

A photograph taken from the interior of a car. A woman with long blonde hair is in the passenger seat, looking towards the driver. The driver is a man with a beard, wearing a dark jacket over a light blue shirt, with his hands raised in a gesture of surrender or panic. The car's interior is visible, including the seats and windows. The text 'KUBERNETES, TAKE THE WHEEL' is overlaid in large, white, bold letters with a black outline at the bottom of the image.

**KUBERNETES,
TAKE THE WHEEL**

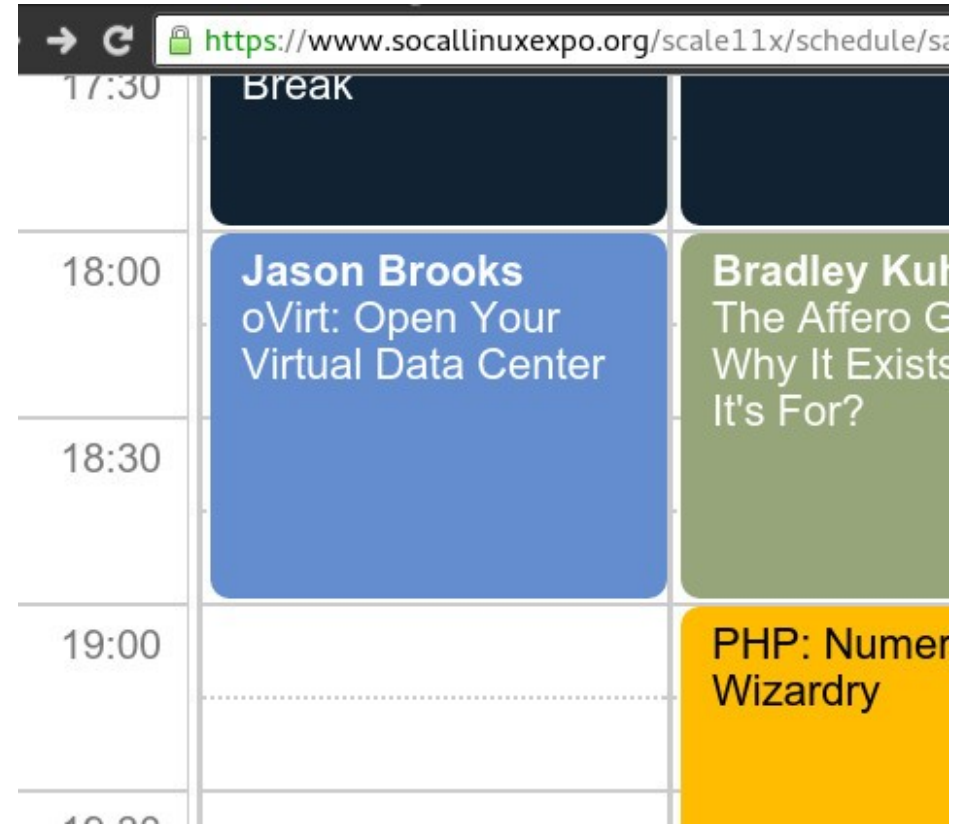
Some Background

- eWEEK Labs
 - research, test & write
 - x86 virtualization
 - operating systems
 - linux & open source
- Red Hat OSAS
 - oVirt, RDO, CentOS, Atomic
 - still writing & testing



User-Mode Me

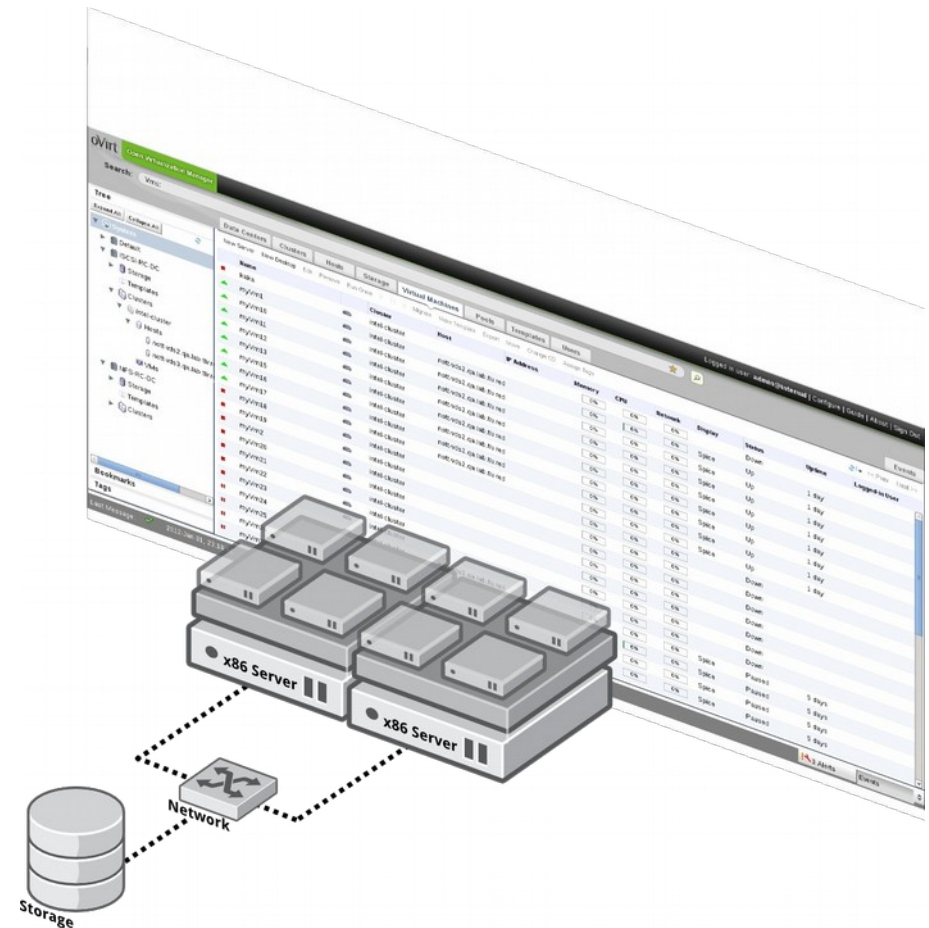
- Everything in a VM
- Nested KVM
- oVirt Rocks



