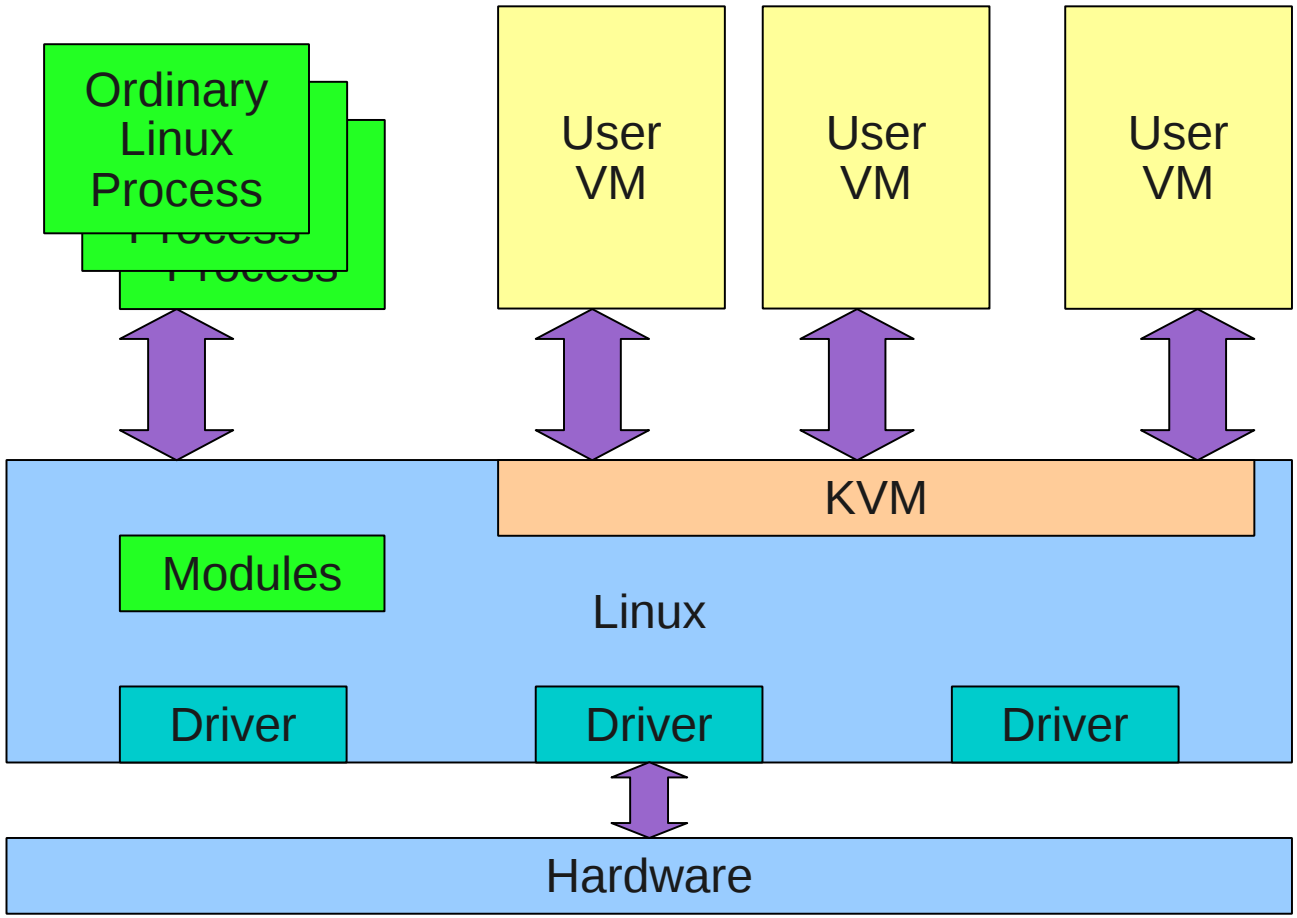
Outline

- KVM
- Architectures / History
- Hardware features
- Enhancements
- RAS
- Supported guests
- Future
- Related talks

