

# Using the KVM hypervisor in CloudStack

Dag Sonstebo

Cloud Architect / Senior Consultant

[dag.sonstebo@shapeblue.com](mailto:dag.sonstebo@shapeblue.com)

Twitter: [@dagsonstebo](https://twitter.com/dagsonstebo)



**Blue**

The Cloud Specialists



# About Me

- **Cloud Architect @ ShapeBlue**
- **Background:**
  - Cloud and virtualization architect with 19 years experience from the service provider, financial and manufacturing industries.
  - Specialize in:
    - Cloud infrastructure architecture and engineering.
    - Virtualization - VMware vSphere, Citrix XenServer, KVM.
    - Automation
  - Involved with CloudStack since version 2.1.



**“ShapeBlue are expert builders of public & private clouds. They are the leading global CloudStack services *company.*”**



# ShapeBlue customers



The Cloud Specialists

[ShapeBlue.com](http://ShapeBlue.com)



@ShapeBlue

# ShapeBlue customers



The Cloud Specialists

[ShapeBlue.com](http://ShapeBlue.com)

 @ShapeBlue

# ShapeBlue customers



The Cloud Specialists

[ShapeBlue.com](https://ShapeBlue.com)

[@ShapeBlue](https://twitter.com/ShapeBlue)

# CloudStack and hypervisor choice



The Cloud Specialists

[ShapeBlue.com](https://ShapeBlue.com)



@ShapeBlue

# CloudStack hypervisor choice

Hypervisor	Pros	Cons
VMware	<ul style="list-style-type: none"><li>• Feature rich</li><li>• Mature</li><li>• Native HA / DRS</li><li>• Bare metal</li></ul>	<ul style="list-style-type: none"><li>• Cost – capex.</li><li>• IP consumption</li><li>• Virtual Centre</li></ul>
XenServer	<ul style="list-style-type: none"><li>• Mature</li><li>• HA</li></ul>	<ul style="list-style-type: none"><li>• Disk chains</li><li>• Poolmaster</li><li>• No workload balancing</li></ul>
KVM	<ul style="list-style-type: none"><li>• Fast</li><li>• Maturing.</li><li>• Cost</li><li>• Non-proprietary</li></ul>	<ul style="list-style-type: none"><li>• Lacking some networking, storage and snapshotting features.</li><li>• No workload balancing</li></ul>



The Cloud Specialists

[ShapeBlue.com](http://ShapeBlue.com)



@ShapeBlue



# KVM background



The Cloud Specialists

[ShapeBlue.com](https://ShapeBlue.com)



@ShapeBlue

# KVM background

- **KVM = Kernel-based Virtual Machine.**
- **In CloudStack KVM is used as a type 2 hypervisor – provides virtualization support on top of a running Linux OS instance.**
- **Included in the Linux kernel since version 2.6.20 (2007).**
- **KVM is Free Software released under the GPL.**
- **Supported in CloudStack since early cloud.com days (version 2.0).**



The Cloud Specialists

**ShapeBlue.com**



**@ShapeBlue**

# KVM background

- **Hardware:**
  - 32 / 64 bit with CPU virtualization support (Intel-VTX / AMD-V).
- **Kernel modules**
  - kvm
  - kvm\_intel
  - kvm\_amd
- **Para-virtualization through the Virtio API.**
- **KVM and Qemu:**
  - KVM abstracts access to the CPU and memory.
  - QEMU emulates the hardware resources (disks, graphics, USB, etc).



The Cloud Specialists

[ShapeBlue.com](https://ShapeBlue.com)



@ShapeBlue

- **Pros:**
  - Flexible - not proprietary.
  - Fast.
  - Lightweight.
  - No SPOF – no poolmasters or VirtualCentre like components.
- **Cons:**
  - No workload balancing (yet...).
  - No network throttling.
  - No VM snapshot support.
  - No native clustered file system.
  - Root volume migration requires VM to be shut down.

