



June 19-21, Seattle

dockercon 16

Docker for Mac and Windows: The Insider's Guide

Justin Cormack

Engineer, Docker

@justincormack





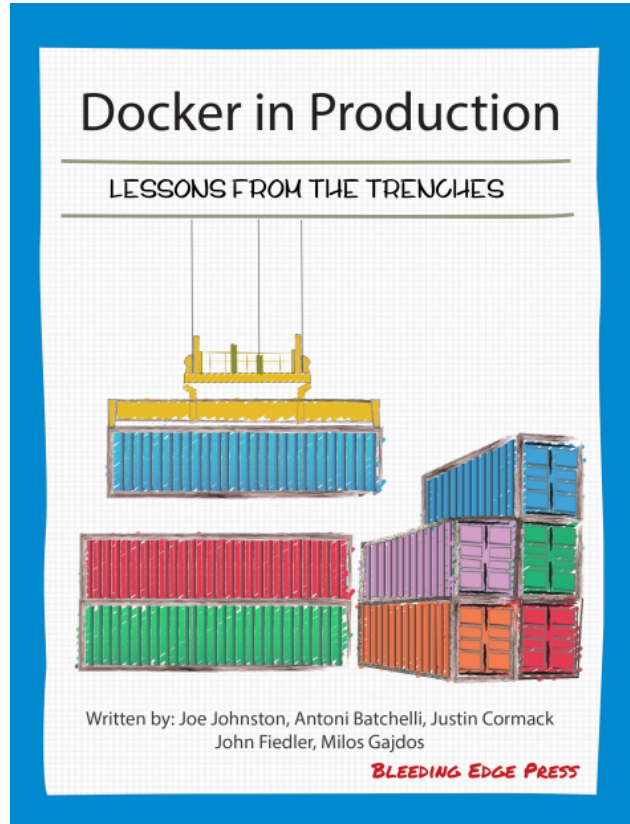
Justin Cormack

Cambridge, UK based engineer at Docker [@justincormack](https://twitter.com/justincormack)





Co-author of Docker in the Trenches: Successful Production Deployment

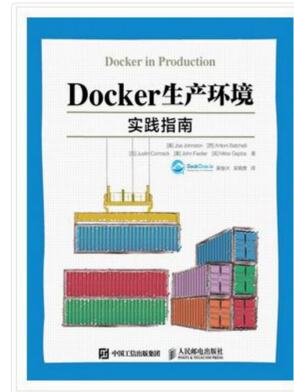


Bleeding Edge Press
@EdgePress



Following

Excited to see our #Docker book in China!
[@milosgajdos](#) [@JohnFiedler](#) [@tbatchelli](#)
[@justincormack](#) [@joejohnston](#)



