Virtual Networking in KVM

Arnd Bergmann
KVM Forum 2010
802.1Qbg and 802.1Qbh
Performance
Open vSwitch
802.1Qbg VEPA

Hypervisor

VM 1  VM 2  VM 3  VM 4  VM 5

Virtual Ethernet Port Aggregator

External Switch
802.1Qbh Port Extender

Hypervisor

VM 1  VM 2  VM 3  VM 4  VM 5

S-VLAN Port Extender Slave

S-VLAN Port Extender Master

External switch
802.1Qbg Multichannel VEPA

Hypervisor

VM 1
VM 2
VM 3
VM 4
VM 5

Virtual Bridge
Virtual Ethernet Port Aggregator
Multichannel Aggregator

External Switch
802.1Qbg protocols

- LLDP: Link-layer discovery protocol, new TLVs
  - EVB: Edge Virtual Bridging
  - CDCP: Channel Discovery and Configuration Protocol
- VDP: VSI discovery protocol

Virtual Station Interface Discovery Protocol

VSI State Machine – Station
(One Instance per VSI)
Network stack alternatives

- **Classic**
  - NIC Hardware
  - Device driver
  - Host kernel (bridge)
  - Host kernel (tap)
  - Host User (qemu)
  - Host kernel (KVM)
  - Guest kernel
  - Guest user space

- **Current**
  - NIC Hardware
  - Device driver
  - macvtap
  - vhost-net
  - Host kernel (KVM)
  - Guest kernel
  - Guest user space

- **Passthrough**
  - NIC Hardware
  - Device driver
  - IOMMU
  - macvtap
  - vhost-net
  - Host kernel (KVM)
  - Guest kernel
  - Guest user space
  - missing features...

- **My Goal**
  - NIC Hardware
  - Device driver
  - IOMMU
  - macvtap
  - vhost-net
  - Host kernel (KVM)
  - Guest kernel
  - Guest user space
Open vSwitch

http://www.openvswitch.org/
Open vSwitch

- OpenFlow standard based
- Alternative to bridge, macvlan/macvtap
- Use cases:
  - Distributed Layer 3 switch
  - Security Appliance
- → drivers/staging
This work represents the view of the author and does not necessarily represent the view of IBM.
IBM, IBM (logo), e-business (logo), pSeries, e (logo) server, and xSeries are trademarks or registered trademarks of International Business Machines Corporation in the United States and/or other countries.
Linux is a registered trademark of Linus Torvalds.
Other company, product, and service names may be trademarks or service marks of others.