Docker ade

Eine schlechte Container-Runtime in 10 Minuten

von neocturne • NooK 2023

Problemstellung

- Ein Docker-Image starten
- Ohne Container-Runtime
- Nur mit Debian-Bordmitteln

Die Lösung

- "Sillycon" - Silly Container Runtime
- ~75 Zeilen Shellscript

Speichern eines Docker-Images

```
$ docker image save -o image.tar my-image
```

```
$ tar tf image.tar
24bf8008…5cbe26ed/
24bf8008…5cbe26ed/VERSION
24bf8008…5cbe26ed/json
24bf8008…5cbe26ed/layer.tar
6a3253cd…c3558d57/
6a3253cd…c3558d57/VERSION
6a3253cd…c3558d57/json
6a3253cd…c3558d57/layer.tar
929c0ba6…b82b58f1.json
manifest.json
repositories
```

manifest.json

```
[
    {
        "Config": "929c0ba6…b82b58f1.json",
        "RepoTags": [
            "my-image:latest"
        ],
        "Layers": [
            "6a3253cd…c3558d57/layer.tar",
            "24bf8008…5cbe26ed/layer.tar"
        ]
    }
]
```

Namespaces

- Mount (Dateisystem/Mounts)
- Network (Netzwerk-Interfaces)
- PID (Prozess-IDs)
- IPC (Interprozess-Kommunikation)
- UTS (Hostname)
- User (User-IDs)
- (Cgroup) (Control Groups)
- (Time) (Monotone Systemzeit)

User-Namespaces

```
$ cat /etc/subuid
neocturne:100000:65536
$ cat /etc/subgid
neocturne:100000:65536
```

```
unshare \
    --map-root-user --map-users=auto --map-groups=auto
    --mount --net --ipc --pid --uts --fork \
    sh -ec '…'
```

```
ip a
1: lo: <LOOPBACK> mtu 65536 qdisc noop state DOWN …
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
```

```
mount -t tmpfs work "${WORKDIR}"
```

config.sh

```
LAYERS='
    6a3253cd…c3558d57/layer.tar
    24bf8008…5cbe26ed/layer.tar
'

CMD='bash'
ENV='PATH=/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:…'
HOSTNAME='sillycon'
```

Entpacken der Layer-Archive

```
i=0
lowerdir=
sep=
for layer in ${LAYERS}; do
    mkdir "${WORKDIR}/layer${i}"
    tar -C "${WORKDIR}/layer${i}" -xf "${DIR}/${layer}"
    lowerdir="${WORKDIR}/layer${i}${sep}${lowerdir}"
    sep=:
    i=$((i+1))
done
```

```
lowerdir=${WORKDIR}/layer1:${WORKDIR}/layer0
```

Overlay-Mount für rootfs

```
mkdir "${WORKDIR}/upper" "${WORKDIR}/work" "${WORKDIR}/root"

mount -t overlay -o\
lowerdir="${lowerdir}",\
upperdir="${WORKDIR}/upper",\
workdir="${WORKDIR}/work" \
    root "${WORKDIR}/root"

cd "${WORKDIR}/root"
```

Essentielle Mounts

```
mount -t proc -o nosuid,nodev,noexec proc proc
mount -t sysfs -o nosuid,nodev,noexec sys sys
mount -t tmpfs -o nosuid,nodev,mode=755 run run
mount -t tmpfs -o nosuid,nodev tmp tmp
```

Setup für /dev: Devices

```
mount -t tmpfs -o nosuid dev dev

for file in null zero full random urandom tty; do
    touch "dev/${file}"
    mount --bind "/dev/${file}" "dev/${file}"
done
```

Setup für /dev: Mounts

```
# Pseudoterminals
mkdir dev/pts
mount -t devpts -o nosuid,noexec devpts dev/pts

# Shared Memory
mkdir dev/shm
mount -t tmpfs -o nosuid,nodev tmpfs dev/shm
```

Setup für /dev: Symlinks

```
ln -s /proc/self/fd dev/fd
ln -s /proc/self/fd/0 dev/stdin
ln -s /proc/self/fd/1 dev/stdout
ln -s /proc/self/fd/2 dev/stderr
ln -s pts/ptmx dev/ptmx
```

Hostname

```
hostname "${HOSTNAME}"
```

```
pivot_root . tmp
cd /

exec /usr/bin/env -i - "${ENV}" /bin/sh -ec "
    umount -l /tmp
    exec ${CMD}
"
```

```
$ ./sillycon
root@sillycon:/#
```

```
# findmnt
TARGET                SOURCE         FSTYPE   OPTIONS
/                     root           overlay  …
|-/proc               proc           proc     …
|-/sys                sys            sysfs    …
|-/run                run            tmpfs    …
|-/tmp                tmp            tmpfs    …
`-/dev                dev            tmpfs    …
  |-/dev/null         dev[/null]     devtmpfs …
  |-/dev/zero         dev[/zero]     devtmpfs …
  |-/dev/full         dev[/full]     devtmpfs …
  |-/dev/random       dev[/random]   devtmpfs …
  |-/dev/urandom      dev[/urandom]  devtmpfs …
  |-/dev/tty          dev[/tty]      devtmpfs …
  |-/dev/pts          devpts         devpts   …
  `-/dev/shm          tmpfs          tmpfs    …
```

neocturne

https://chaos.social/@neocturne

https://github.com/neocturne

Slides und Code:

https://github.com/neocturne/sillycon