What's New in Virtualization

Presented by
Amit Shah
Red Hat, Inc.

22 Nov 2014
CentOS Dojo Pune

About Me

- Senior Software Engineer at Red Hat
- Working on KVM since 2007
- Working on Linux since 2001

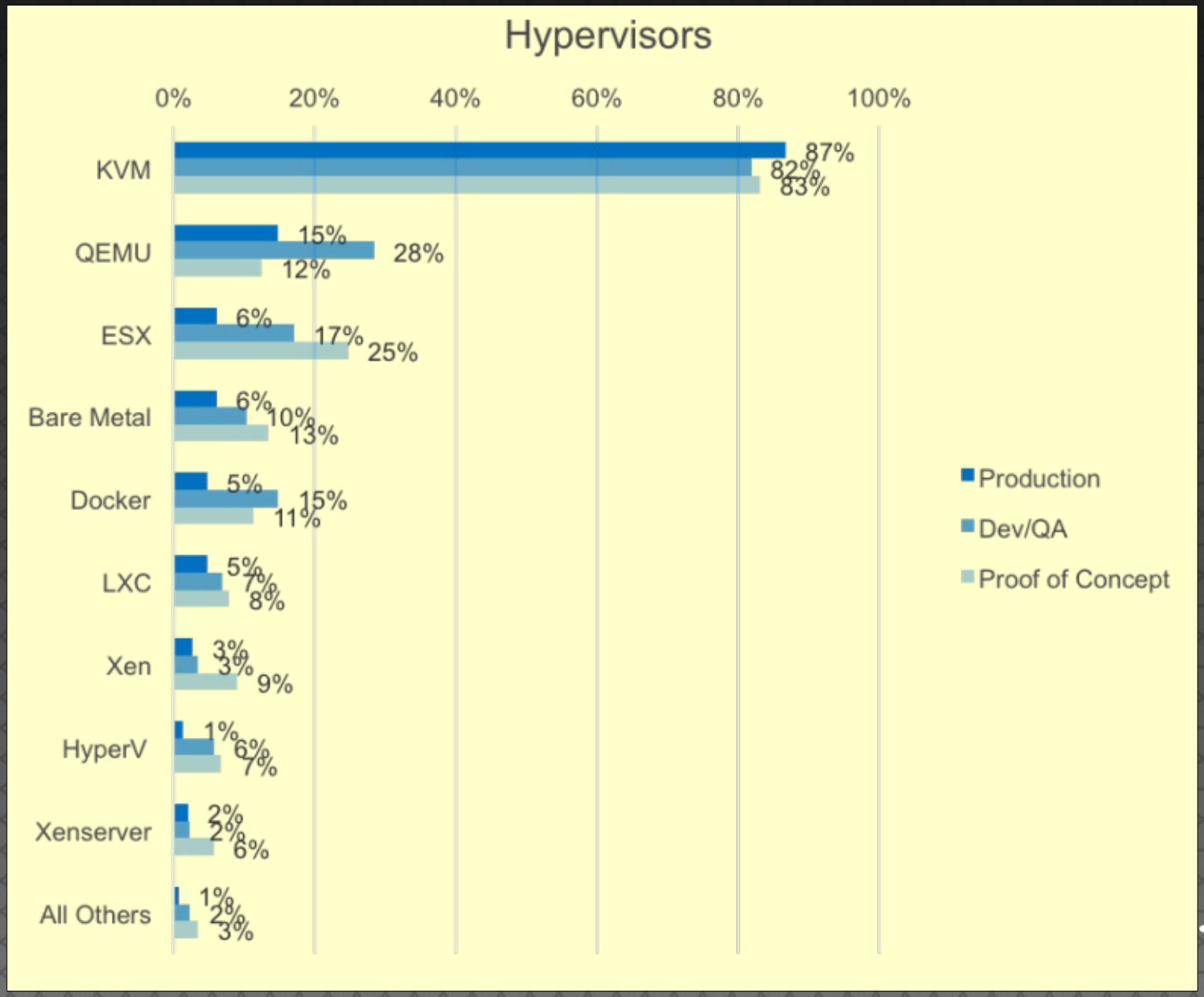
fedora

About KVM

- Default hypervisor since EL6
  - Also available as an option since EL5.4
- Default hypervisor for OpenStack
- Consistent top-performer in SPECVirt

