IaaS: 开源组件入门介绍

Star Guo

guohuaxing@96900.com.cn

什么是IaaS

- IaaS, Infrastructure as a Service(基础设施即服务).
- 资源弹性：计算、网络、存储和带宽；
- 快速部署：模板化部署和应用商店；
- 数据持久：分布式存储多副本与异地容灾；
- 自动监控：友好的面板和及时通知；
- 网络安全：能抵御一定限度攻击；
- 付费模式：便捷支付手段与合理计费；

IaaS开源方案

Others …

常用的IaaS组件项目

- 操作系统: Ubuntu, Debian, CentOS …
- 虚拟层: kvm, xen, openvz …
- 虚拟管理软件: Libvirt …
- 虚拟交换机： OpenvSwitch …
- 存储系统： Btrfs, ZFS, Gluster, Ceph …
- 监控系统： Nagios, Zabbix, Cacti …

KVM虚拟机

```
[root@rhel ~]# rpm -qi qemu-kvm
Name        : qemu-kvm                     Relocations: (not relocatable)
Version     : 0.12.1.2                         Vendor: Red Hat, Inc.
Release     : 2.415.el6                     Build Date: Fri 25 Oct 2013 10:26:38 PM CST
Install Date: Fri 29 Nov 2013 05:51:38 PM CST      Build Host: x86-001.build.bos.redhat.com
Group       : Development/Tools             Source RPM: qemu-kvm-0.12.1.2-2.415.el6.src.rpm
Size        : 4541632                          License: GPLv2+ and LGPLv2+ and BSD
Signature   : RSA/8, Tue 29 Oct 2013 04:06:15 PM CST, Key ID 199e2f91fd431d51
Packager    : Red Hat, Inc. <http://bugzilla.redhat.com/bugzilla>
URL         : http://www.linux-kvm.org
Summary     : Userspace component of KVM
Description :
KVM (for Kernel-based Virtual Machine) is a full virtualization solution
for Linux on x86 hardware.

Using KVM, one can run multiple virtual machines running unmodified Linux
or Windows images. Each virtual machine has private virtualized hardware:
a network card, disk, graphics adapter, etc.
```