Time	Session 1	Session 2
17:30	Break	
18:00	Jason Brooks oVirt: Open Your Virtual Data Center	Bradley Kuhl The Affero GPL Why It Exists It's For?
18:30		
19:00		PHP: Numerical Wizardry
19:30		

Hello, oVirt

- Large scale, centralized management for server and desktop virtualization
- Based on KVM
- Provides an open source alternative to vCenter/vSphere
- Upstream for RHEV





That's all Folks!

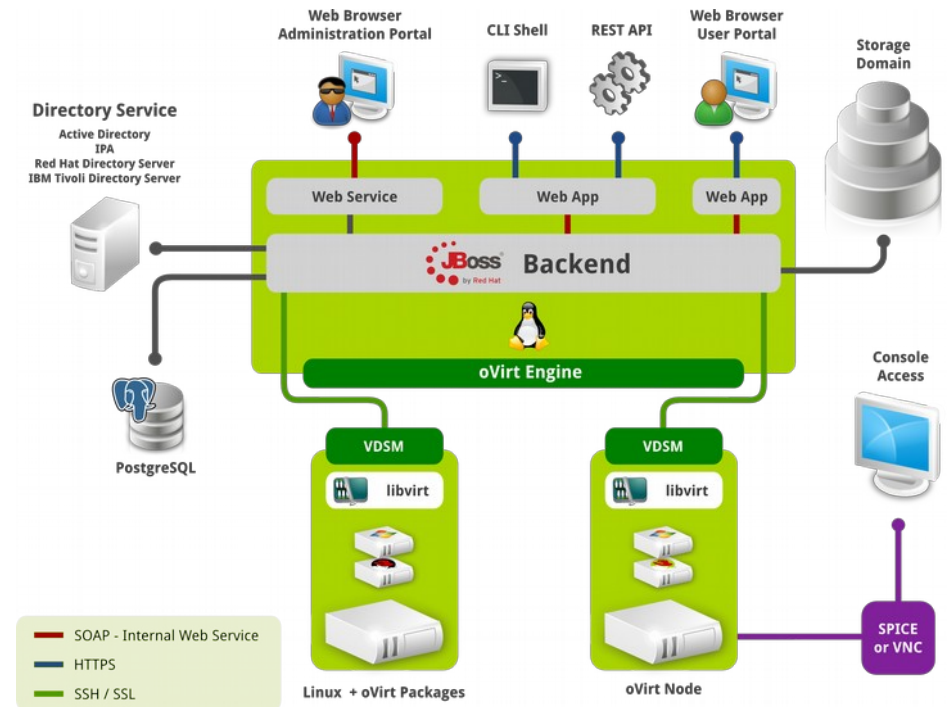
Admin-Mode Me

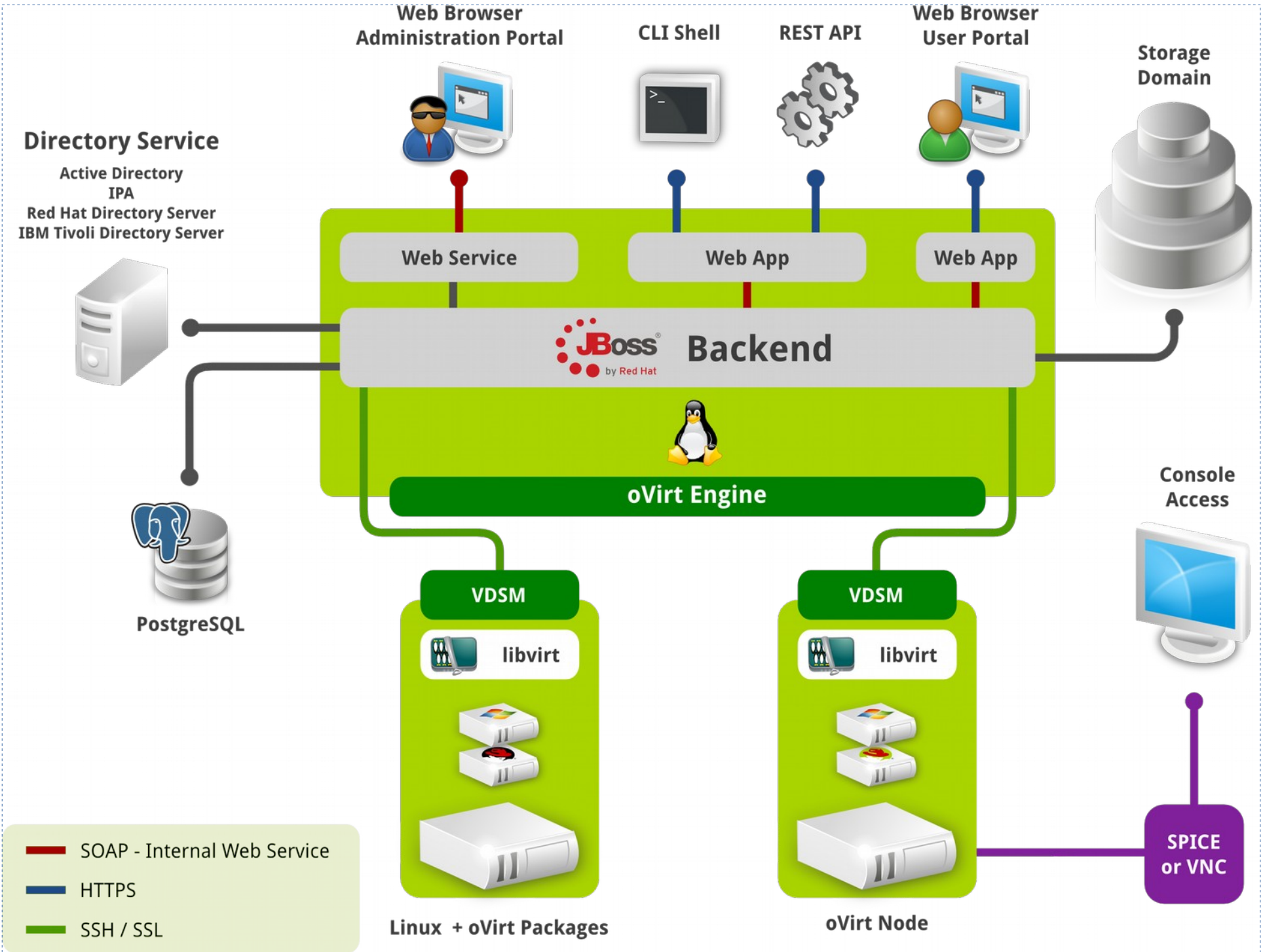
- I run my lab
- All upstream software
- Stack of regular servers
- Try to stay on latest versions
- Keep my configs close enough to “stock” to be helpful



oVirt Needs

- Shared storage
- Identity services
- Management server
- Additional pieces





Hyperconvergence

- Gluster Storage
 - replica 3 for split BRAAAAINS
- oVirt Virt Hosts
 - regular CentOS 7
- Self-Hosted Engine
- Assorted Vms
 - Freeipa, Glance, Neutron, Cinder, Optaplanner

Up and Running with oVirt 3.5

by [Jason Brooks](#) - Wednesday 29 October 2014



Last week, version 3.5 of oVirt, the open source virtualization management platform, hit [FTP mirrors](#) sporting a slate of fixes, a new web interface, and support for using CentOS 7 as a virtualization host.

As with every new oVirt release, I'm happy to see the project on single server, with an option for expanding to additional hosts. In this quick rundown of the different single-machine options for trying out oVirt 3.5, I'll be walking through the steps you can follow to get up and running.