RETWEETS
20

LIKES
22



8:19 PM - 3 Jun 2016



20



22



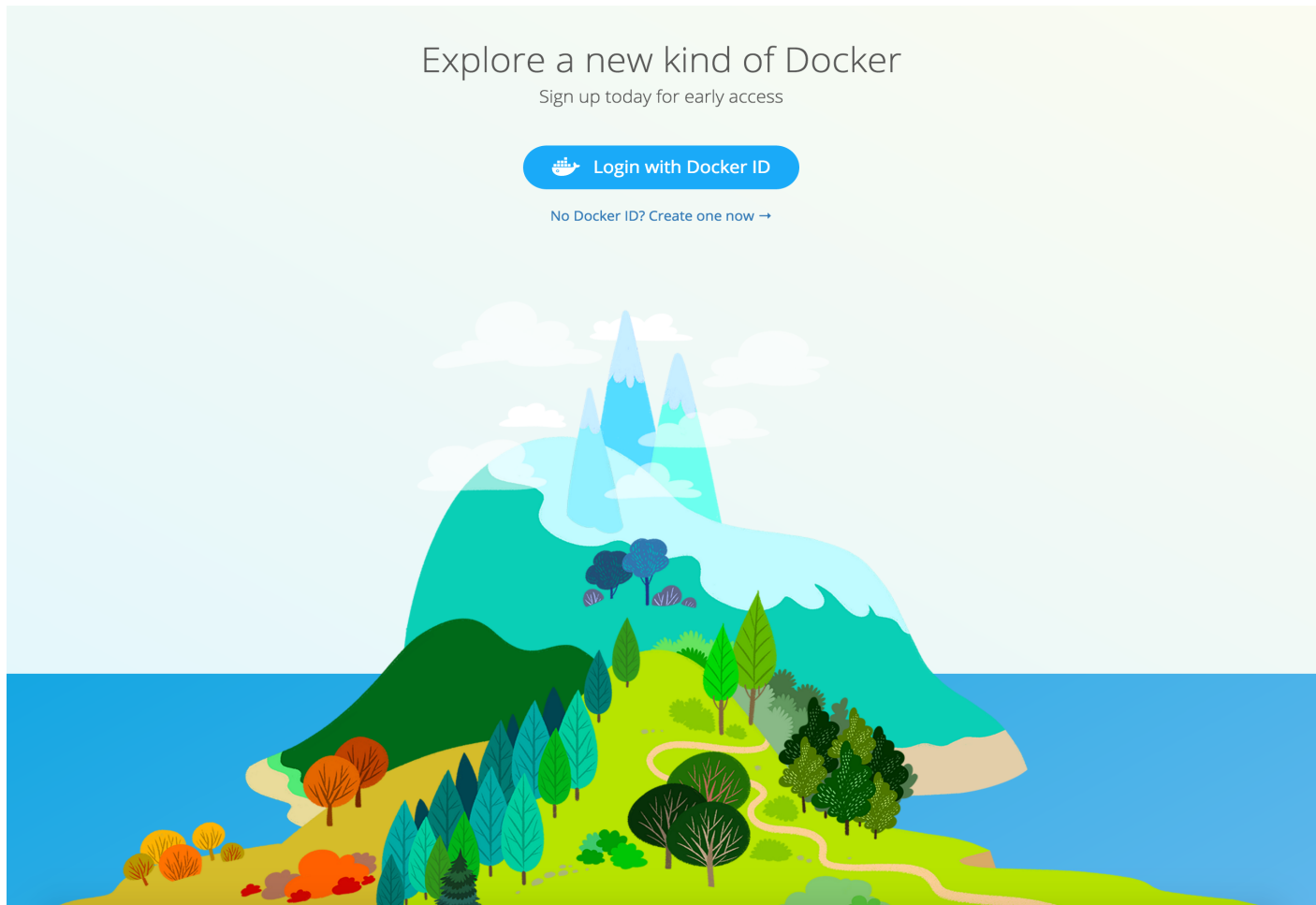


Explore a new kind of Docker

Sign up today for early access

 [Login with Docker ID](#)

[No Docker ID? Create one now →](#)





Make it native

- Install Docker and it is just there and working
- Make everything work just like a Linux install
- Make file notifications work
- No Virtualbox
- Get out of the way
- Make the powerful simple



People liked it!

- 30,000 signups on the first day!
- 70,000 by DockerCon
- The team building it used it from the start



Jeff Turcotte
@jeffturcotte



Follow

Docker for Mac has replaced VirtualBox/docker-machine for me. Great experience and it gets better with every update. Thanks [@docker!!](#)