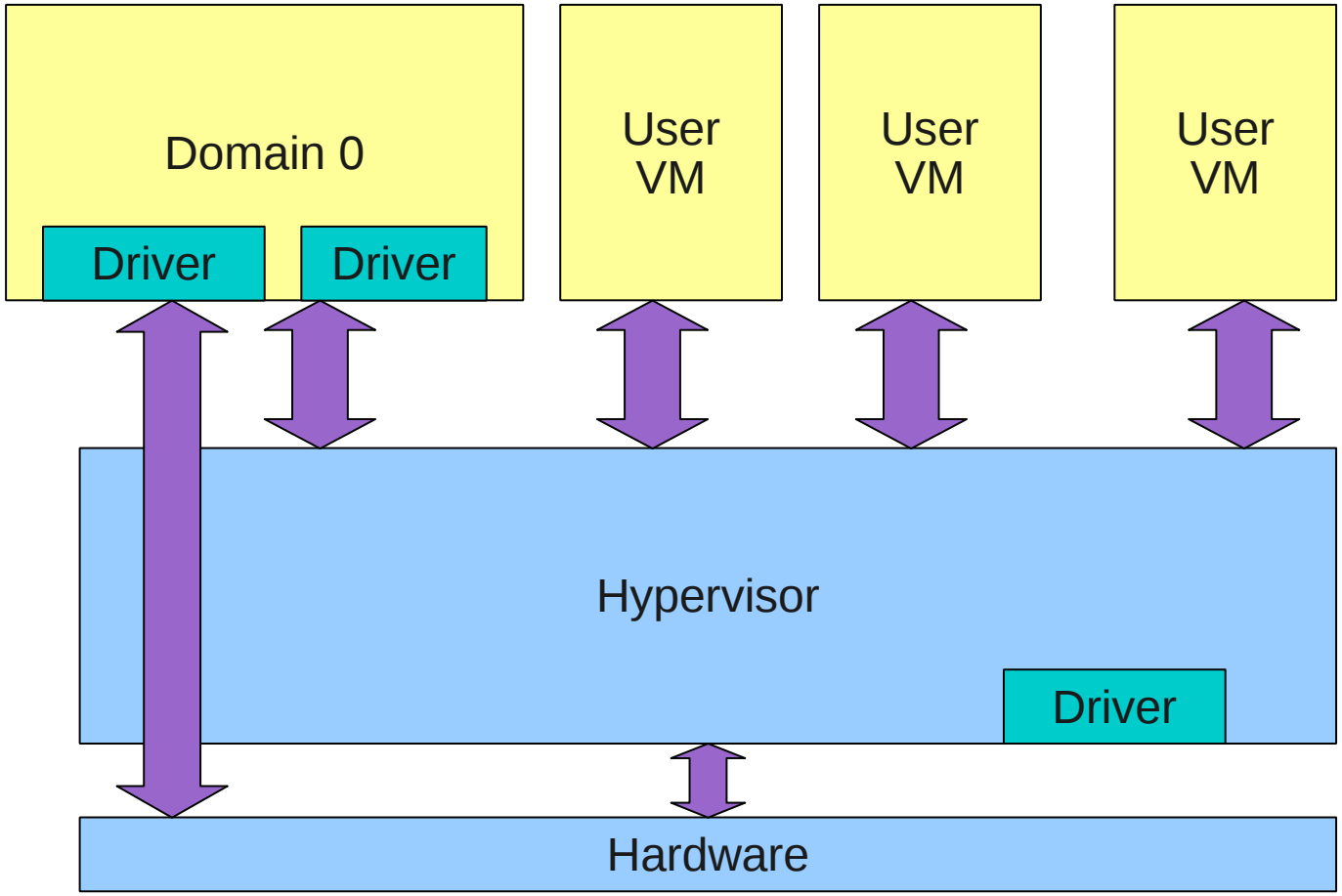
KVM

- Kernel-based virtual machine
- Linux Kernel modules turn Linux into a hypervisor
  - In Linux since 2.6.20
  - User space component is qemu-kvm
- Leverages Linux kernel features and support
- Leverages hardware virtualization support