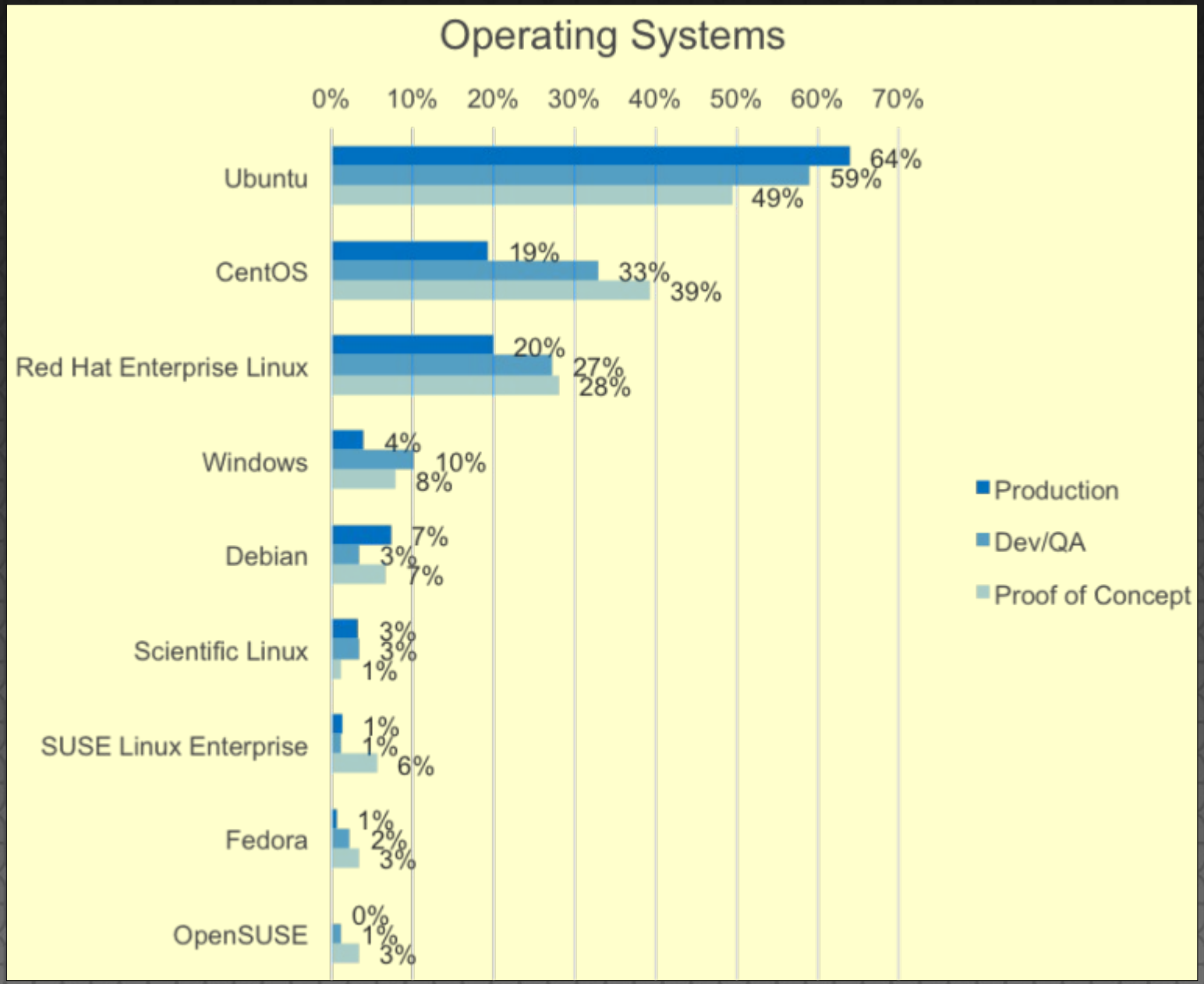
How Is This Relevant?