- [oVirt Live ISO](#): A LiveCD image that you can burn onto a blank CD or DVD and run oVirt. This is probably the fastest way to get up and running, but it's the lowest-performance option, and not suitable for extended use.
- [oVirt All in One plugin](#): Run the oVirt management server and engine on the same machine with local storage. This is a more permanent version of the All-in-One option favored by many users until the rise of... [the Hosted Engine](#).
- [oVirt Hosted Engine](#): The self-hosted engine approach consists of running the oVirt management engine on its own management engine. This route is a bit more complicated, but it offers the most flexibility. Running oVirt Engine in a separate VM allows you to run oVirt on CentOS 6 around for the engine. With the All-in-One approach, your management engine is on the same machine, limiting your expansion options. The Hosted Engine can be run on CentOS 7 as a virtualization host, but it's not supported on CentOS 6.

For this howto, I'll be walking through the steps you can follow to get up and running.

Convergence Challenges

- Packaging conflicts
- Resource management
- Deployment complexity
- One-off management processes

[Add an attachment](#) (proposed patch, testcase, etc.)

Jason Brooks 2012-07-13 13:22:40 EDT

Description of problem:

freeipa-server conflicts with 1:mod_ssl-2.2.22-4.fc17.x86_64

Version-Release number of selected component (if applicable):

2.2.0-1.fc17

How reproducible:

On Fedora 17 system with mod_ssl-2.2.22-4.fc17.x86_64 installed, freeipa-server.

