rkt Deep Dive

Interesting technical aspects of rkt

Iago López Galeiras, Kinvolk
Alban Créquy, Kinvolk

About us

- Living in Berlin
- Kinvolk co-founders
- Worked a lot on rkt and rktlet
- Currently maintainers of rkt and rktlet

```
* github: https://github.com/iaguis
          https://github.com/alban
* twitter: @iaguis
           @albcr
```

Cgroup mounts

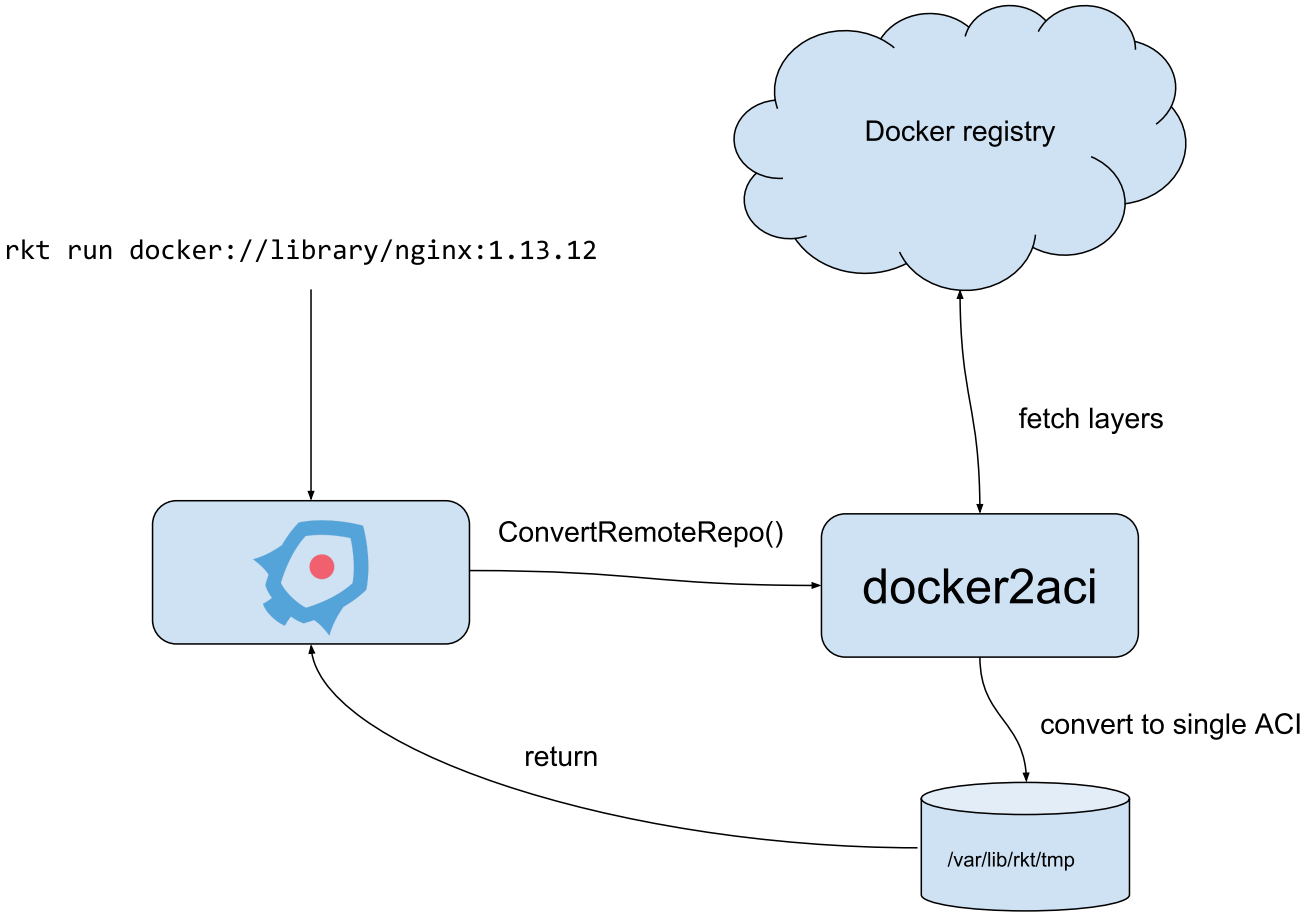
Now (v1.30.0)

```
/sys/fs/cgroup/cpu (RO)
└── machine.slice
    └── machine-rkt\x2da4d0d08e\x2d4944\x2d428c\x2daf55\x2d1d1ab1782560.scope (RW)
        └── system.slice
            └── alpine.service
                ├── cpu.cfs_period_us
                ├── cpu.cfs_quota_us
                ├── cgroup.procs
                └── ...
```