# Installation and configuration



The Cloud Specialists

[ShapeBlue.com](https://ShapeBlue.com)



@ShapeBlue

# Base installation

- **Host OS:**
  - CentOS / RHEL 6.x / **7.x**
  - Ubuntu 12.04 / **14.04** / 16.04
  - 4GB memory / 64-bit X86.
- **Components (CloudStack 4.9):**
  - NTP + DNS
  - Libvirt 1.2.0 or higher
  - Qemu / KVM: 2.0 or higher
- **Installing the CloudStack agent installs KVM and libvirt:**
  - `yum install cloudstack-agent`
  - `apt-get install cloudstack-agent`



The Cloud Specialists

[ShapeBlue.com](http://ShapeBlue.com)



@ShapeBlue

# Configuration

- **/etc/libvirt/libvirtd.conf:**

```
listen_tls = 0
listen_tcp = 1
tcp_port = "16509"
auth_tcp = "none"
mdns_adv = 0
```

- **/etc/sysconfig/libvirtd:**

```
LIBVIRT_ARGS="--listen"
libvirtd_opts="-d -1"
```

- **Security policies:**

- CentOS: Selinux: permissive
- Ubuntu: Apparmor

- **Firewall ports:**

- TCP/22 (SSH)
- TCP/1798 (KVM)
- TCP/16509 (Libvirt)
- TCP/5900-6100 (VNC)
- TCP/49152-49216 (Libvirt migration)

# Networking and storage



The Cloud Specialists

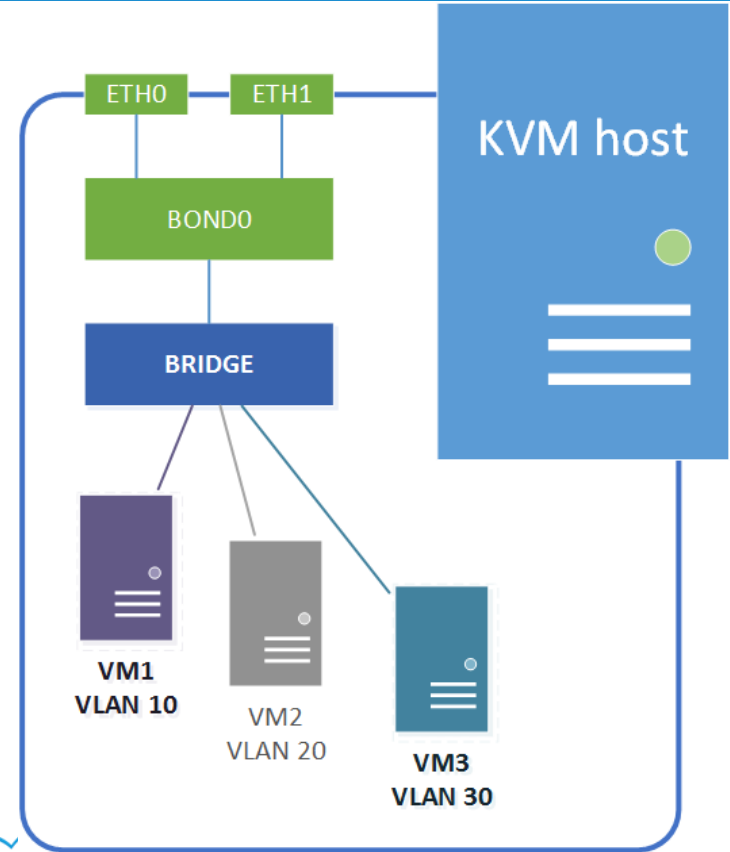
[ShapeBlue.com](https://ShapeBlue.com)



@ShapeBlue



- **KVM utilizes network bridges – similar to VMware vSwitches and XenServer networks.**
- **Bridging back ends:**
  - Linux bridge
  - Open Vswitch (OVS)



- **Linux bridge**
  - In Linux kernel since version 2.2.
  - Fast, simple, reliable and mature.
  - Provides L2 requirements for CloudStack.
  - Lacks automation options and scalable tunneling.