Actual results:

Package conflict, freeipa-server won't install.

Expected results:

freeipa-server installs

Additional info:

mod_ssl is a dependency of ovirt-engine. I'm attempting to install on the same machine as ovirt-engine 3.1 (from <http://www.ovirt.org/releases/beta/fedora/17/>) to use as an identifier.

What I Want

- Support upstream packaging
- Resource controls
- Avoid unnecessary overhead
- Stick close to our projects
- ONE to Rule Them



Other Options?

- oVirt / OpenStack Overlord?
 - VM overhead
- OpenShift?
 - v2...
- Hand-wavy Container Option?



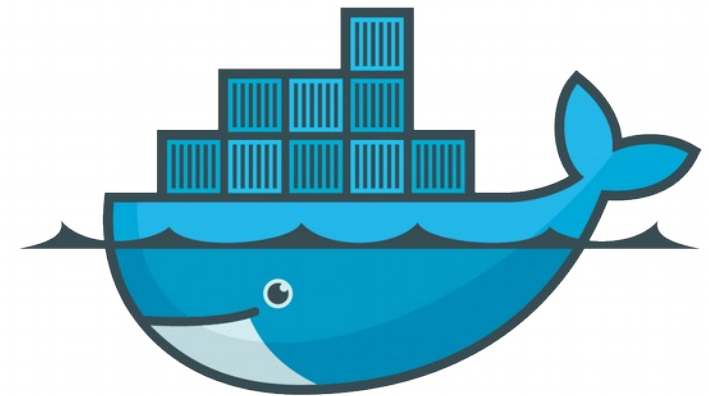
Containerization

- BYO dependencies; decent isolation; low-overhead
- I've been interested since Solaris 10
 - but, Solaris...
- LXC, OpenVZ, Linux-Vservers
 - not enough traction

David Rodriguez Practical Application Debugging Tracing	Jérôme Petazzoni Lightweight Virtualization with namespaces, cgroups, and unioning filesystems	Sofia Kelly How To Be A Picasso Using Tux Paint
Lu Lavigne Extending NAS Functionality with FreeNAS	Joseph Guarino Linux and Windows Inter-operability – Making Linux and	Philip Ballew How Open Source Benefits Youth

Enter Docker

- Containers, executed well
- Use whatever packaging you like
- Docker Hub full of examples
- Magical traction



docker

Enter Kubernetes

- Container orchestrator
- Manage applications, not machines
- Based on Google's experiences and internal systems



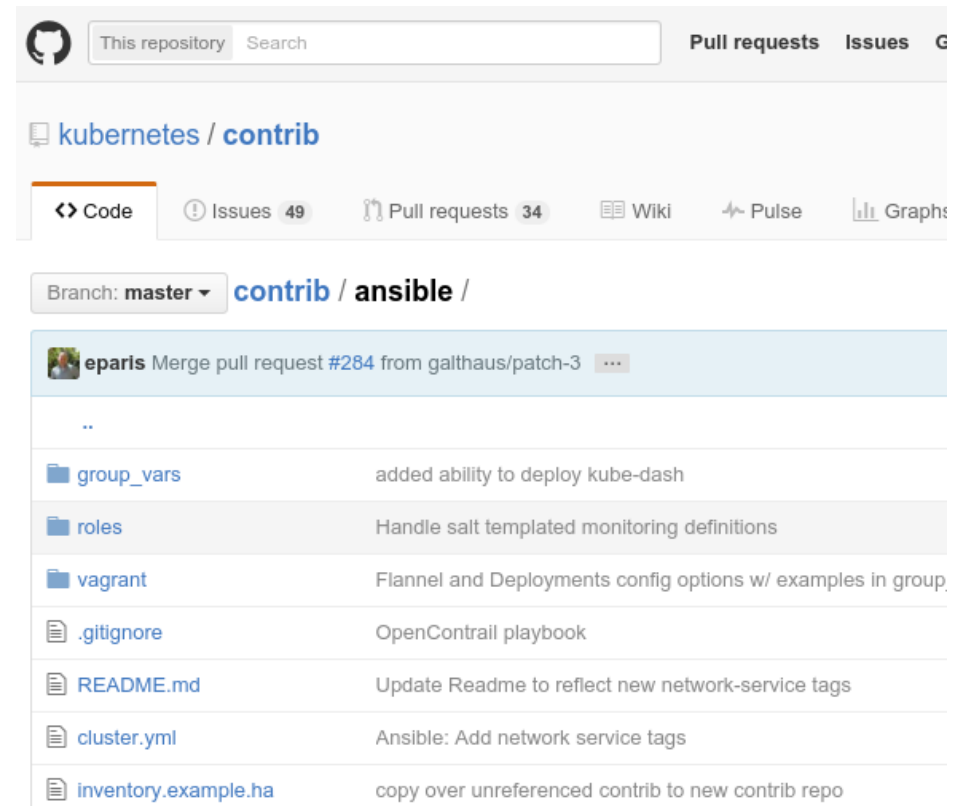
Kubernetes App Anatomy

- Container: A sealed application package (Docker)
- Pod: A small group of tightly coupled Containers
- Controller: A loop that drives current state towards desired state
- Service: A set of running pods that work together