Nov 2014 OpenStack Survey



fedora

Nov 2014 OpenStack Survey



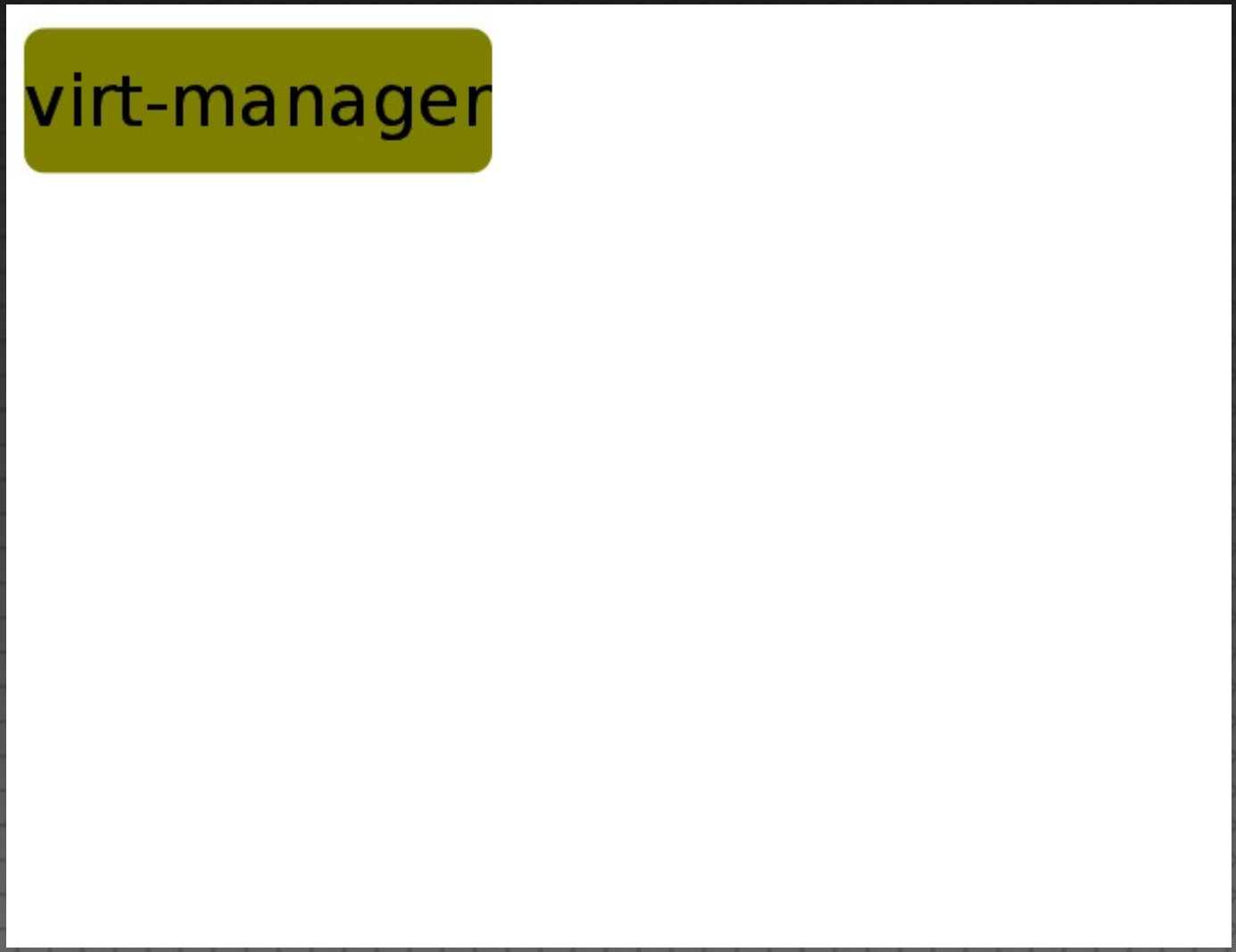
fedora

Overview of the Virt Stack

Stack

- No Swiss-Army knife
- Do one thing, do it right

Stack

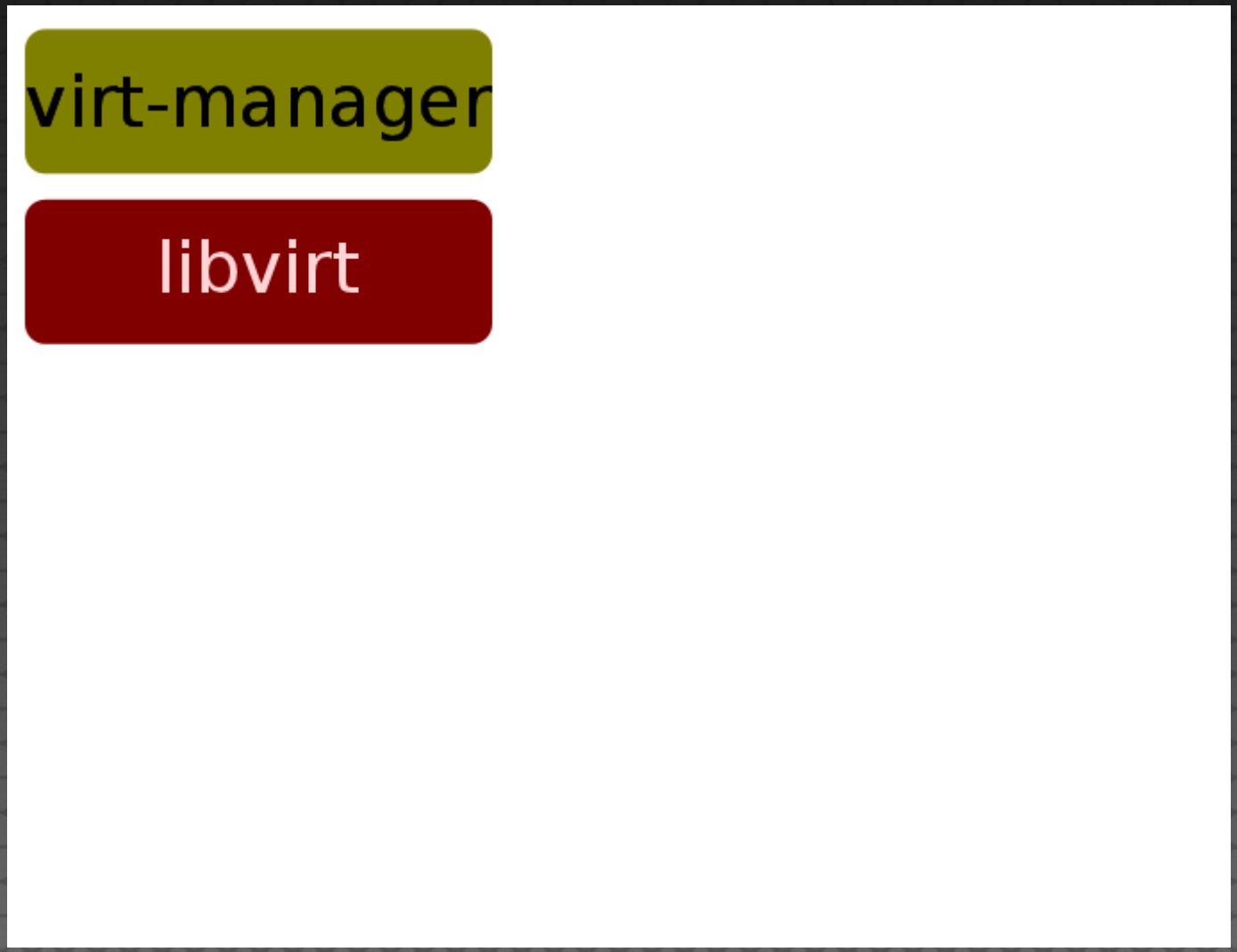


fedora

virt-manager

- User-visible interface
- Install and run guests
- Select optimal settings according to guest type
- Connects to libvirt
  - Local and remote

