Painting a Picture Of KVM Use-Cases In The Container World

Fabian Deutsch, Red Hat

KVM Forum, Edinburgh, 2018 (link)

(Image: vidalcuglietta.com)

redhat

Myself

fabiand@redhat.com
twitter.com/dummdida
github.com/fabiand
https://dummdida.tumblr.com

qemu + kvm = 💪

Versatile and battle proven building blocks.

Machine level abstraction.

Strong isolation.


Containers.
KVM?
(Photo: Itsuo Inouye/Associated Press)
redhat

Containers:

Focus on applications and user workflows.

Admins enjoy virtualization ~ Developers enjoy containers

redhat.

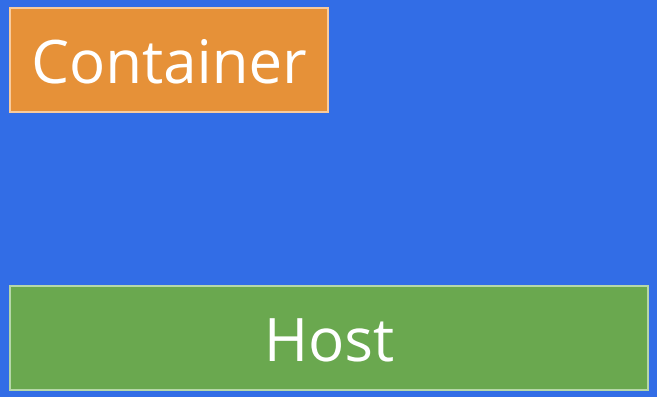
Solves Everything!

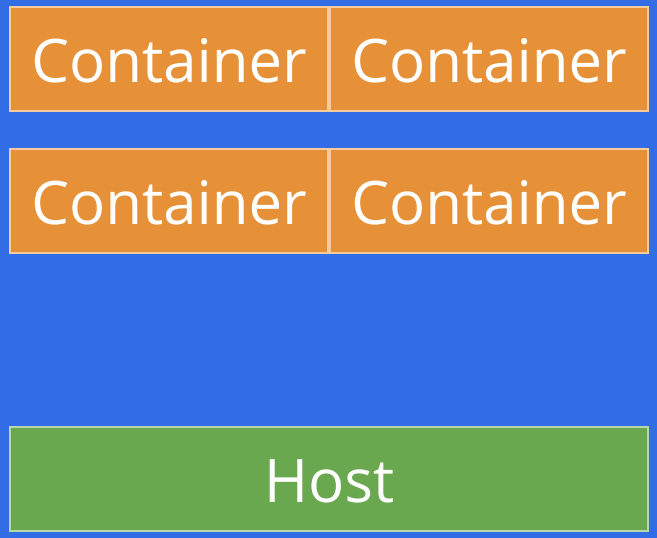
Brought to you by "new technology".

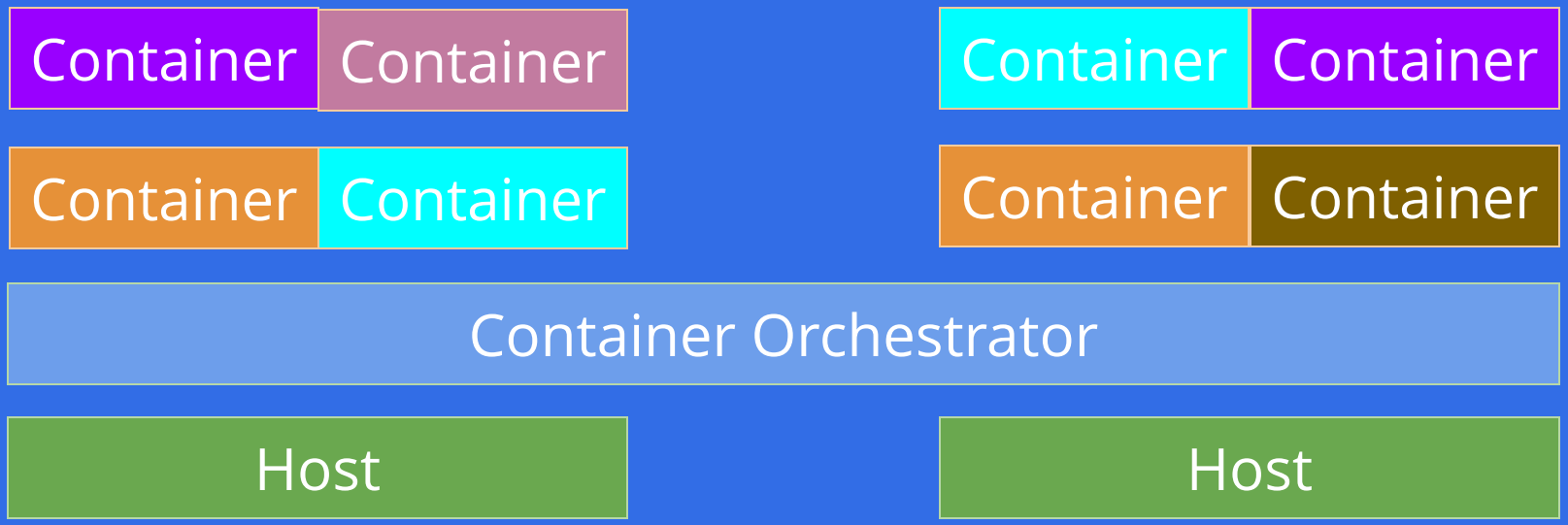
With a brand new - well known (ha!) - featureset ...