RETWEETS
3

LIKES
2



8:40 PM - 10 Jun 2016



3

2



Kelsey Hightower
@kelseyhightower



Following

Docker hit a home run with Docker for Mac. Smooth install process and it just works.

```
Server Version: 1.11.1
Storage Driver: aufs
Root Dir: /var/lib/docker/aufs
Backing Filesystem: extfs
Dirs: 8
Dirs Supported: tmpfs
Logging Driver: json-file
Group Driver: cgroups
Plugins:
Volumes: local
Network: null host bridge
Kernel Version: 4.4.11-moby
Operating System: Alpine Linux v3.3
OSType: linux
Architecture: amd64
CPUs: 2
Total Memory: 1.955 GiB
Name: moby
ID: 09057C32-25A2-7E9B-6632-1A55-C75P-AD5N-LFSK-NDFY-S07N-TXWG
Docker Root Dir: /var/lib/docker
Debug mode (client): false
Debug mode (server): true
File Descriptors: 48
GoRoutines: 34
System Time: 2016-06-04T00:18:34.641885748Z
No Proxy: *local*, 109.224.16
Registry: https://index.docker.io/v1/
```

RETWEETS
94

LIKES
185



1:21 AM - 4 Jun 2016



94

185





Let's Go Inside



```
justin      43426  /Applications/Docker.app/Contents/MacOS/com.docker.hyperkit -A -m 2G -c 4 -u -s 0:0,hostbridge -s 31,lpc -s 2:0,virtio-vpnkit,u
justin      42732  /Applications/Docker.app/Contents/MacOS/com.docker.osx.hyperkit.linux
justin      42728  com.docker.backend
justin      23338  com.docker.db --url=file:///Users/justin/Library/Containers/com.docker.docker/Data/s40 --git /Users/justin/Library/Containers/c
justin      23337  /Applications/Docker.app/Contents/MacOS/com.docker.osx.hyperkit.linux -watchdog fd:0
justin      23334  /Applications/Docker.app/Contents/MacOS/com.docker.osx.hyperkit.linux -watchdog fd:0
root        3355   /Library/PrivilegedHelperTools/com.docker.vmnetd
justin      43684  /Applications/Docker.app/Contents/Resources/bin/com.docker.frontend {"action":"vmstateevent","args":{"vmstate":"running"}}
justin      43683  /Applications/Docker.app/Contents/Resources/bin/com.docker.frontend {"action":"vmstateevent","args":{"vmstate":"running"}}
justin      43425  /Applications/Docker.app/Contents/MacOS/com.docker.driver.amd64-linux -db /Users/justin/Library/Containers/com.docker.docker/Da
justin      43424  com.docker.driver.amd64-linux -db /Users/justin/Library/Containers/com.docker.docker/Data/s40 -osxfs-volume /Users/justin/Libra
justin      43423  com.docker.slirp --db /Users/justin/Library/Containers/com.docker.docker/Data/s40 --socket fd:3 --port-control fd:4 --vsock-pat
justin      43422  com.docker.osxfs --address fd:3 --connect /Users/justin/Library/Containers/com.docker.docker/Data/@connect --volume-control fd:
```




Hyperkit



Hyperkit: A toolkit for embedding hypervisor capabilities in your application

- Only used on the Mac – Windows uses Hyper-V
- Based on xhyve, which in turn is based on bhyve from FreeBSD
- A hypervisor is just a single process per emulated CPU core
- <https://github.com/docker/hyperkit>



Hyperkit

- Sparse block device, using qcow2 format
- Virtio devices: network, block, 9p, socket, rng
- Configure the amount of memory and CPU cores in the preferences



Datakit



Git database

- Can do really interesting stuff – it is git for datastructures
- Simple use cases here for storing configuration
- Yes there is a git tree at

```
~/Library/Containers/com.docker.docker/Data/database
```

- There is a filesystem view inside the VM using 9p
- <https://github.com/docker/datakit>



Plumbing



VSock and HVSock

- Socket families for communication with VMs, like Unix sockets
- VSocket originally developed by VMWare
- HVSocket developed by Microsoft, similar design, but different addressing
- New socket families so not well supported yet, eg no Go support have to use C bindings
- Communication not over network, so works even if network issues
- Used to transport the Docket socket to the VM not via https, and other uses.



Moby



Alpine Linux is a security-oriented, lightweight Linux distribution based on musl libc and busybox.