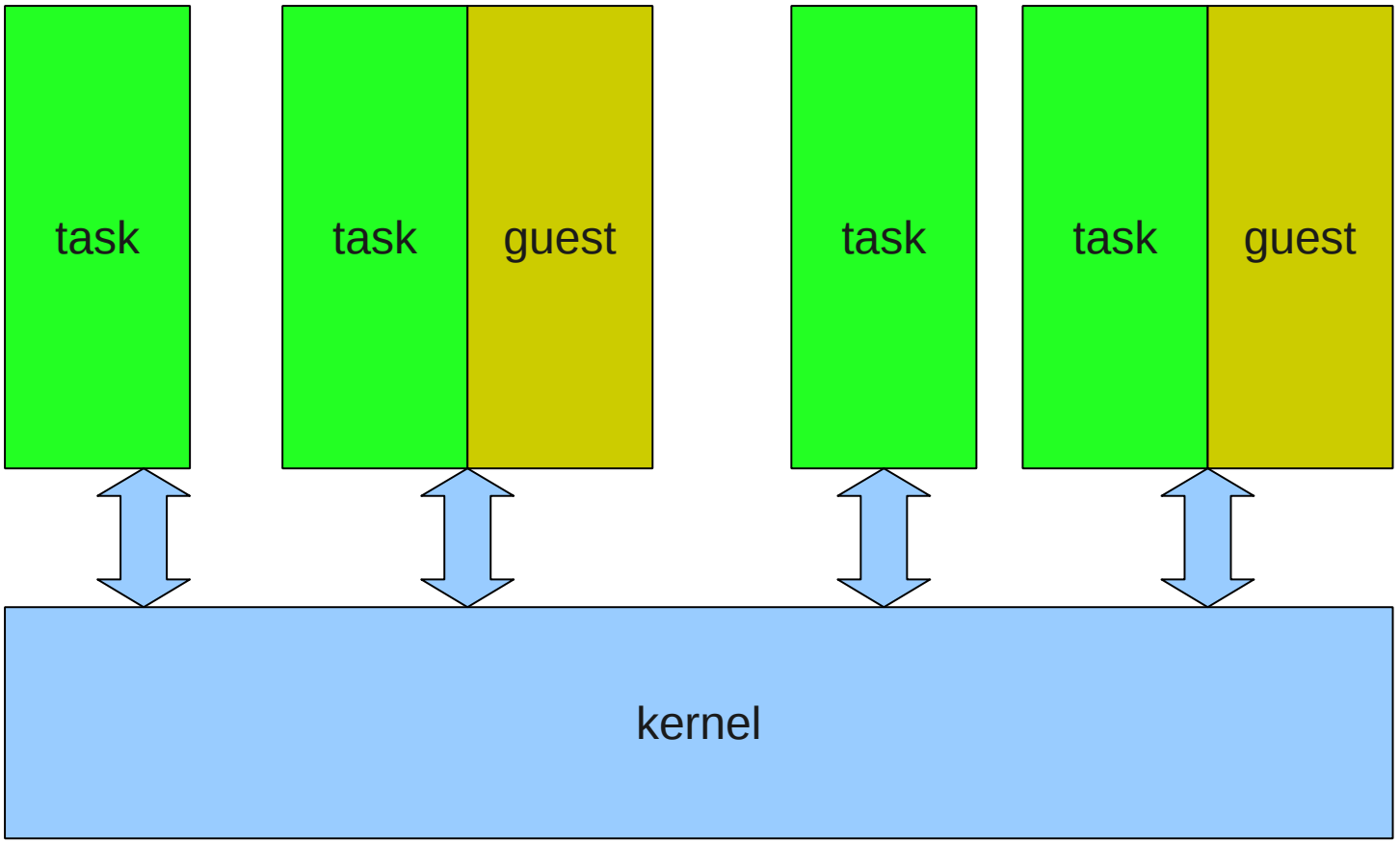
General Architecture (KVM)



General Architecture (Xen)



General Architecture (process model)



General Architecure (process model cont.)