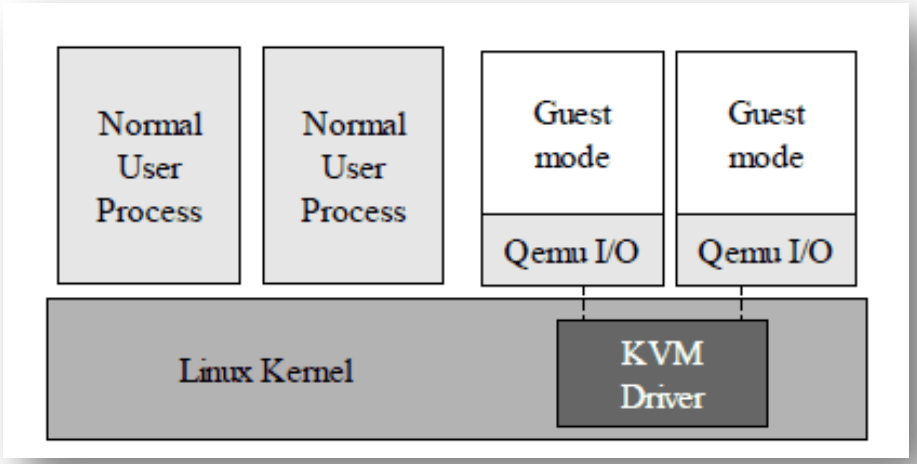
KVM虚拟机

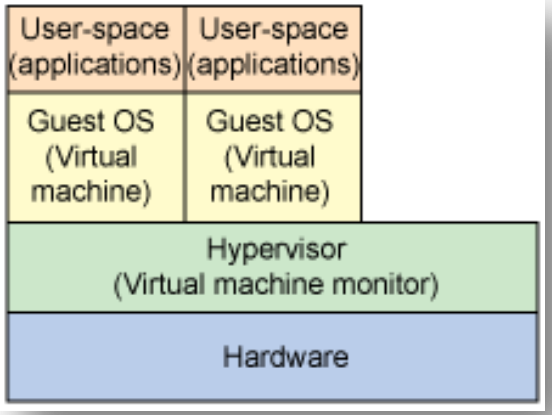
```
[root@rhel ~]# modinfo kvm
filename:       /lib/modules/2.6.32-431.el6.x86_64/kernel/arch/x86/kvm/kvm.ko
license:        GPL
author:         Qumranet
srcversion:     CC8A9FE27345BE5E7968E69
depends:
vermagic:       2.6.32-431.el6.x86_64 SMP mod_unload modversions
parm:           min_timer_period_us:uint
parm:           oos_shadow:bool
parm:           ignore_msrs:bool
parm:           allow_unsafe_assigned_interrupts:Enable device assignment on platforms
 (bool)
[root@rhel ~]# modinfo kvm_intel
filename:       /lib/modules/2.6.32-431.el6.x86_64/kernel/arch/x86/kvm/kvm-intel.ko
license:        GPL
author:         Qumranet
srcversion:     B81B8A09ACE1F572C590EF5
depends:        kvm
vermagic:       2.6.32-431.el6.x86_64 SMP mod_unload modversions
parm:           bypass_guest_pf:bool
parm:           vpid:bool
parm:           flexpriority:bool
parm:           ept:bool
parm:           unrestricted_guest:bool
parm:           eptad:bool
parm:           emulate_invalid_guest_state:bool
parm:           yield_on_hlt:bool
parm:           vmm_exclusive:bool
parm:           ple_gap:int
parm:           ple_window:int
```

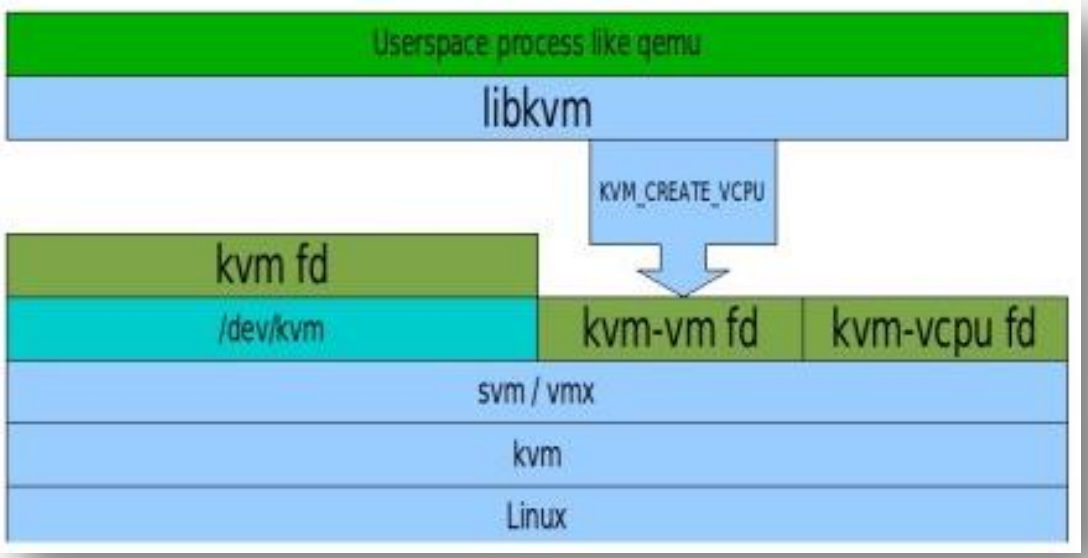
KVM虚拟机

```
oot     27764     1  0 Dec01 ?        00:10:50 /usr/libexec/qemu-kvm -name r
61-VM -S -M rhel6.4.0 -cpu core2duo,+lahf_lm,+dca,+pdcm,+xtpr,+cx16,+tm2,+vmx,