Stateless

- Alpine was designed to boot from init ramdisk
- Reboot and everything is back to the same state, except `/var` with Docker state on it.
- Configuration stored in the datakit database, currently mounted on `/Database`
- Phoenix like, just rebuild a new image, no upgrades
- Only runs Docker and supporting infrastructure.



No user serviceable parts?

- Designed to “just work” and keep Docker running.
- Database sets all the Docker and system config (UI coming soon)
- Use privileged containers or capabilities to do other host changes, persist with `--restart always`
- eg [install sysdig kernel module](#)
- root shell: `docker run -it --privileged --pid=host debian nsenter -t 1 -m -u -n -i sh`



Kernel

- Currently stable 4.4.x series kernel with aufs, vsock and hvsock patches
- No modules, everything built in for fast boot, but you can add modules
- Supports aufs and overlay storage drivers (overlay2 soon)
- Supports NFS, SMB, CRIU, many other things requested by users on the forum
- `binfmt_misc` set up so you can run arm and other binaries with qemu emulation
- config in `/proc/config.gz` patches in `/etc/kernel-patches`



Userspace

- Alpine 3.4
- Docker, statically linked build with seccomp.
- Reasonable set of utility programs, not just base system, eg Kubernetes had some requirements which were added
- You can install new packages, again they will not persist over a reboot, but we use this to debug Docker sometimes
- Some diagnostics code, setup code, file sharing, time synchronisation, ...
- Will be open sourced when a bit more stable, after GA.



VPNKit



**Come to Mindy Preston's talk tomorrow at
14:25 Ballroom 6C for more about VPNkit**



VPNkit

- <https://github.com/docker/vpnkit>
- Currently being used on Mac, being tested for Windows
- Take the ethernet traffic from the Linux VM
- Reconstruct it as application traffic on OSX
- So `socket(); ... listen() ... send()` in Linux is reconstructed from ethernet traffic as the same series of calls on OSX.
- No OS interfaces needed, complete control, appears just like another application so works with VPN, firewall.



VPNkit

- VPNkit uses the network stack from the Mirage unikernel
- Mirage is a set of low level system libraries
- Written in OCaml, a functional programming language
- Easy to repurpose for completely different use cases



OSXfs



OSXfs

- Currently only being used on the Mac, SMB used in Windows
- Uses FUSE on the Linux side to get the system calls
- Transports these over VSocket
- Converts into OSX filesystem calls.
- Listens for filesystem notifications in OSX and replays events in Linux.



OSXfs

- As you may see it is somewhat similar model to VPNkit but more complex
- Filesystems are more different than network stacks
- Filesystem notifications are quite different as not very standard
- eg OSX does not provide read notifications, but Linux does.
- Soon will be able to specify which parts of the filesystem to share.



Windows file sharing

- Uses a SMB mount from Linux to Windows
- SMB protocol supports filesystem notifications but Linux does not unfortunately
- Will have to solve this by adding notification support or porting OSXfs