- Guests are ordinary processes
  - Each virtual cpu is thread
- Like a new operating mode: kernel, user and “guest”
- “Guest” mode can hypercall, but not syscall
- Leverages Linux kernel features like
  - Scheduling, Accounting, cgroups
  - KSM (Kernel Samepage Merging)
  - Power management

Brief virtualization history

- Mid 1960s IBM developed mainframe virtualization
- 2001 VmWare 1st x86 virtualization, binary translation
- 2003 Xen, para-vitualization
- 2005/6 Intel/AMD x86 hardware virtualization
- 2007 Linux 2.6.20 includes KVM
- 2009 Red Hat supports KVM in Red Hat Enterprise Linux 5.4
- 2010 RHEL 6

Hardware features

- CPU support (Intel VMX, AMD SVM)
  - EPT/NPT
- IOMMU/VT-d
  - Protection for devices that are passed through to guests
- SR-IOV
  - Safe sharing of real hardware
  - Getting real traction with NICs
- NPIV
  - Allows sharing storage

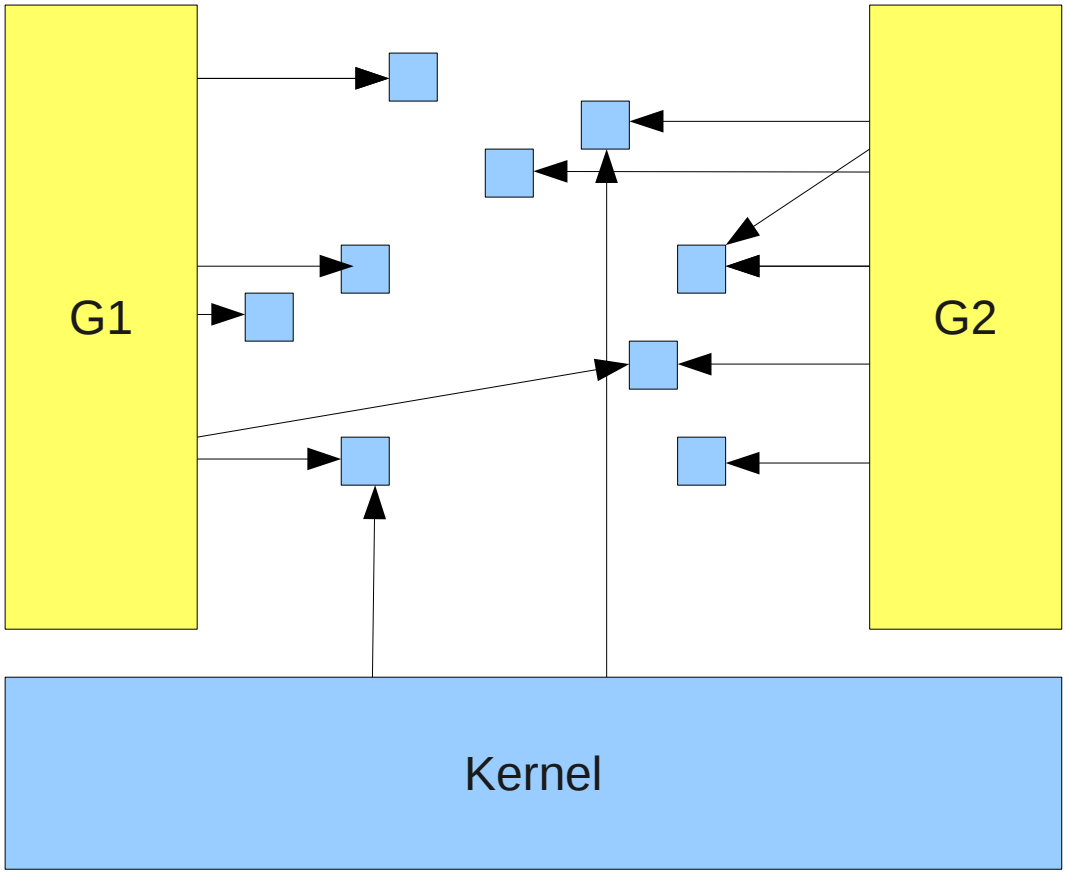
RHEL 6 CPU Enhancements

- 64 Virtual CPUs per guest
- Minimized CPU overhead
  - RCU kernel “locking” improves large SMP performance
  - User space notifiers
  - X2apic, a virtual interrupt controller