Stack

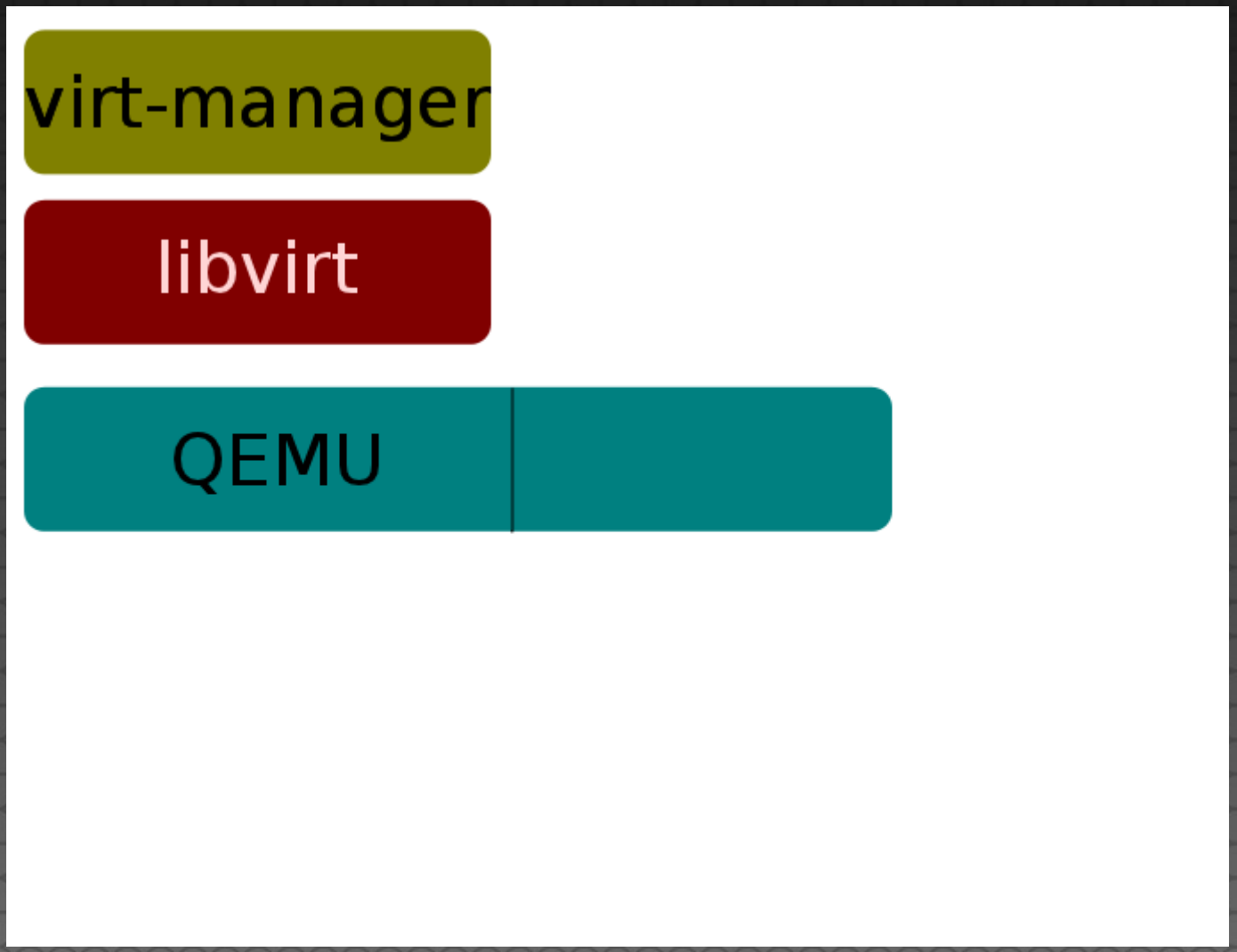


fedora

libvirt

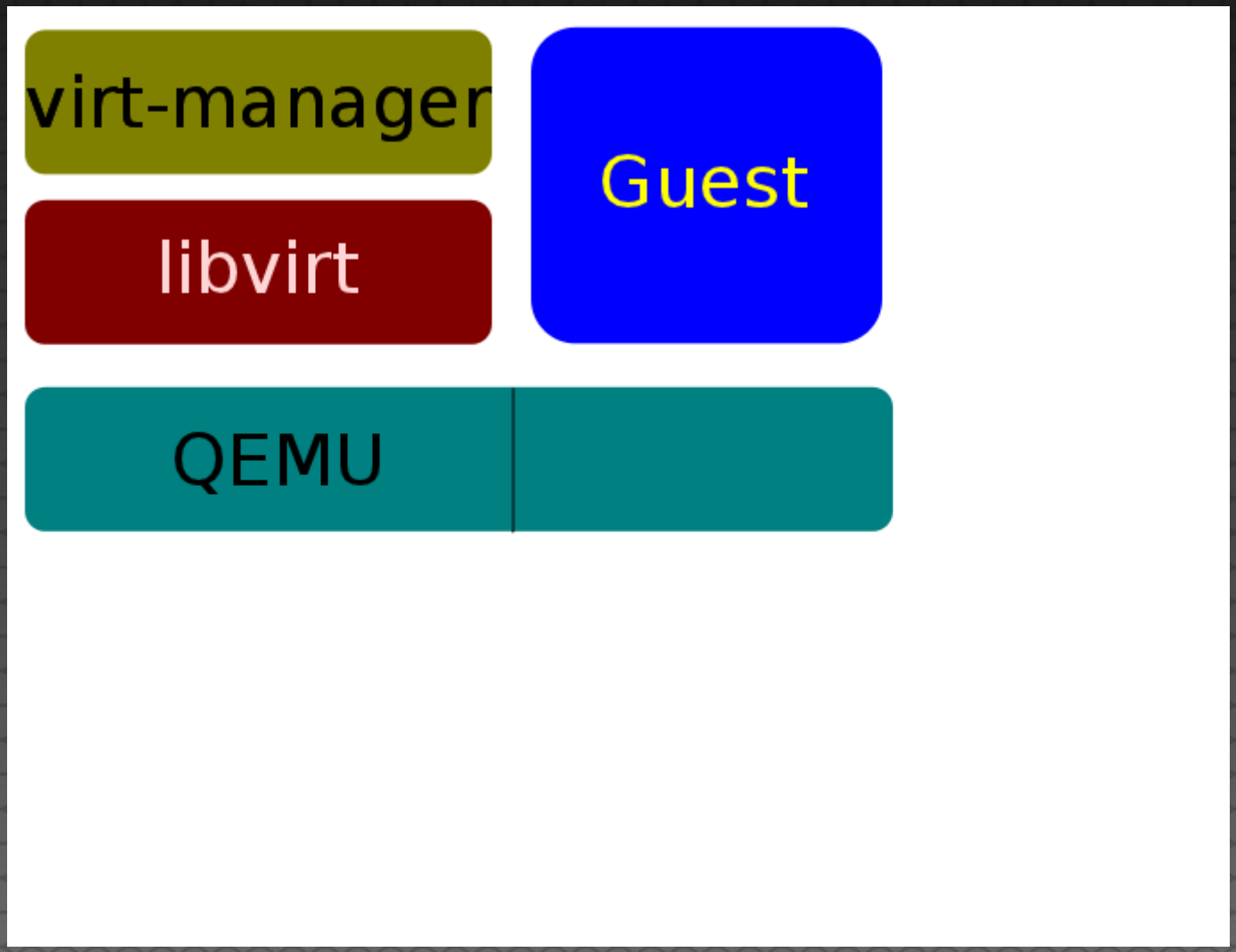
- Hypervisor-neutral API
- Mechanisms, not policies
- Handles various storage types
- Handles networking
- Talks to QEMU