redhat


FFWD Today.
Aka today we know it better ...
(Photo: gsshow)
redhat.


Container
Host


Container
Container
Container
Container
Host


Container
Container
Container
Container
Container
Container
Container
Container
Container Orchestrator
Host
Host

Apache MESOS™

kubernetes

A container orchestrator.

redhat.

Containers in production environments have production critical problems.

"*Hum* Workloads can be insecure ..."

Building sources from github in a container, processing data retrieved from a webform in a container, ...

"**Erm** No - We can't containerize it."

Doing kernel module testing, my very old application, vendor appliances, ...

Kata Containers

Adds an **isolation** layer to containers
by using an optimized qemu and kvm.

The containers are running inside a VM.

(Merger of: frakti, hypersh, ClearContainers)

```yaml
apiVersion: v1
kind: Pod  # <---------------------------------- HERE (1)
metadata:
  name: nginx-untrusted
  annotations:
    io.kubernetes.cri.untrusted-workload: "true"  # <-- HERE (2)
spec:
  containers:
  - name: nginx
    image: nginx
```

nemu

A forked and **optimized** qemu and kvm
which can be used with Kata Containers or elsewhere.

Aims to be a modern hypervisor.

gVisor

Adds an **isolation** layer around containers
one option (beyond others) is to use kvm.

Proxying syscalls trhough a KVM sandbox.

runq

Adds an **isolation** layer around containers using stock qemu and kvm.

Favors simplicity over efficiency.

virtlet

Allows you to run **VM** appliances with a container API using qemu and kvm.

redhat

```yaml
apiVersion: v1
kind: Pod                                          # <-- HERE (1)
metadata:
  name: ubuntu-vm-rdb-block-pv
  annotations:
    kubernetes.io/target-runtime: virtlet.cloud  # <-- HERE (2)
    …
spec:
  containers:
  - name: ubuntu-vm
    image: virtlet.cloud/cloud-images.ubuntu.com/xenial/cloudimg-amd64-disk1.img
    …
```

KubeVirt

Allows you to run **VM** images
using libvirt, qemu, and kvm.

As close as you can get to a clustered libvirt.

```
apiVersion: kubevirt.io/v1alpha2
kind: VirtualMachineInstance  # <-- HERE (1)
…
spec:
  domain:  # <---------------------- HERE (2)
    cpu:
      cores: 2
    devices:
      disks:
      - disk:
          bus: virtio
        name: fedoracore1
    …
```

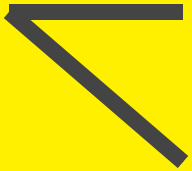

Container Isolation

gVisor

Kata Containers ➪ nemu

runq

Legacy VM Workloads



KubeVirt

virtlet

It's not just a one way ticket.

- virtfs / 9p (kata containers)
- firmware & devices (nemu)
- machine type discussion
- guest details in libosinfo (KubeVirt)
- ...

Take away.

- Virtualization is still here - one way or the other.

- Containers impact the KVM ecosystem

- New use-cases and shifted requirements

Questions?

fabiand@redhat.com
@dummdida

kubevirt.io

redhat.