The Cluster

- “regular” CentOS 7
 - kubernetes from CentOS pkg repos
- contrib/ansible scripts
 - w/ or w/o vagrant
- using single master
 - scripts offer multi-master option

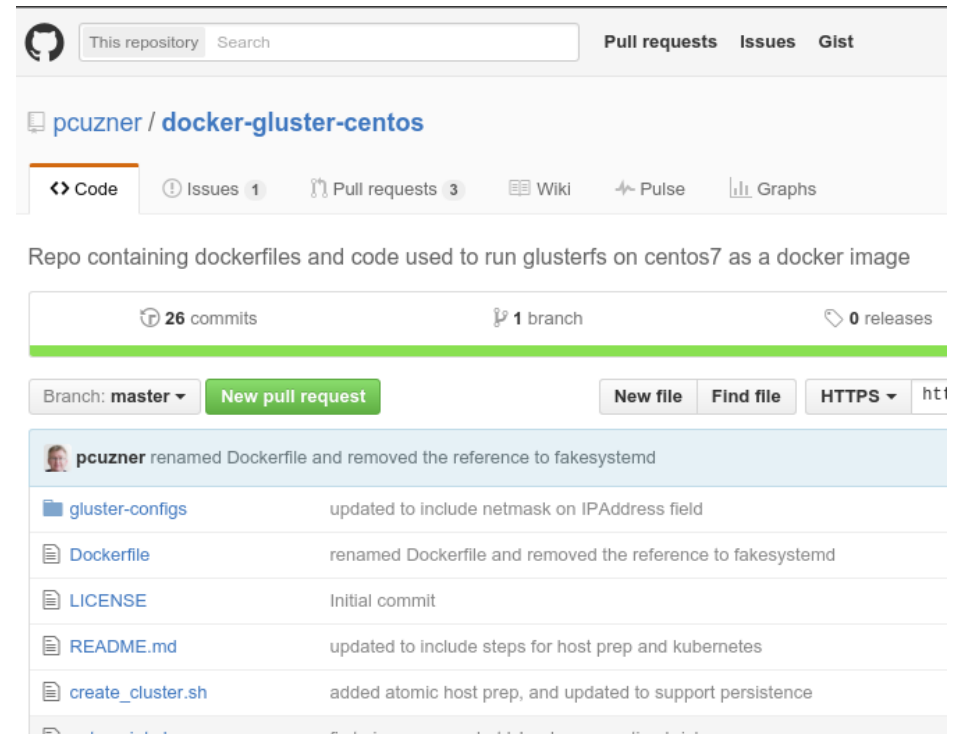


The screenshot shows the GitHub interface for the `kubernetes/contrib` repository. The repository is on the `master` branch, and the current view is for the `contrib/ansible` directory. A pull request #284 by `eparis` is being merged from `galthaus/patch-3`. The repository structure is as follows:

File/Folder	Description
<code>..</code>	
<code>group_vars</code>	added ability to deploy kube-dash
<code>roles</code>	Handle salt templated monitoring definitions
<code>vagrant</code>	Flannel and Deployments config options w/ examples in group.
<code>.gitignore</code>	OpenContrail playbook
<code>README.md</code>	Update Readme to reflect new network-service tags
<code>cluster.yml</code>	Ansible: Add network service tags
<code>inventory.example.ha</code>	copy over unreferenced contrib to new contrib repo

Storage

- Gluster in a container
- Etcd configuration
- Host attachment
 - Data stored in specific brick device on host
 - Configs live in dirs mounted on host



```
FROM centos:latest
MAINTAINER Paul Cuzner <pcuzner@redhat.com>
ENV container docker

RUN curl -o /etc/yum.repos.d/glusterfs-epel.repo \
    http://download.gluster.org/pub/gluster/glusterfs/3.7/3.7.6/EPEL.repo/glusterfs-epel.repo

RUN yum -y install epel-release

RUN yum --setopt=tsflags=nodocs -y install xfsprogs nfs-utils nmap-ncat \
    openssh-server openssh-clients attr iputils iproute net-tools \
    glusterfs glusterfs-server glusterfs-fuse glusterfs-geo-replication \
    glusterfs-cli glusterfs-api && yum clean all -y

VOLUME [ "/sys/fs/cgroup" ]

RUN systemctl enable glusterd.service sshd.service

RUN mkdir -p /build/config/{etc/glusterfs,var/lib/glusterd,var/log/glusterfs}

RUN cp -pr /etc/glusterfs/* /build/config/etc/glusterfs && \
    cp -pr /var/lib/glusterd/* /build/config/var/lib/glusterd && \
    cp -pr /var/log/glusterfs/* /build/config/var/log/glusterfs

ADD entrypoint.sh /build/entrypoint.sh
ADD utils.sh /build/utils.sh
ADD create_cluster.sh /build/create_cluster.sh

RUN echo "root:password" | chpasswd

EXPOSE 22 111 245 443 24007 2049 8080 6010 6011 6012 38465 38466 38468 \
    38469 49152 49153 49154 49156 49157 49158 49159 49160 49161 49162

ENTRYPOINT ["/build/entrypoint.sh"]
```