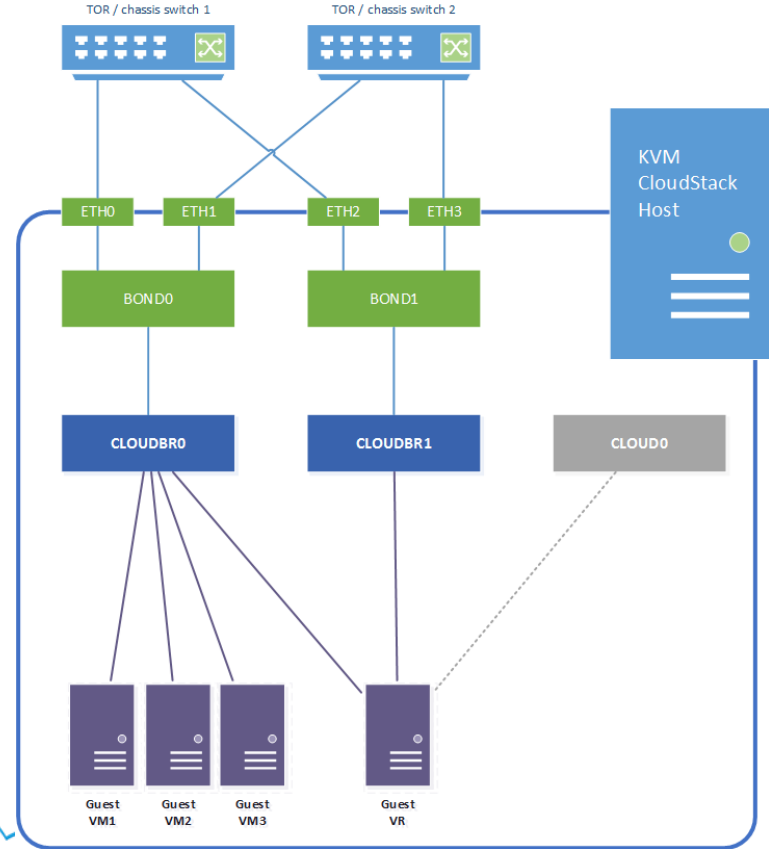
- **Open vSwitch**

- Written for multi-host virtualization environments with dynamic end points, higher level of abstraction and potential for hardware offloading.
- Network state kept in a network state database (OVSDB).
- Allows for better automation.
- SDN options (tunneling).
- External controllers: OpenDaylight, Nicira, VMware NSX.
- Default bridge backend in XenServer 6.0 and later versions (does not work with basic zones).
- Some issues with VLAN configuration and stability (depending on host OS and version).



# Networking example

- **Cloudbr0:**
  - Bond0: Eth0 + Eth1
  - Carries management, guest and storage traffic.
- **Cloudbr1:**
  - Bond1: Eth2 + Eth3
  - Carries public traffic.
- **Cloud0:**
  - Internal bridge, carries system VM management traffic.
  - Managed by CloudStack – does not need configuration.



# Linux bridge configuration (CentOS)

## ifcfg-eth0:

```
DEVICE=eth0  
TYPE=Ethernet  
USERCTL=no  
MASTER=bond0  
SLAVE=yes  
BOOTPROTO=none  
NM_CONTROLLED=no  
ONBOOT=yes
```

## ifcfg-bond0:

```
DEVICE=bond0  
ONBOOT=yes  
BONDING_OPTS='mode=1  
miimon=100'  
BRIDGE=cloudbr0  
NM_CONTROLLED=no
```

## ifcfg-cloudbr0:

```
DEVICE=cloudbr0  
ONBOOT=yes  
TYPE=Bridge  
IPADDR=192.168.100.20  
NETMASK=255.255.255.0  
GATEWAY=192.168.100.1  
NM_CONTROLLED=no  
DELAY=0
```



