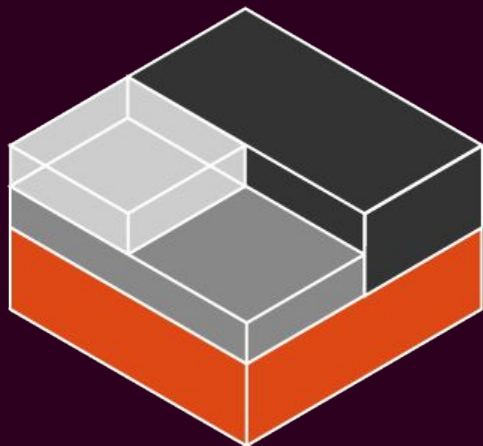


mixing cgroupfs v1 & v2

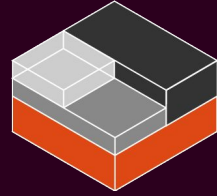
ubuntu 

Linux Piter 2017
Санкт-Петербург, Россия



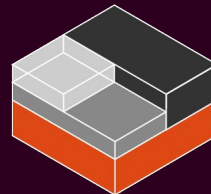
Christian Brauner
Software Engineer, Canonical Ltd.
christian.brauner@ubuntu.com
[@brauner](https://www.instagram.com/brauner)
<https://cbrauner.wordpress.com>

cgroups



- Pseudo kernel filesystem following the vfs
- Collection of processes
- Resource management and tracking

cgroup v1



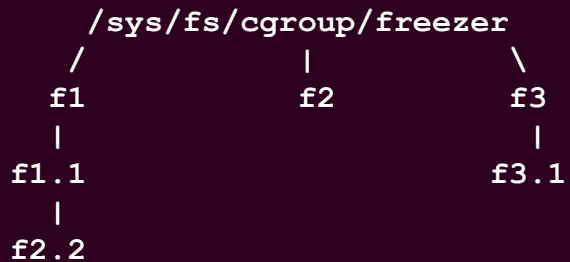
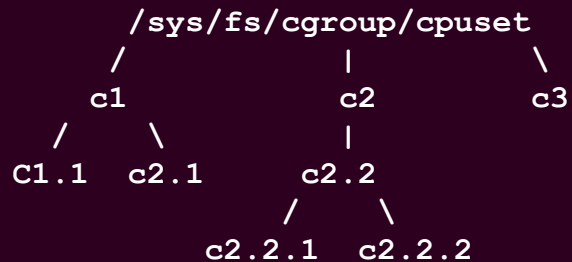
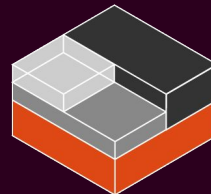
```
chb@conventiont|~
```

```
> ls -al /sys/fs/cgroup/
```

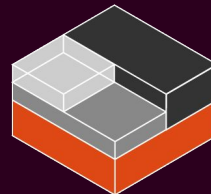
```
total 0
```

```
drwxr-xr-x 14 root root 360 May  8 14:24 .
drwxr-xr-x  9 root root   0 May  8 14:24 ..
dr-xr-xr-x  6 root root   0 May  8 14:24 blkio
lrwxrwxrwx  1 root root  11 May  8 14:24 cpu -> cpu,cpuacct
lrwxrwxrwx  1 root root  11 May  8 14:24 cpuacct -> cpu,cpuacct
dr-xr-xr-x  6 root root   0 May  8 14:24 cpu,cpuacct
dr-xr-xr-x  3 root root   0 May  8 14:24 cpuset
dr-xr-xr-x  6 root root   0 May  8 14:24 devices
dr-xr-xr-x  4 root root   0 May  8 14:24 freezer
dr-xr-xr-x  3 root root   0 May  8 14:24 hugetlb
dr-xr-xr-x  7 root root   0 May  8 14:24 memory
lrwxrwxrwx  1 root root  16 May  8 14:24 net_cls -> net_cls,net_prio
dr-xr-xr-x  3 root root   0 May  8 14:24 net_cls,net_prio
lrwxrwxrwx  1 root root  16 May  8 14:24 net_prio -> net_cls,net_prio
dr-xr-xr-x  3 root root   0 May  8 14:24 perf_event
dr-xr-xr-x  6 root root   0 May  8 14:24 pids
dr-xr-xr-x  7 root root   0 May  8 14:24 systemd
```