Stack



fedora

Stack

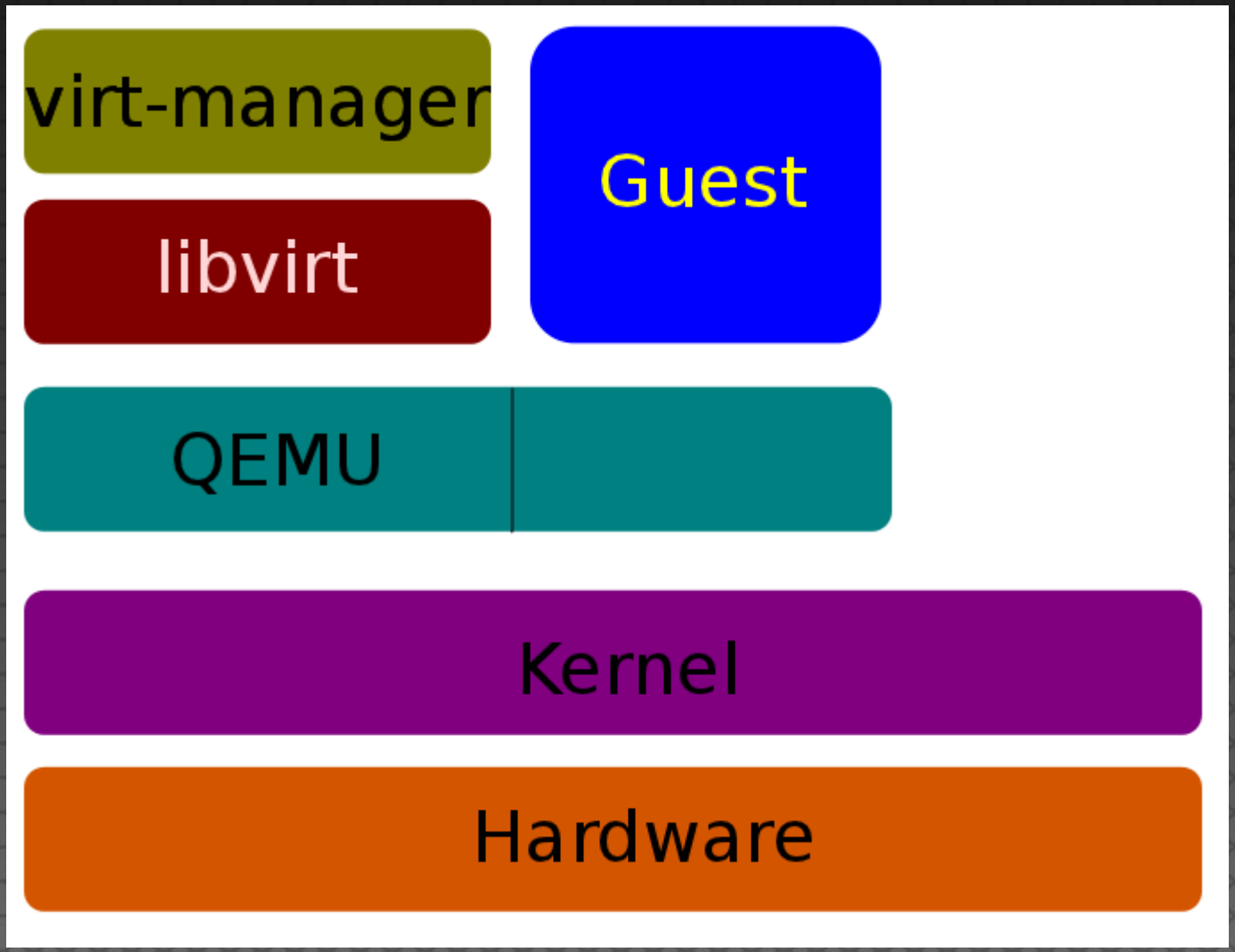


fedora

QEMU

- Hardware emulation
  - CPU, Devices
- Management interface to guests
- Interfaces with the kernel