The Cloud Specialists

[ShapeBlue.com](https://ShapeBlue.com)



@ShapeBlue

## ifcfg-cloudbr1:

DEVICE=cloudbr1

ONBOOT=yes

TYPE=Bridge

NM\_CONTROLLED=no

DELAY=0

## Storage VLAN

### ifcfg-bond.100:

```
DEVICE=bond0.100  
VLAN=yes  
BOOTPROTO=static  
ONBOOT=yes  
TYPE=Unknown  
BRIDGE=cloudbri100
```

## Storage VLAN

### ifcfg-cloudbri100:

```
DEVICE=cloudbri100  
ONBOOT=yes  
TYPE=Bridge  
VLAN=yes  
IPADDR=10.0.100.20  
NETMASK=255.255.255.0  
NM_CONTROLLED=no  
DELAY=0
```

# Linux bridge configuration (CentOS)

- **Check bridges:**

```
# brctl show
```

bridge name	bridge id	STP enabled	interfaces
brbond0-113	8000.000c29b43c4d	no	bond0.113
cloudbr0	8000.000c29b55932	no	bond0
cloudbr1	8000.000c29b45956	no	bond1
cloudbr100	8000.000c29b43c4d	no	bond0.100

- **Check bonds:**

```
# cat /proc/net/bonding/bond0
```

```
Ethernet Channel Bonding Driver: v3.7.1 (April 27, 2011)
```

```
Bonding Mode: fault-tolerance (active-backup)
```

```
Primary Slave: None
```

```
Currently Active Slave: eth0
```

```
MII Status: up
```

```
MII Polling Interval (ms): 100
```

```
Up Delay (ms): 0
```

```
Down Delay (ms): 0
```

```
Slave Interface: eth0
```

```
MII Status: up
```



The Cloud Specialists

[ShapeBlue.com](https://ShapeBlue.com)



@ShapeBlue



# OVS configuration (CentOS)

- **Requires OVS installation:**

```
# apt-get install openvswitch-switch
```

```
# yum install openvswitch-<version>.rpm
```

```
# yum install openvswitch-kmod-<version>.rpm
```

- **Add bridges and bonds with OVS command line tools:**

```
# ovs-vsctl add-br cloudbr0
```

```
# ovs-vsctl add-br cloudbr1
```

```
# ovs-vsctl add-bond cloudbr0 bond0 eth0 eth1
```

```
# ovs-vsctl add-bond cloudbr1 bond1 eth2 eth3
```



# OVS configuration (CentOS)

## ifcfg-eth0:

```
DEVICE=eth0  
ONBOOT=yes  
NM_CONTROLLED=no  
BOOTPROTO=none  
HOTPLUG=no
```

## ifcfg-bond0:

```
DEVICE=bond0  
ONBOOT=yes  
DEVICETYPE=ovs  
TYPE=OVSBond  
OVS BRIDGE=cloudbr0  
BOOTPROTO=none  
BOND IFACES="eth0 eth1"  
OVS OPTIONS="bond mode=  
active-backup lacp=off  
other config:bond-  
detect-mode=miimon  
other config:bond-  
miimon-interval=100"  
HOTPLUG=no
```

## Ifcfg-cloudbr0:

```
DEVICE=cloudbr0  
ONBOOT=yes  
DEVICETYPE=ovs  
TYPE=OVSBridge  
BOOTPROTO=static  
IPADDR=192.168.100.20  
NETMASK=255.255.255.0  
HOTPLUG=no
```

## ifcfg-cloud0:

```
DEVICE=cloud0  
ONBOOT=yes  
DEVICETYPE=ovs  
TYPE=OVSBridge  
BOOTPROTO=static  
HOTPLUG=no  
IPADDR=169.254.0.1  
NETMASK=255.255.0.0
```

## VLAN configuration:

```
# ovs-vsctl add-br cloudbr100 cloudbr0 100
```

## ifcfg-cloudbr100

```
DEVICE=cloudbr100  
ONBOOT=yes  
DEVICETYPE=ovs  
TYPE=OVSBridge  
BOOTPROTO=static  
IPADDR=10.0.100.20  
NETMASK=255.255.255.0  
HOTPLUG=no
```