+ds_cpl,+dtes64,+pbe,+tm,+ht,+ss,+acpi,+ds -enable-kvm -m 128 -smp 1,sockets=1
,cores=1,threads=1 -uuid 566b1efd-abaf-4f22-a930-e8bbd6568a18 -nodefconfig -no
defaults -chardev socket,id=charmonitor,path=/var/lib/libvirt/qemu/r-61-VM.mon
itor,server,nowait -mon chardev=charmonitor,id=monitor,mode=control -rtc base=
utc -no-shutdown -device piix3-usb-uhci,id=usb,bus=pci.0,addr=0x1.0x2 -device
virtio-serial-pci,id=virtio-serial0,bus=pci.0,addr=0x4 -drive file=/gfs/324b60
41-6d43-4426-ae36-690c40c9cd0f,if=none,id=drive-virtio-disk0,format=qcow2,cach
e=writeback -device virtio-blk-pci,scsi=off,bus=pci.0,addr=0x5,drive=drive-vir
tio-disk0,id=virtio-disk0,bootindex=2 -drive file=/usr/share/cloudstack-common
/vms/systemvm.iso,if=none,media=cdrom,id=drive-ide0-1-0,readonly=on,format=raw
,cache=writeback -device ide-drive,bus=ide.1,unit=0,drive=drive-ide0-1-0,id=id
e0-1-0,bootindex=1 -netdev tap,fd=24,id=hostnet0,vhost=on,vhostfd=33 -device v
irtio-net-pci,netdev=hostnet0,id=net0,mac=0e:00:a9:fe:00:2b,bus=pci.0,addr=0x3
 -chardev pty,id=charserial0 -device isa-serial,chardev=charserial0,id=serial0
 -chardev socket,id=charchannel0,path=/var/lib/libvirt/qemu/r-61-VM.agent,serv
er,nowait -device virtserialport,bus=virtio-serial0.0,nr=1,chardev=charchannel
0,id=channel0,name=r-61-VM.vport -device usb-tablet,id=input0 -vnc 0.0.0.0:8,p
assword -vga cirrus -device virtio-balloon-pci,id=balloon0,bus=pci.0,addr=0x6
```