- nodeSelector to attach to particular host
- using host network

```
spec:  
  hostNetwork: true  
  nodeSelector:  
    GlusterNode: gluster-6  
  containers:  
    - name: glusterfs  
      image: jasonbrooks/glusterfs-centos  
      ports:  
        - name: web
```


- hostPath volumes
- mounts brick device on host during container startup

```
volumeMounts:  
  - name: glusterfs-etc  
    mountPath: "/etc/glusterfs"  
  - name: glusterfs-logs  
    mountPath: "/var/log/glusterfs"  
  - name: glusterfs-config  
    mountPath: "/var/lib/glusterd"  
  - name: glusterfs-devtree  
    mountPath: "/dev"  
  - name: glusterfs-cgroup  
    mountPath: "/sys/fs/cgroup"  
securityContext:  
  capabilities: {}  
  privileged: true  
volumes:  
  - name: glusterfs-etc  
    hostPath:  
      path: "/etc/glusterfs"  
  - name: glusterfs-logs  
    hostPath:  
      path: "/var/log/glusterfs"
```

Engine, et al

- systemd-based CentOS images
- Persistent volumes for data that needs it
- Exposed via kubernetes service

Pods

Name	Namespace
 engine	
Pod	Containers
Logs	Shell

```
Jan 19 16:34:38 engine systemd[1]: Starting oVirt Engine.
Jan 19 16:34:39 engine systemd[1]: Started oVirt Engine.
sh-4.2# systemctl status ovirt-engine -l
● ovirt-engine.service - oVirt Engine
   Loaded: loaded (/usr/lib/systemd/system/ovirt-engine.service; vendor preset: disabled)
   Active: active (running) since Tue 2016-01-19 16:34:39 UTC; 1min 10s ago
     Main PID: 23787 (ovirt-engine.py)
       CGroup: /system.slice/docker-677cbf54079fe533c7401b10a9d4e41e96.scope/system.slice/ovirt-engine.service
               └─23787 /usr/bin/python /usr/share/ovirt-engine/ovirt-engine.py --redirect-output --systemd=notify
                   └─23891 ovirt-engine -server -XX:+TieredCompilation -XX:PermSize=256m -XX:MaxPermSize=256m -Djava.gc.client.gcInterval=3600000 -Dsun.rmi.dgc.server.leSNIExtension=false -Djava.security.krb5.conf=/etc/krb5.conf -Djava.security.sasl.conf=/etc/sasl.conf -Djboss.logging.manager=org.jboss.logmanager -Dlogging.ovirt-engine/jboss.runtime/config/ovirt-engine-logging.properties -Djboss.solver.warning=true -Djboss.modules.system.pkgs=org.jboss.byteman -Djboss.server.default.config=/usr/share/ovirt-engine-wildfly -Djboss.server.config.dir=/usr/share/ovirt-engine-wildfly -Djboss.server.data.dir=/var/lib/ovirt-engine -D
```

Gluster PV

```
apiVersion: "v1"
kind: "PersistentVolume"
metadata:
  name: "vol5"
spec:
  capacity:
    storage: "10Gi"
  accessModes:
    - "ReadWriteMany"
  glusterfs:
    endpoints: "glusterfs-cluster"
    path: "vol5"
    readOnly: false
  persistentVolumeReclaimPolicy: "Recycle"
```

Specific PV Claim

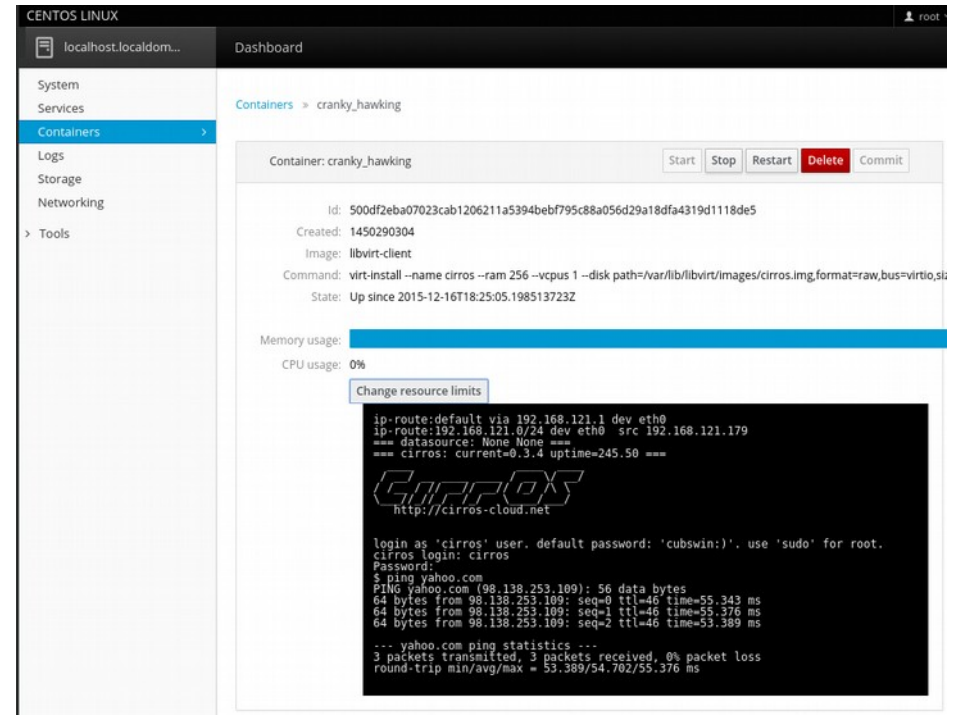
```
kind: PersistentVolumeClaim
apiVersion: v1
metadata:
  name: engine-pgsql
spec:
  accessModes:
    - ReadWriteMany
  resources:
    requests:
      storage: 9Gi
```