RHEL 6 Memory Enhancements

- Transparent Hugepages
  - Now dynamic, no boot time preallocation required
  - Can be broken down and swapped
- Extended/Nested page table aging improves swap choices
- Linux feature KSM (Kernel Samepage Merging) coalesces common pages.
  - A real win for Windows guests, where they zero all pages on boot

KSM example



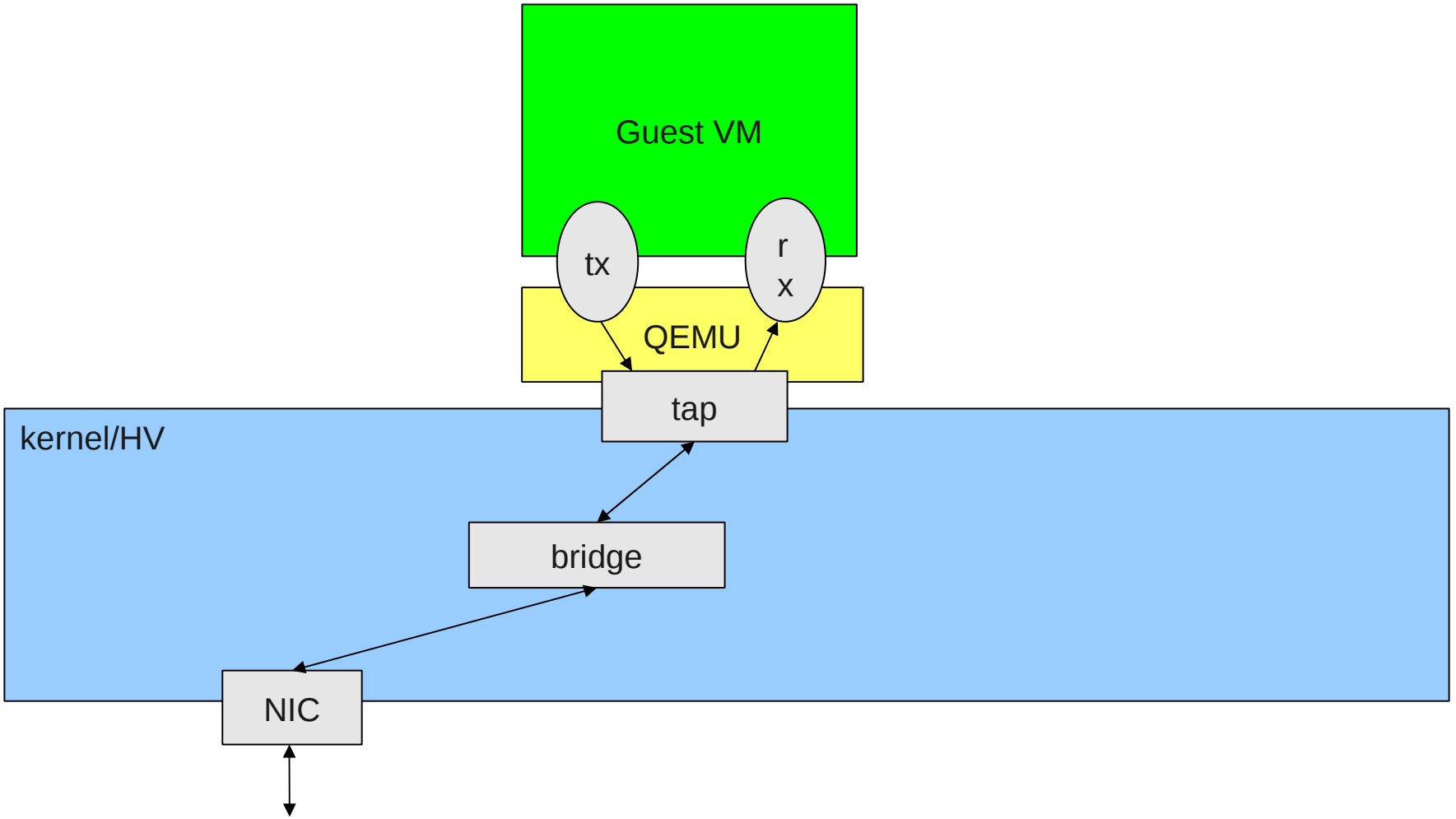
RHEL 6 Block I/O Enhancements

- Native AIO, and preadv/writev