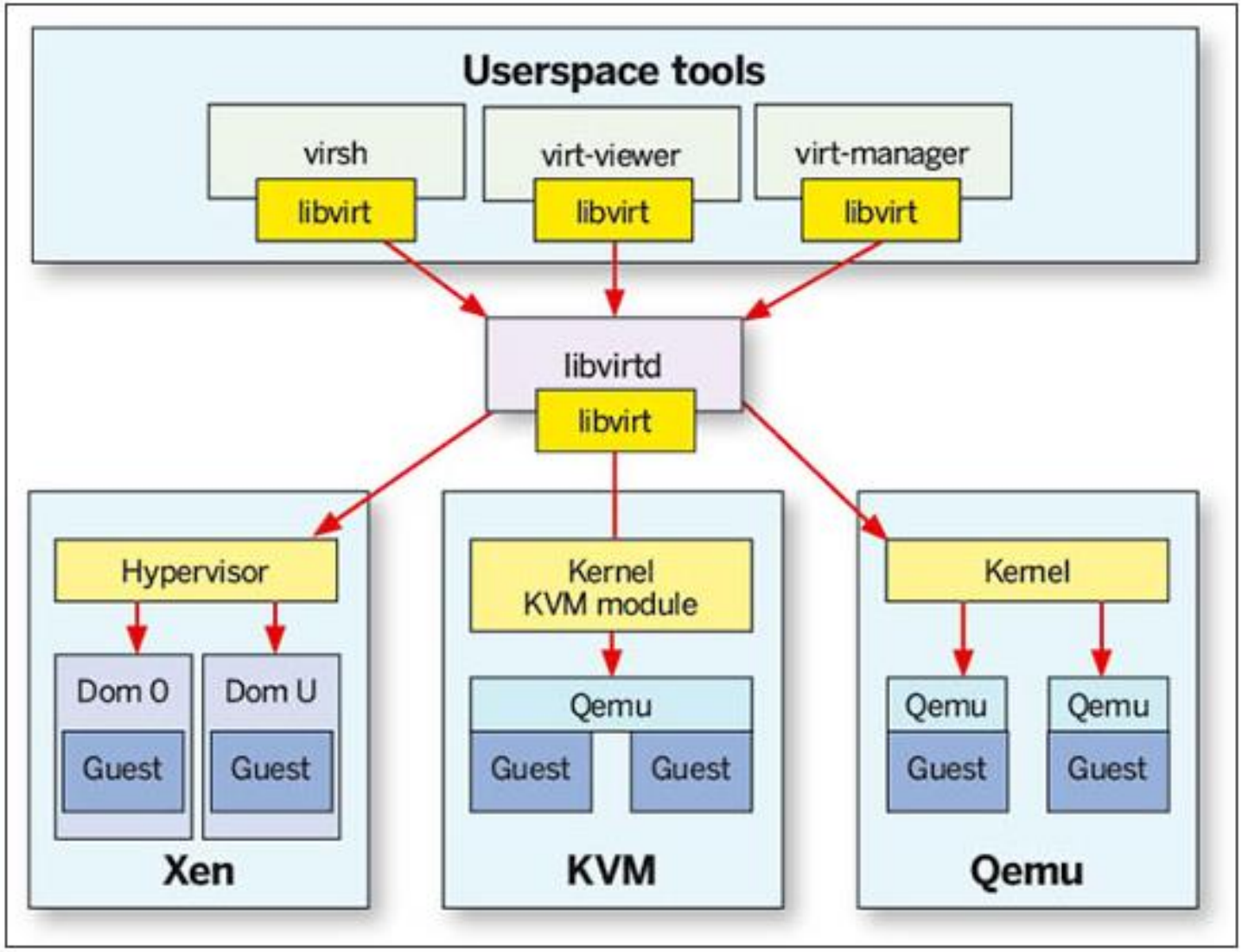
KVM虚拟机


Normal User Process
Normal User Process
Guest mode
Guest mode
Qemu I/O
Qemu I/O
Linux Kernel
KVM Driver

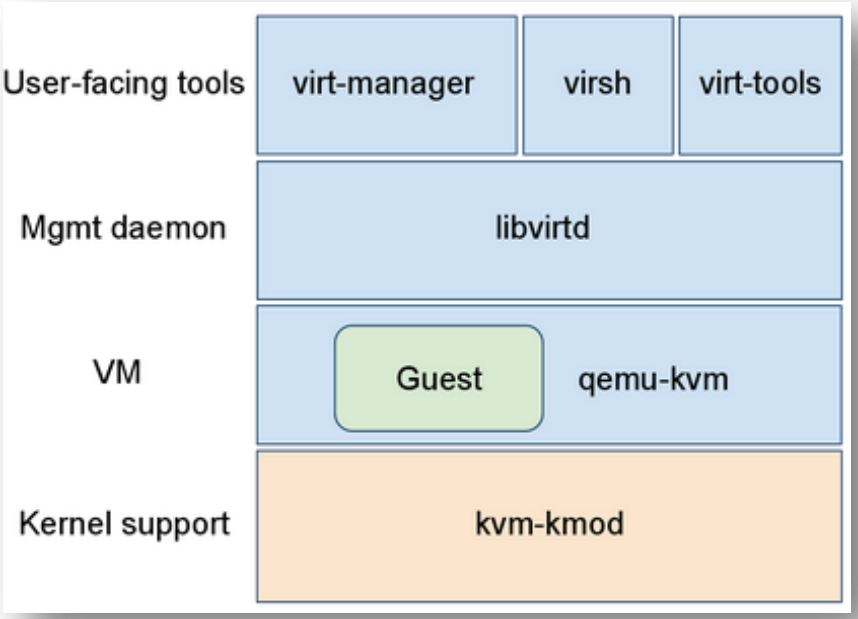

User-space (applications)
User-space (applications)
Guest OS (Virtual machine)
Guest OS (Virtual machine)
Hypervisor (Virtual machine monitor)
Hardware


Userspace process like qemu
libkvm
KVM_CREATE_VCPU
kvm fd
/dev/kvm
kvm-vm fd
kvm-vcpu fd
svm / vmx
kvm
Linux