User Interface

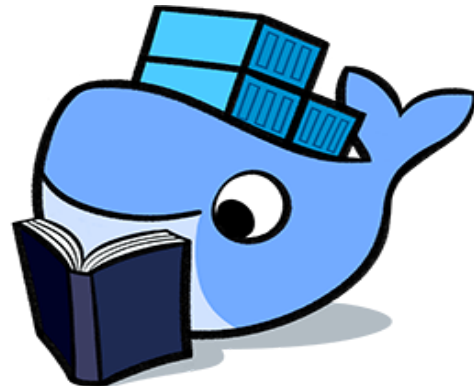
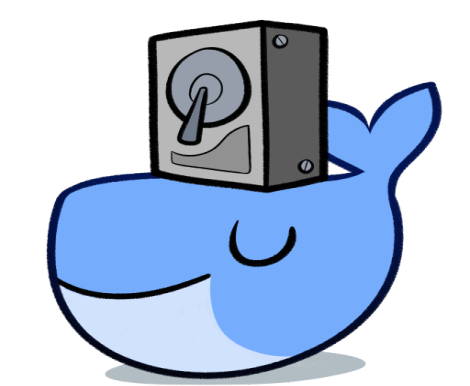
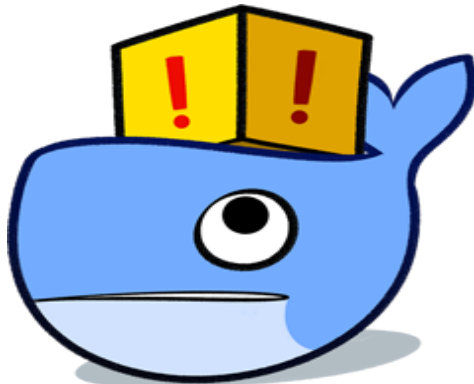
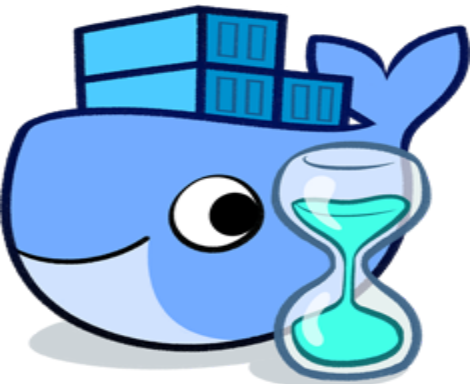
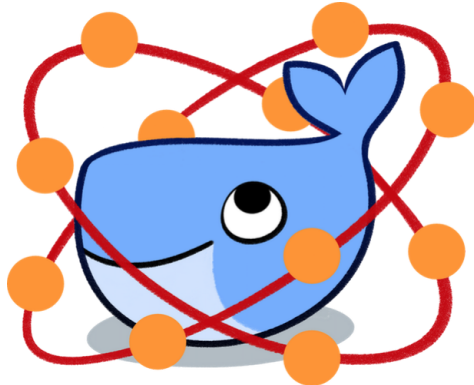


User interface

- Minimal at the moment but more plans for the future
- Native code: Swift on OSX, C# on Windows
- Communicates through database for configuration.
- Plans to integrate Kitematic or similar graphical interface later



Design







**Why doesn't *X*
work?**



--net=host

- With VPNkit, when you `--publish` a port in a container, we use a custom userland proxy
- This communicates through the `/port` 9p filesystem to the Mac
- The Mac opens a listening socket on the host, so you can connect to your container on `localhost:nnn`
- This is also used for `docker service` in Swarm mode
- But listening to a port with `--net=host` does not notify Docker so we cannot easily intercept it



I cannot connect to container ports

- On OSX there is no direct routing to the Linux VM, and so neither can you reach the machines on a bridge network
- If you want to debug, do it from another container, or `docker run --net=container:name`
- You need to do this for overlay networks anyway, so good idea to get used to this.



Unix socket between host and container doesn't work

- Well yes, they are different computers...
- But we do plan to make it work via osxfs and transparently proxying over VSocket for you
- In the mean time, make do with a TCP socket or run everything in containers.



Sound and Vision

- X Windows, RDP, VNC, or HTTP graphical interfaces can be made to work
- Audio should be possible too.
- Quite a few people want this, please help out!



Roadmap



Stability first, features second

- Stability and bug fixing priorities for GA
- People are using it every day for their work. Including us.
- Features driven by what you want
- More components will be open sourced.



Thanks



The team, many are here at DockerCon

Many many more people at Docker who contributed

A very long list of contributors to open source projects, especially bhyve, xhyve, Mirage, Alpine Linux



Questions?

- [@justincormack](https://twitter.com/justincormack)
- justin.cormack@docker.com
- `docker run -d -P justincormack/dockercon2016`