# OVS configuration (CentOS)

- **Some issues with network drivers cause VLANs to not propagate:**

```
# ovs-vsctl set interface eth0 other-config:enable-vlan-splinters=true
```

- **List bridges:**

```
# ovs-vsctl show
27daed4e-52f3-4177-9827-550f0e7df452
```

```
Bridge "cloudbr1"
    Port "vnet2"
        Interface "vnet2"
    Port "bond1"
        Interface "eth3"
        Interface "eth2"
    Port "cloudbr1"
        Interface "cloudbr1"
            type: internal
Bridge "cloud0"
```

...

- **List bonds:**

```
ovs-appctl bond/show bond0
---- bond0 ----
bond_mode: active-backup
bond may use recirculation: no,
Recirc-ID : -1
bond-hash-basis: 0
updelay: 0 ms
downdelay: 0 ms
lacp_status: off
active slave mac:
00:0c:xx:xx:xx:xx(eth0)
    slave eth0: enabled
        active slave
            may_enable: true
    slave eth1: enabled
        may_enable: true
```



# OVS configuration (CentOS)

- **Agent properties file - /etc/cloudstack/agent/agent.properties:**

```
guest.network.device=cloudbr0  
workers=5  
private.network.device=cloudbr0  
network.bridge.type=openvswitch  
port=8250  
resource=com.cloud.hypervisor.kvm.resource.LibvirtComputingResource  
pod=3  
zone=2  
hypervisor.type=kvm  
guid=c3c6a2fc-13d3-3af1-ae2e-e48a21d9b883  
public.network.device=cloudbr1  
cluster=5  
local.storage.uuid=2f2220ef-2624-4e69-b442-0a4ae5c5add6  
domr.scripts.dir=scripts/network/domr/kvm  
LibvirtComputingResource.id=28  
host=192.168.100.20
```

- **Disks, templates and snapshots use QCOW2 format.**
- **Primary storage options:**
  - NFS – recommended if performance is sufficient, easy option, supports CloudStack managed thin provisioning.
  - Local storage (does not support storage migration).
  - SharedMountpoint option used for iSCSI / FC.
  - CEPH.
- **SharedMountpoint:**
  - Used to access pre-defined block storage.
  - Must be running a clustered file system.
  - Preconfigured on each host.
  - Consistent – same on each host – e.g. /mnt/primary



- **KVM lacks a native clustered file system like VMFS or SR (CLVM).**
- **Clustered file system options:**
  - OCFS2 (Oracle)
  - GFS2 (RedHat)
  - CLVM (not officially supported in CloudStack)
- **CEPH – Wido is wizard...😊**

# Management and troubleshooting



The Cloud Specialists

[ShapeBlue.com](https://ShapeBlue.com)



@ShapeBlue



# How do I manage KVM

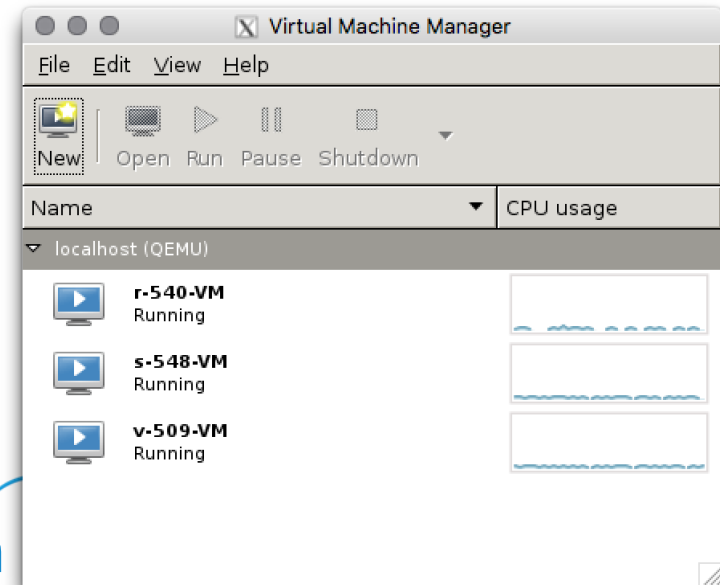
- **Countless tools (40+).**
- **Most common:**
  - Virsh command line tool:
  - Virt-manager: linux native but works well with SSH X session forwarding.
- **Configuration management:**
  - Anything that utilises libvirt.
  - Ansible: virt modules
  - Puppet modules and Chef Cookbooks.

```
# ssh -X kvmhost1
```

```
root@kvmhost1:~# virsh list
```