Libvirt管理组件



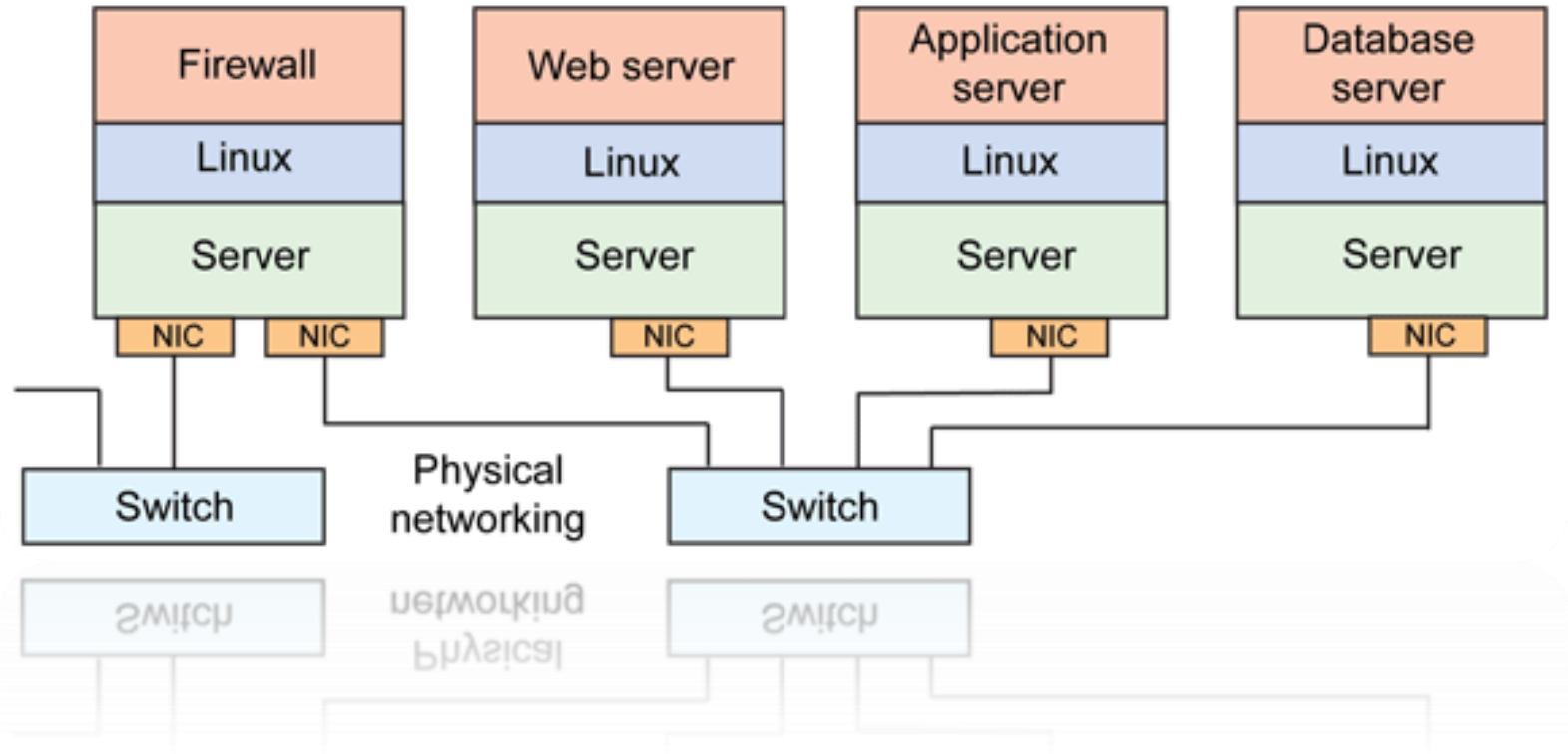
Libvirt管理组件




```
[root@kvm-02 ~]# virsh list --all
 Id    Name                           State
----------------------------------------------------
 1     v-27-VM                        running
 2     s-23-VM                        running
 4     r-4-VM                         running
 7     r-58-VM                        running
 8     i-6-29-VM                      running
 11    r-61-VM                        running
 12    r-63-VM                        running
 13    i-8-10-VM                      running
 18    i-11-70-VM                     running
```