Engine Pod

```
volumeMounts:
  - name: engine-pgsql
    mountPath: "/var/lib/pgsql/"
  - name: etc-ovirtengine
    mountPath: "/etc/ovirt-engine"
  - name: etc-ovirtpki
    mountPath: "/etc/pki/ovirt-engine"
  - name: etc-ovirtsysconfig
    mountPath: "/etc/sysconfig/ovirt-engine"
  - name: var-logengine
    mountPath: "/var/log/ovirt-engine"
securityContext:
  capabilities: {}
  privileged: true
volumes:
  - name: engine-pgsql
    persistentVolumeClaim:
      claimName: "engine-pgsql"
```

Virtualization

- This *can* work
 - Libvirt in a container
 - RancherVM
 - Kolla



The screenshot displays a web-based dashboard for managing containers on a CentOS Linux system. The interface includes a sidebar with navigation options like System, Services, Containers, Logs, Storage, Networking, and Tools. The main content area shows details for a container named 'cranky_hawking', including its ID, creation time, image, and command. It also features resource usage bars for memory and CPU, and a terminal window showing the container's internal state, including network configuration, uptime, and ping results.

```
CENTOS LINUX
localhost.localdom... Dashboard
Containers > cranky_hawking

Container: cranky_hawking [Start] [Stop] [Restart] [Delete] [Commit]

id: 500df2eba07023cab1206211a5394beb795c88a056d29a18dfa4319d1118de5
Created: 1450290304
Image: libvirt-client
Command: virt-install --name cirros --ram 256 --vcpu 1 --disk path=/var/lib/libvirt/images/cirros.img,format=raw,bus=virtio,si
State: Up since 2015-12-16T18:25:05.198513723Z

Memory usage: [Progress Bar]
CPU usage: 0%
[Change resource limits]

ip-route:default via 192.168.121.1 dev eth0
ip-route:192.168.121.0/24 dev eth0 src 192.168.121.179
=== datasource: None None ===
=== cirros: current=0.3.4 uptime=245.50 ===

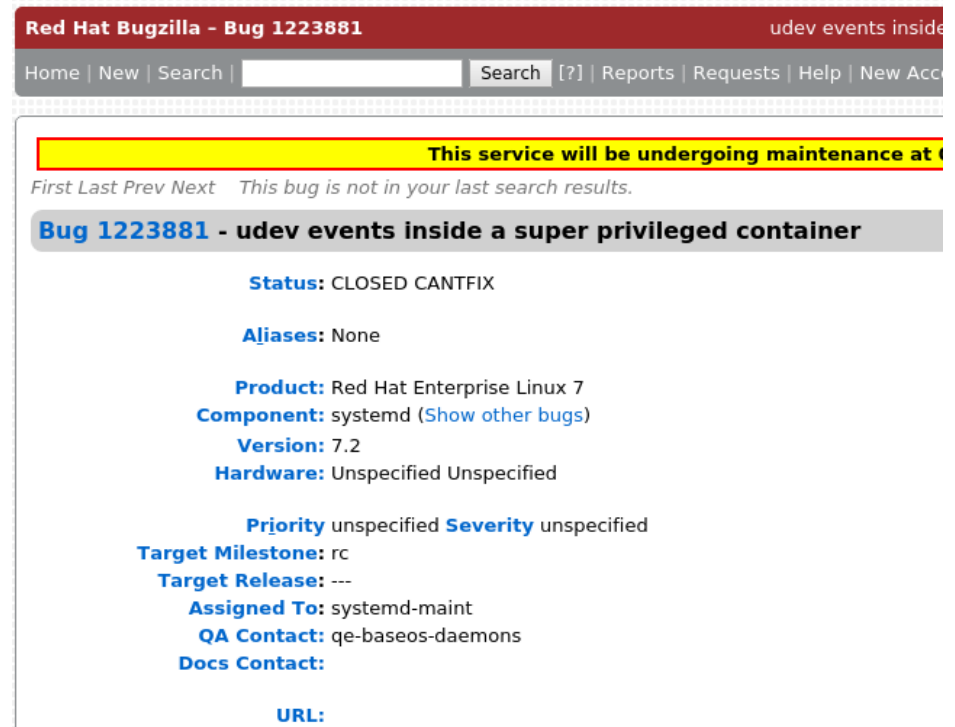
CIRRROS
http://cirros-cloud.net

login as 'cirros' user, default password: 'cubswin:)', use 'sudo' for root.
cirros login: cirros
Password:
$ ping yahoo.com
PING yahoo.com (98.138.253.109): 56 data bytes
64 bytes from 98.138.253.109: seq=0 ttl=46 time=55.343 ms
64 bytes from 98.138.253.109: seq=1 ttl=46 time=55.376 ms
64 bytes from 98.138.253.109: seq=2 ttl=46 time=53.389 ms

--- yahoo.com ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 53.389/54.702/55.376 ms
```

However...