- External ring buffers in guest/host interface
- Virtio barrier support
- MSI interrupt support
- Intelligent block alignment changes, better default
- Near native performance

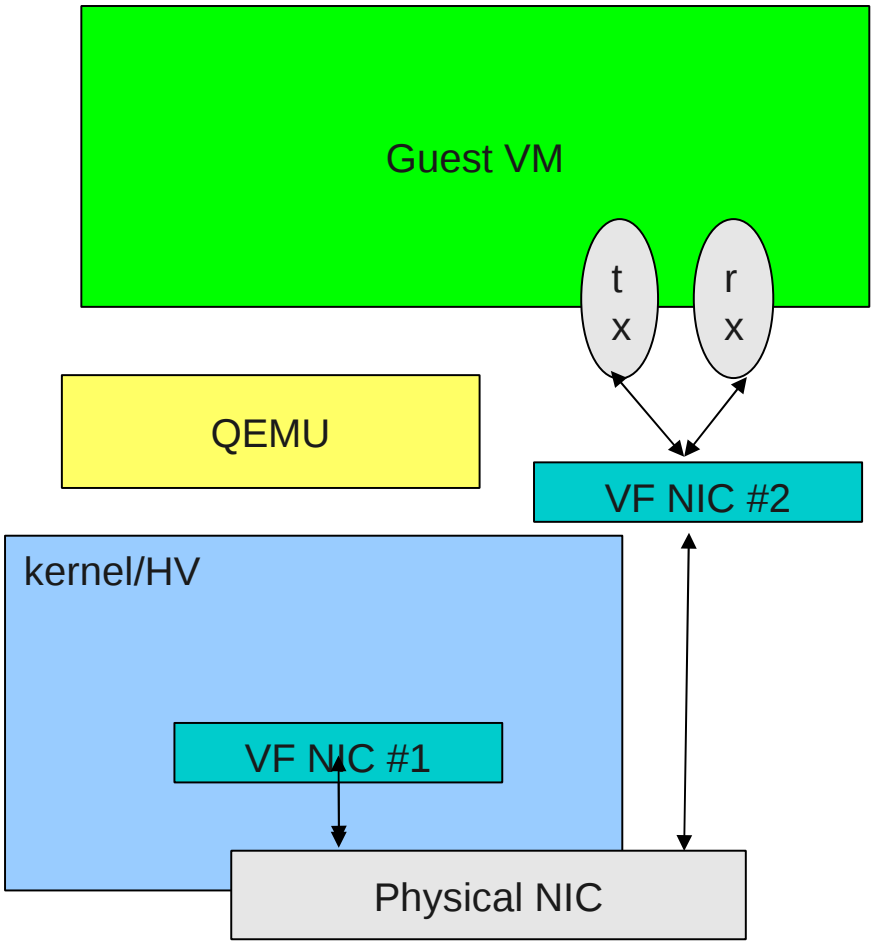
RHEL 6 Network I/O Enhancements

- Networking
  - Vhost-net, moves a portion of networking from user space to kernel.
    - More migratable than pass through
    - Can be used on top of SR-IOV devices, while preserving migratability
  - GPXE network boot supported.