Id	Name	State
2	r-540-VM	running
3	s-548-VM	running
4	v-509-VM	running

```
root@kvmhost1:~# virt-manager &
```



# Troubleshooting

- **Check KVM is running:**

```
# lsmod | grep kvm
```

```
kvm_intel          151552    9
kvm                479232    1 kvm_intel
```

- **Log file: /var/log/cloudstack/agent/agent.log**
- **Increase logging verbosity:**  

```
# sed -i 's/INFO/DEBUG/g' /etc/cloudstack/agent/log4j-cloud.xml
```
- **KVM uses link local networking, hence connect to system VMs using:**  

```
# ssh -i /root/.ssh/id_rsa.cloud -p 3922 root@<linklocalIP>
```



## What's next

- **HA (and IPMI control) is being developed for KVM.**

<https://cwiki.apache.org/confluence/display/CLOUDSTACK/KVM+HA+with+IPMI+Fencing>

- **DRS functionality has been discussed but is not yet in development.**



The Cloud Specialists

**ShapeBlue.com**



**@ShapeBlue**

# Questions?



The Cloud Specialists

[ShapeBlue.com](https://ShapeBlue.com)



@ShapeBlue

- **Background:**
  - <http://www.linux-kvm.org>
  - <http://wiki.qemu.org/KVM>
  - <https://libvirt.org/>
- **Installation:**
  - <http://docs.cloudstack.apache.org/projects/cloudstack-installation/en/4.9/hypervisor/kvm.html>
- **Management tools:**
  - [http://www.linux-kvm.org/page/Management\\_Tools](http://www.linux-kvm.org/page/Management_Tools)

- **Networking:**

- <http://openvswitch.org>
- <http://openvswitch.org/support/dist-docs/WHY-OVS.md.txt>

- **Storage:**

- CEPH: <http://docs.ceph.com/docs/hammer/rbd/rbd-cloudstack/>
- OCFS2: <https://oss.oracle.com/projects/ocfs2/>
- GFS2: [https://access.redhat.com/documentation/en-US/Red\\_Hat\\_Enterprise\\_Linux/6/html-single/Global\\_File\\_System\\_2/index.html](https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/html-single/Global_File_System_2/index.html)

## More information

- Slide deck: <http://www.slideshare.net/shapeblue>
- Blog: <http://shapeblue.com/blog>  
<http://dsonstebo.wordpress.com>
- Email: [dag.sonstebo@shapeblue.com](mailto:dag.sonstebo@shapeblue.com)
- Twitter: @dagsonstebo
- Web: <http://shapeblue.com>



The Cloud Specialists

**ShapeBlue.com**



**@ShapeBlue**

- My next talk:
  - What: **Apache CloudStack Upgrade Best Practices**
  - When: **Tomorrow – Thursday @ 2:40PM**
  - Where: **Rafael (here....)**



# Using the KVM hypervisor in CloudStack

Dag Sonstebo

Cloud Architect / Senior Consultant

[dag.sonstebo@shapeblue.com](mailto:dag.sonstebo@shapeblue.com)

Twitter: [@dagsonstebo](https://twitter.com/dagsonstebo)



**Blue**

The Cloud Specialists