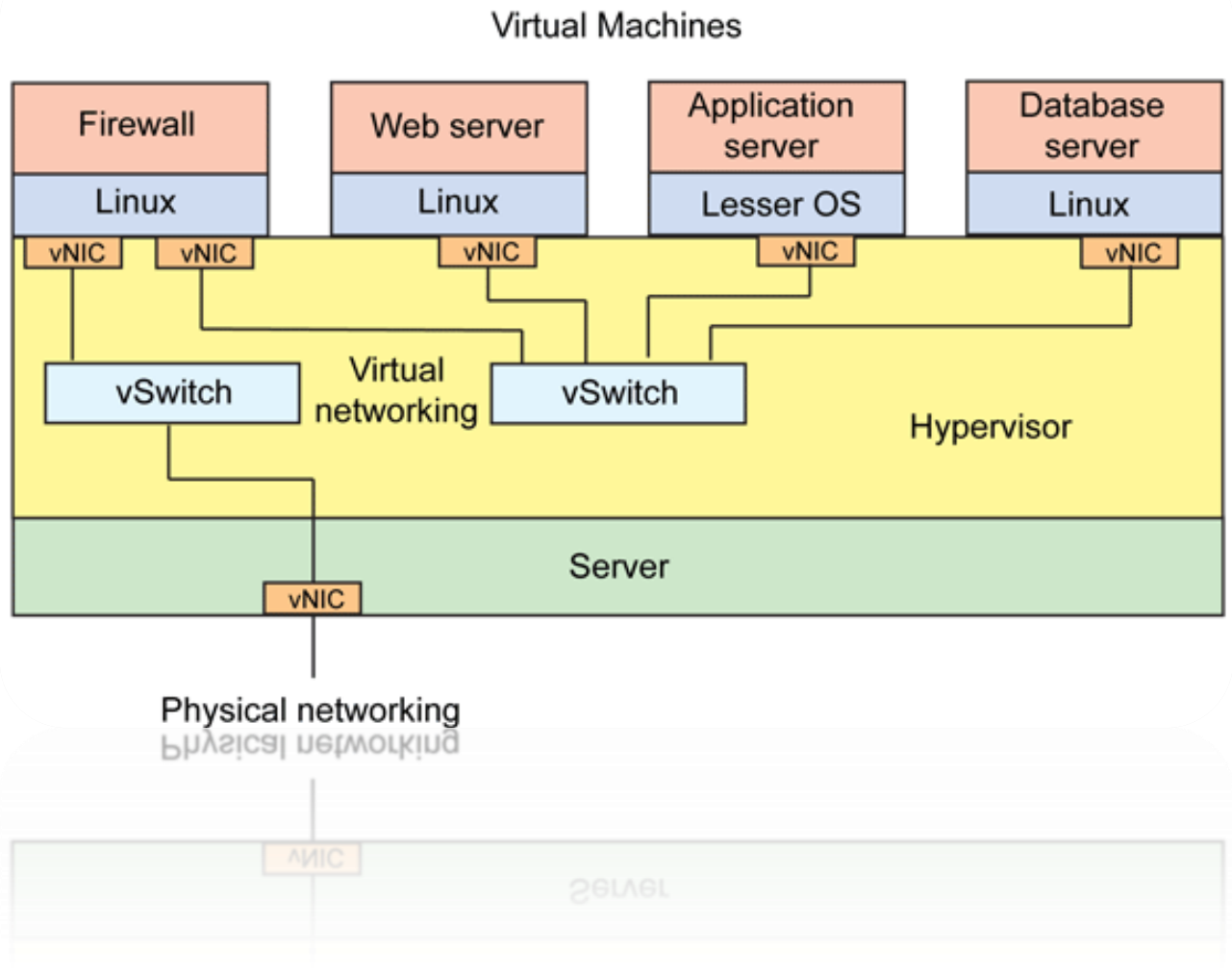
传统网络结构

图 1. 传统网络基础架构



OpenvSwitch网络

2. 虚拟的网络设施



OpenvSwitch网络

```
[root@kvm-02 ~]# ovs
ovs-appctl            ovsdb-server          ovs-parse-backtrace  ovs-tcpundump
ovs-benchmark         ovsdb-tool            ovs-parse-leaks      ovs-vlan-test
ovs-bugtool           ovs-dpctl             ovs-pcap             ovs-vsctl
ovsdb-client          ovs-ofctl             ovs-pki              ovs-vswitchd
```

```
[root@kvm-02 ~]# ovs-vsctl show
8ecebec4-78ed-43e7-955a-6242b126ab45
    Bridge "cloudbr1"
        Port "eth1"
            Interface "eth1"
        Port "vnet13"
            tag: 100
            Interface "vnet13"
        Port "vnet11"
            tag: 100
            Interface "vnet11"
        Port "vnet16"
            tag: 100
            Interface "vnet16"
```

```
Bridge "cloudbr0"
    Port "vnet1"
        Interface "vnet1"
    Port "vnet6"
        Interface "vnet6"
    Port "cloudbr0"
        Interface "cloudbr0"
            type: internal
    Port "vnet4"
        Interface "vnet4"
    Port "eth0"
        Interface "eth0"
```

OpenvSwitch组件

Our Practice	Bridge	Open vSwitch
LACP	Support(bond0)	Support(LACP Command)
VLAN	Support (bond0.100 for br0, bond0.101 for br1, etc… if I have many customers >_<#)	Support(Tag in vSwitch, ^_^)
Response on Live Migration	Slow (windows remote drop)	Fast (windows remote hold)

Somebody says:

虛擬交換 / 比較項目	OpenvSwitch	LinuxBridge
可移動性	有	有限
即時反應	有	無
邏輯標籤	有	無
處理效能	佳	較差
遠端控制	有	有限

演示操作

- 操作系统: CentOS 7
- 虚拟层: KVM
- 虚拟管理软件:Libvirt
- 虚拟交换机：OpenvSwitch
- 存储系统：Btrfs

Q&A