Stack

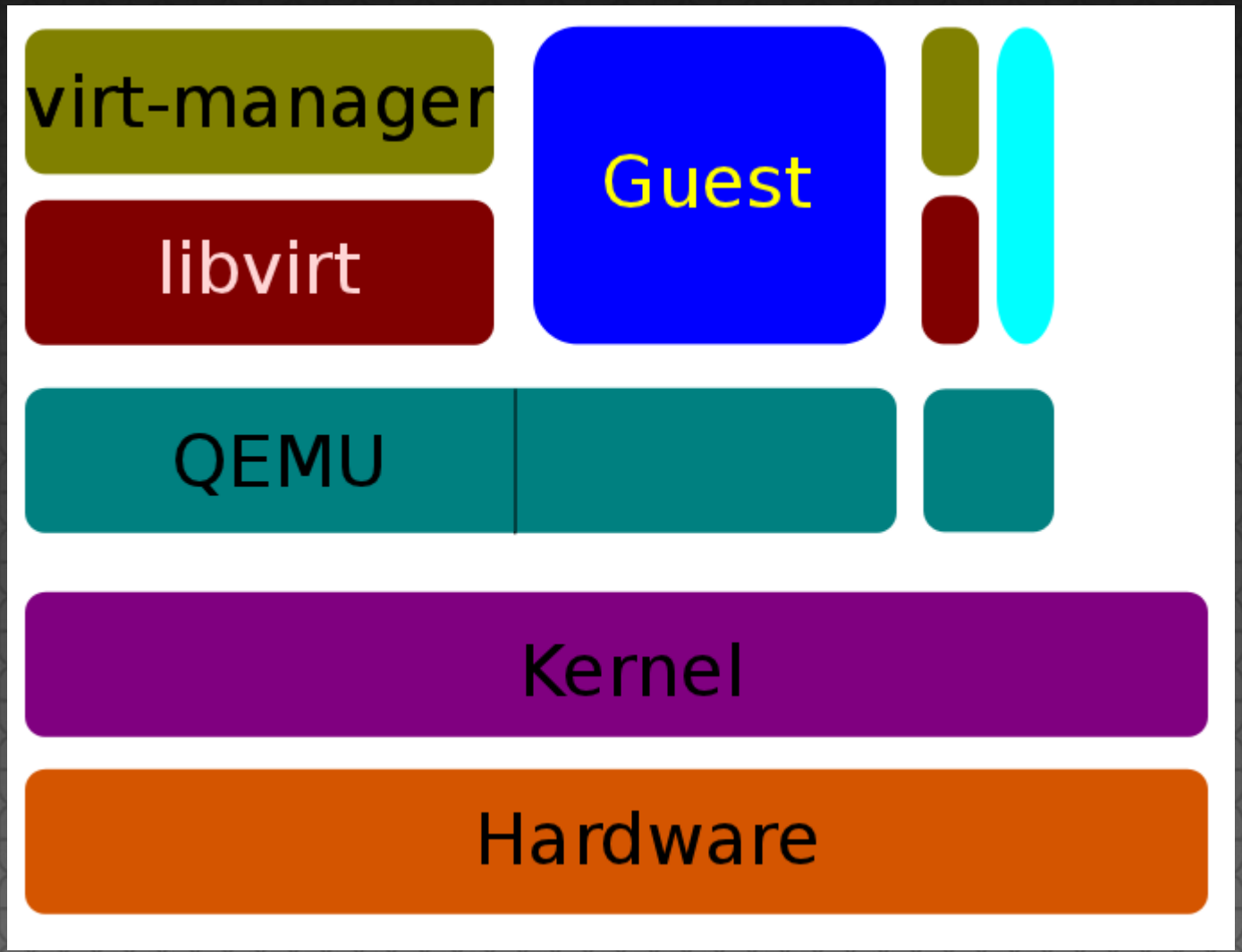


fedora

Linux Kernel

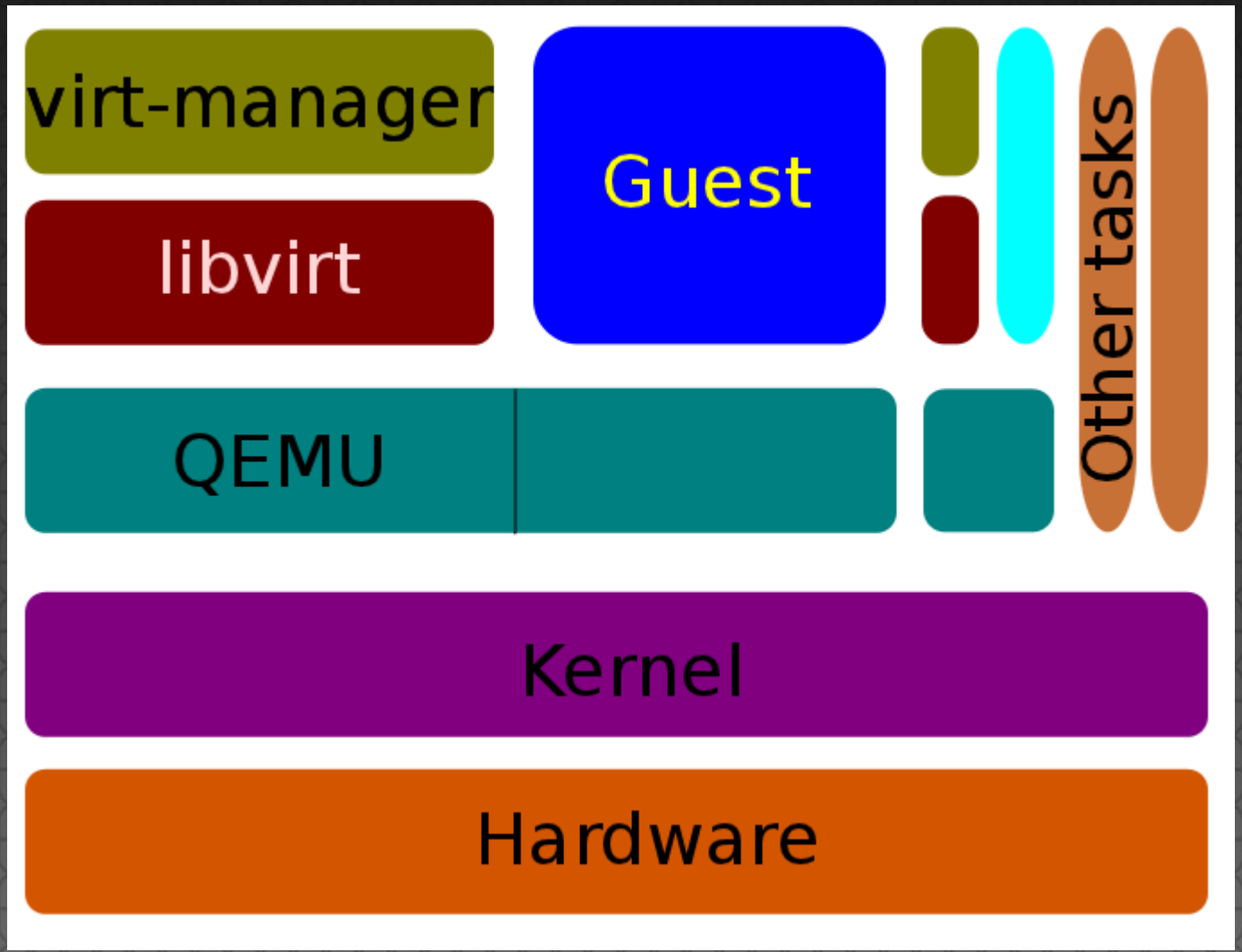
- Houses KVM
- Exploits hardware virtualization features
- Allows QEMU guests to run on local CPUs
- Paravirtualized drivers

Stack



fedora

Stack



fedora

Host Tools

- `guestfish`
- `libguestfs`
- `virt-top`
- `virt-df`
- `virt-ls`
- `virt-cat`
- `virt-*`
- diff between guests
- add users
- change passwords
- security audits
- resize disk images
- create, clone templates
- edit Windows registry
- `guestfs-recipes(1)`