Virtio drivers

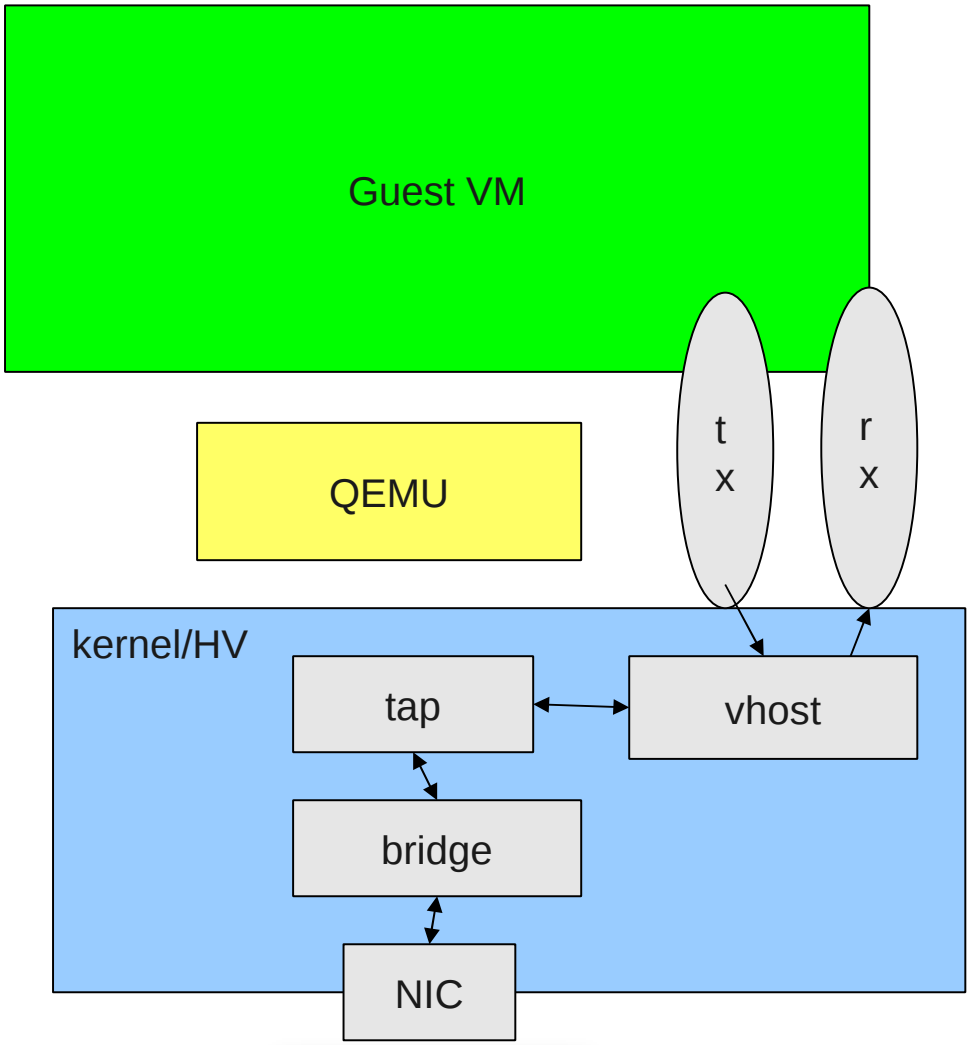


Device assignment – SR-IOV, VT-d/IOMMU



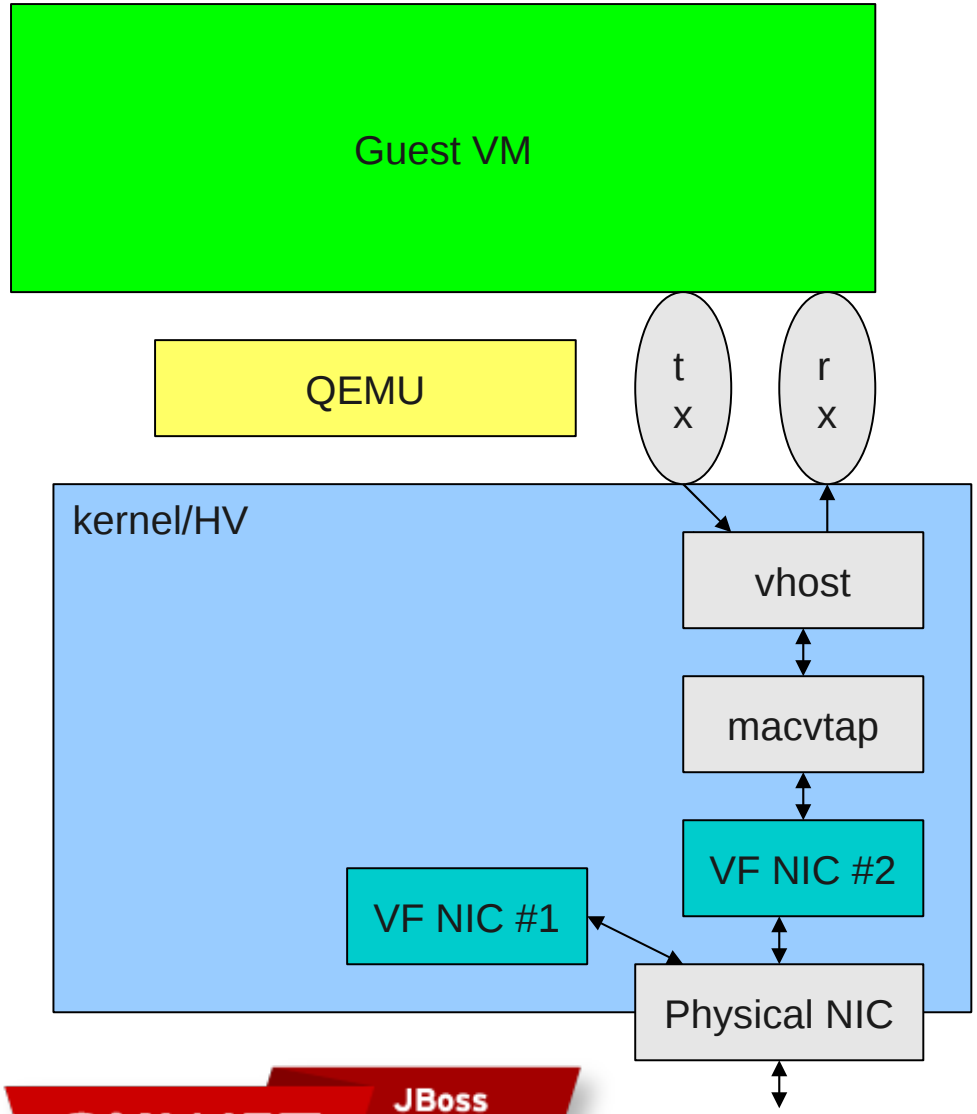
- Low overhead
  - Best throughput
  - Lowest latency
- Complicates migration

In-Kernel Vhost-net



- Less context switching
- Low latency
- MSI
- One less copy

Vhost over SR-IOV using macvtap



- Guest only knows virtio
- Migration friendly
- Excellent performance
- Future zero copy

RHEL 6 RAS Enhancements

- QMP – QEMU Monitor protocol
- Virtio serial (vmchannel)
- Improved migration protocol
- Kvm-clock
- Cgroups
- sVirt
- Power management – tickless kernel
- Static PCI slots to allow easier migration

Supported Guests

- RHEL 3/4/5/6
- Windows 2003, 2008, XP and 7

Future (beyond 6.0)

- PV spin locks
- Vhost zero copy
- Vswitch, VEPA
- Nested VMX
- UIO PCI device assignment
- Deep C-state power management

Related sessions

- Achieving Peak Performance from Red Hat KVM-Based Virtualization
  - Mark Wagner (2pm today)
- Kernel Virtualization Optimizations for KVM
  - Rik van Riel (4:20pm tomorrow)

FOLLOW US ON TWITTER

www.twitter.com/redhatsummit

TWEET ABOUT IT

#summitjbw

READ THE BLOG

http://summitblog.redhat.com/