- oVirt expects a lot, in terms of “real” hardware
- SPCs are a continuing area of study
- I'm leaving this bit uncontained for now...



The screenshot shows a Red Hat Bugzilla page for bug 1223881. The page title is "Red Hat Bugzilla - Bug 1223881" and the bug title is "Bug 1223881 - udev events inside a super privileged container". The status is "CLOSED CANTFIX". The product is "Red Hat Enterprise Linux 7", the component is "systemd", and the version is "7.2". The hardware is "Unspecified Unspecified". The priority is "unspecified" and the severity is "unspecified". The target milestone is "rc" and the target release is "---". The assigned to is "systemd-maint", the QA contact is "qe-baseos-daemons", and the docs contact is not specified. The URL is not specified. A yellow banner at the top of the bug page reads "This service will be undergoing maintenance at".

Red Hat Bugzilla - Bug 1223881 udev events inside

Home | New | Search | Search [?] | Reports | Requests | Help | New Acc

This service will be undergoing maintenance at

First Last Prev Next This bug is not in your last search results.

Bug 1223881 - udev events inside a super privileged container

Status: CLOSED CANTFIX

Aliases: None

Product: Red Hat Enterprise Linux 7
Component: systemd ([Show other bugs](#))
Version: 7.2
Hardware: Unspecified Unspecified

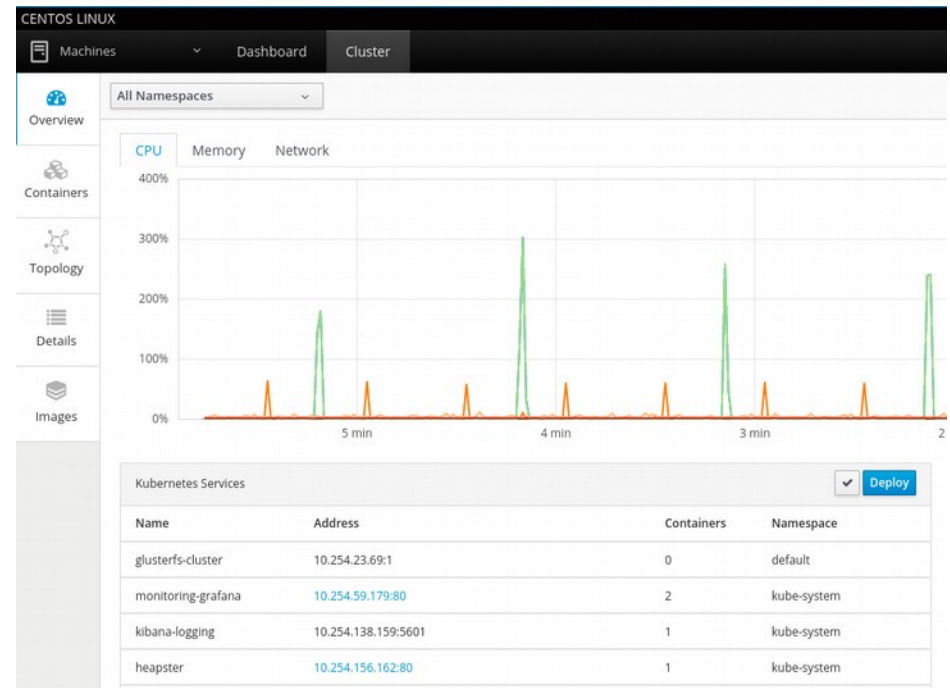
Priority: unspecified **Severity:** unspecified

Target Milestone: rc
Target Release: ---
Assigned To: systemd-maint
QA Contact: qe-baseos-daemons
Docs Contact:

URL:

Progress?

- Simpler view into my infra
- Can update components independently
- No radical reshuffling required
- A place to host new components



Looking Ahead

- setenforce 1
- Nicer systemd integration
- Contained virt
 - Needs work upstream
 - openstack/kolla an option
- Side-by-side ceph/gluster

Looking Ahead, cont

- cleaner network setup
 - Openvswitch
 - oVirt / Neutron Integration
 - Looking to Atomic Enterprise
- Atomic hosts
- Smooth out deployment/automation
 - Ansible
 - AtomicApp / Nucleule

Questions?

- Or, ask me later:
 - @jasonbrooks
 - jbrooks@redhat.com
 - jbrooks on freenode
 - jebpages.com