Before v1.23.0

```
/sys/fs/cgroup/cpu (RO)
└── machine.slice
    └── machine-rkt\x2da4d0d08e\x2d4944\x2d428c\x2daf55\x2d1d1ab1782560.scope
        └── system.slice
            └── alpine.service
                ├── cpu.shares (RW)
                ├── cpu.cfs_quota_us (RW)
                ├── cgroup.procs (RW)
                ├── cpu.stat
                └── ...
```

Running OCI images



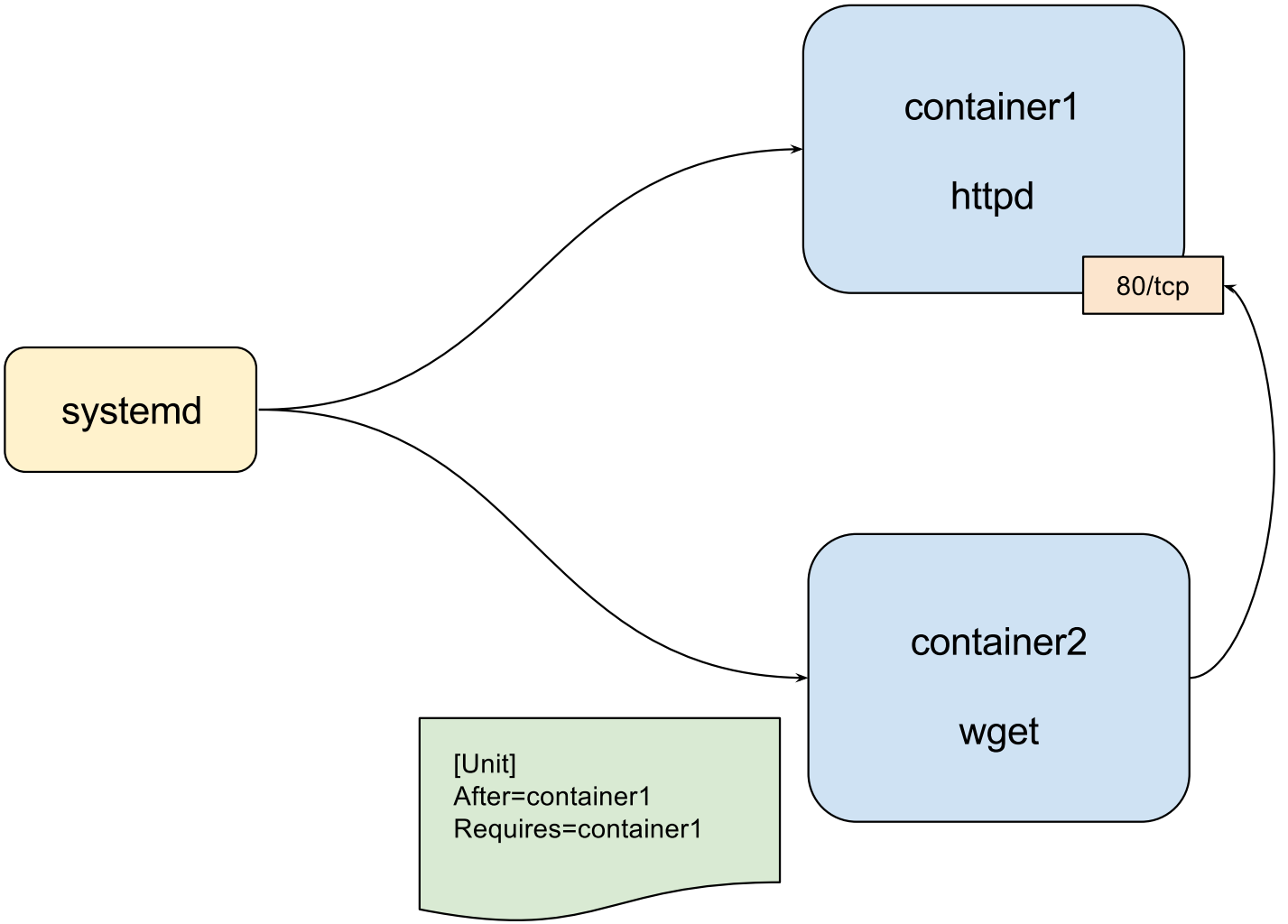
CLI flags -> systemd Directives

Flag	systemd Directive
`--caps-{remove,retain}=`	`CapabilityBoundingSet=`