Client tools

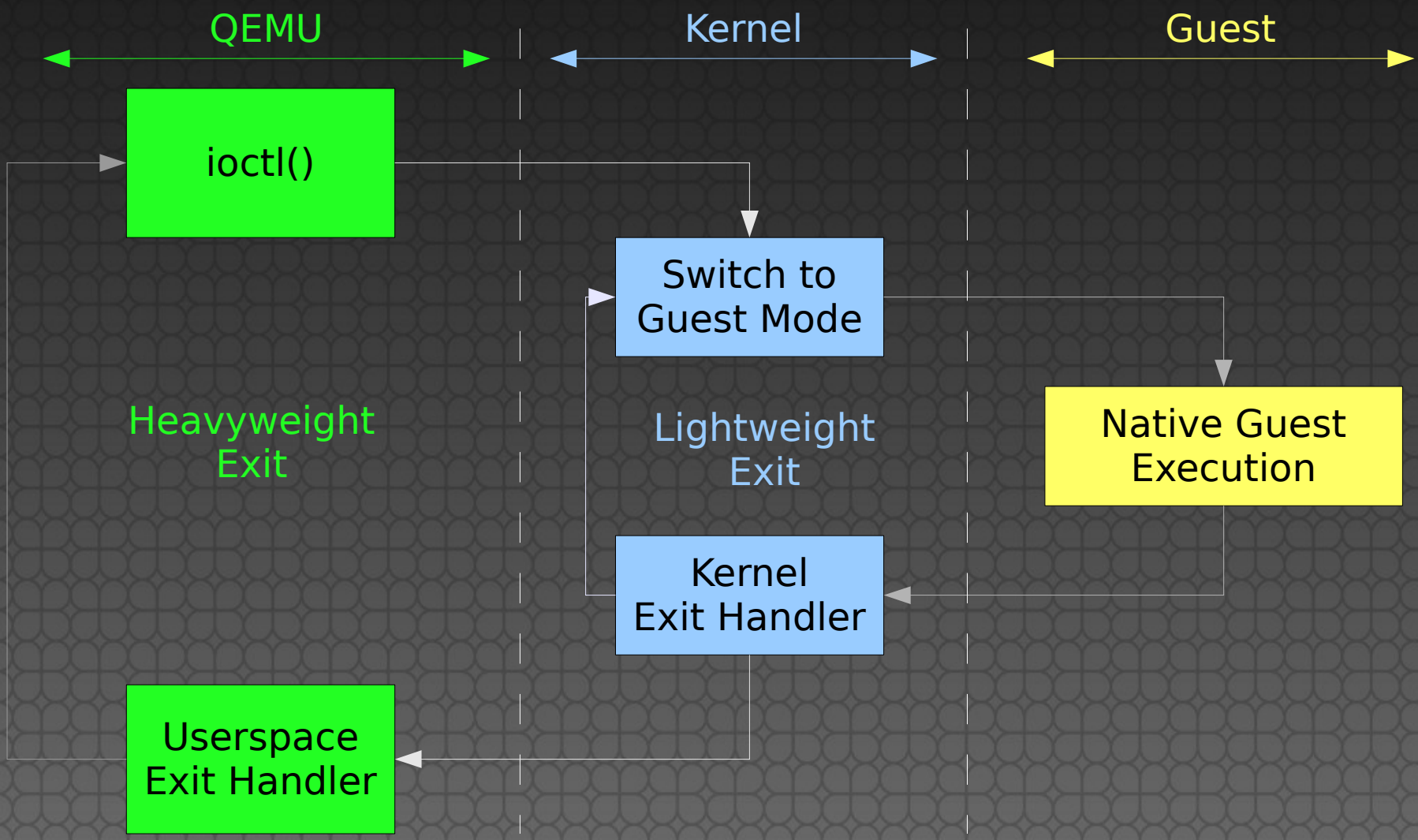
- VNC viewer
- SPICE client

Guest Agents

- Spice Video
- Spice Audio
- Clipboard copy/paste
- Lock screen
- Reboot / Shutdown
- USB devices

Behind the Scenes

KVM Execution Model



What's New

Performance

- EOI acceleration
- Integrated ballooning and THP
- HyperV Enlightenments
- Paravirtualized ticketlocks
- kvmclock vsyscall improvements
- NUMA-aware scheduling

PCI Improvements

- VFIO
  - New way of doing device assignment
- NVIDIA GPU device assignment
- PCI Express
- PCI Bridges
  - No more 32 device limit
- AER for assigned devices

Block I/O

- virtio-blk dataplane
- Multi-queue virtio-scsi
- IO throttling support
- discard support
- qcow2v3 support
- Live block copy / storage migration
- Live snapshot

Networking

- Multi-queue virtio-net
- Bridge zero-copy transmit

Security

- Syscall sandboxing
- virtio-rng

Reliability, Availability, Scalability

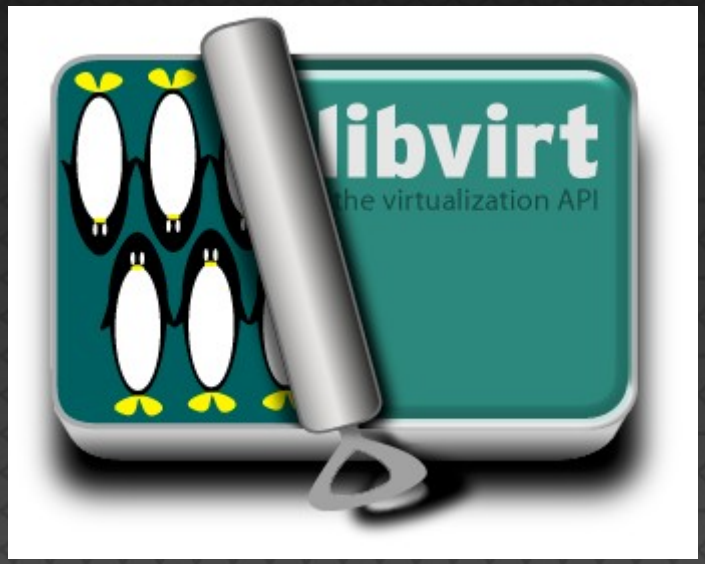
- vCPU hot-add
- Page compression for live migration
- Live migration from EL6.5 to 7.0
- Improved live migration statistics
- Live migration thread
- pvpanic

Resources

- http://opensource.com/business/14/11/lessons-openstack-user-survey
- http://linux-kvm.org
- http://qemu-project.org
- http://libvirt.org
- http://virt-tools.org
- http://libguestfs.org

fedora

Questions?



Contact:
amit.shah@redhat.com
http://log.amitshah.net