`--cpu=`	`CPUQuota=`
`--cpu-shares=`	`CPUShares=`
`--exec=`	`ExecStart=`
`--group=`	`Group=`
`--memory=`	`MemoryLimit=`

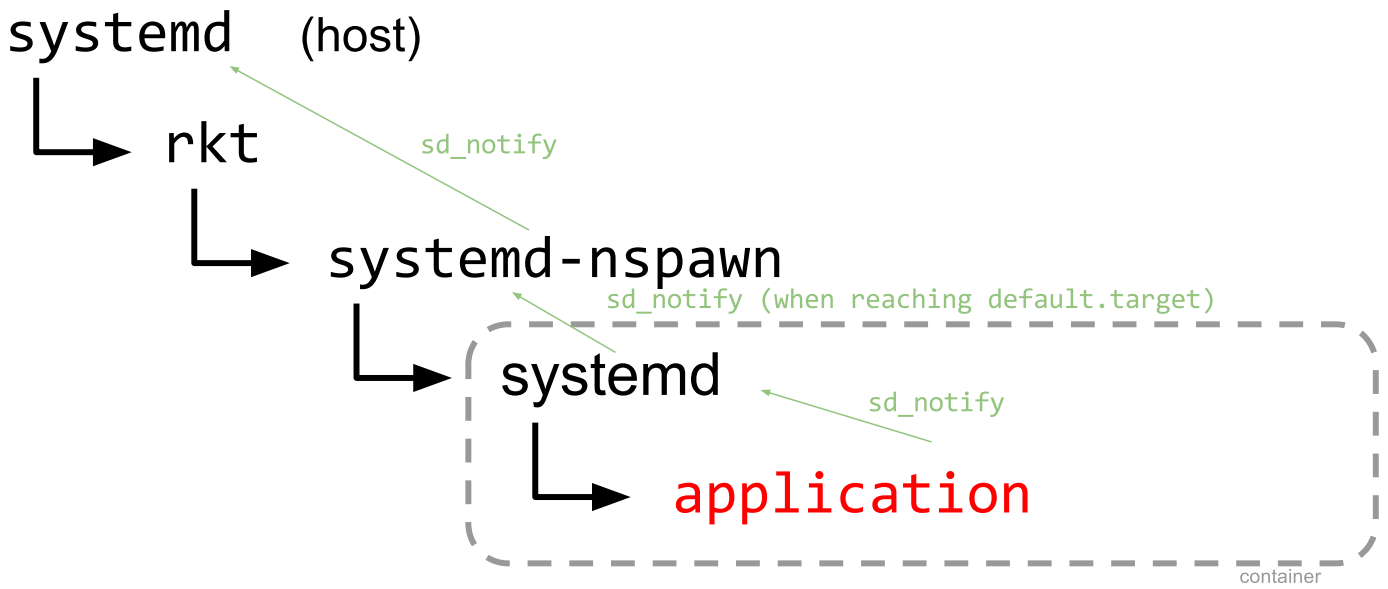
CLI flags -> systemd Directives (2)

Flag	systemd Directive
`--oom-score-adj=`	`OOMScoreAdjust=`
`--readonly-rootfs=`	`ReadOnlyDirectories=`
`--seccomp=`	`SystemCallFilter=`
`--private-users`	`--private-users=` (nspawn)
`--supplementary-gids=`	`SupplementaryGroups=`
`--user=`	`User=`
`--working-dir=`	`WorkingDirectory=`

SD Notify usecase



SD Notify propagation



Thank You!

Questions?