cgroup v1

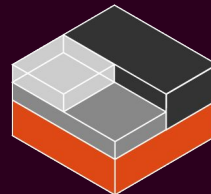


cgroup v2



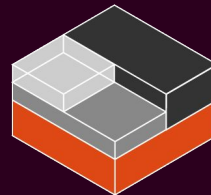
```
chb@conventiont|~  
> ls -al /sys/fs/cgroup/  
total 0  
-r--r--r--  1 root root    0 Oct 25 10:57 cgroup.controllers  
-rw-r--r--  1 root root    0 Oct 25 10:57 cgroup.max.depth  
-rw-r--r--  1 root root    0 Oct 25 10:57 cgroup.max.descendants  
-rw-r--r--  1 root root    0 Oct 24 23:54 cgroup.procs  
-r--r--r--  1 root root    0 Oct 25 10:57 cgroup.stat  
-rw-r--r--  1 root root    0 Oct 25 10:57 cgroup.subtree_control  
-rw-r--r--  1 root root    0 Oct 25 10:57 cgroup.threads
```

legacy != unified



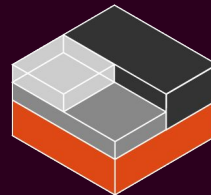
(1) there can only be one

legacy != unified



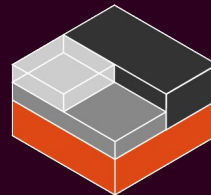
(2) internal process constraint

legacy != unified



(3) resource distribution

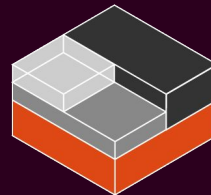
legacy != unified



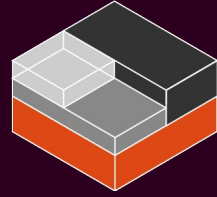
(4) delegation containment



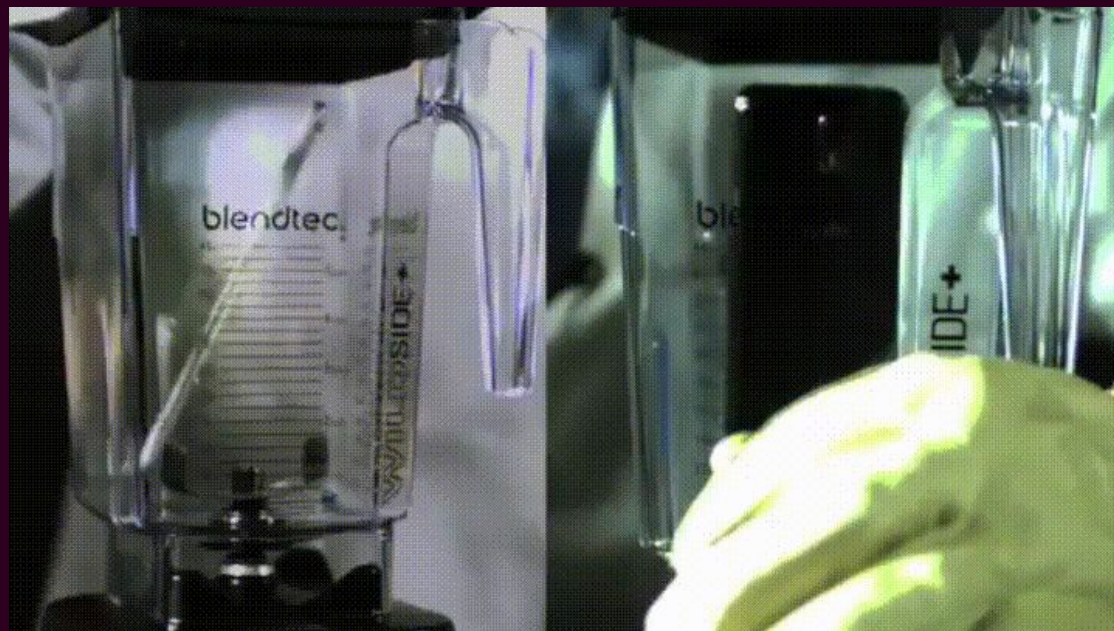
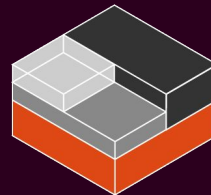
userspace



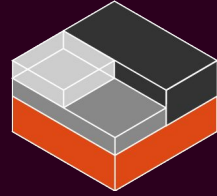
container runtimes



mixing legacy & unified



mixing legacy & unified



creating & populating cgroups

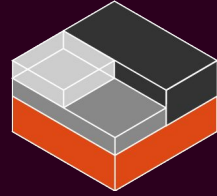
1. the easy cases

- *privileged containers*
- *unprivileged containers started by root*
- *unprivileged containers managed by a privileged daemon*

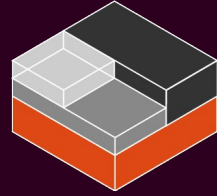
2. the hard cases

- *lacking privilege*

mixing legacy & unified

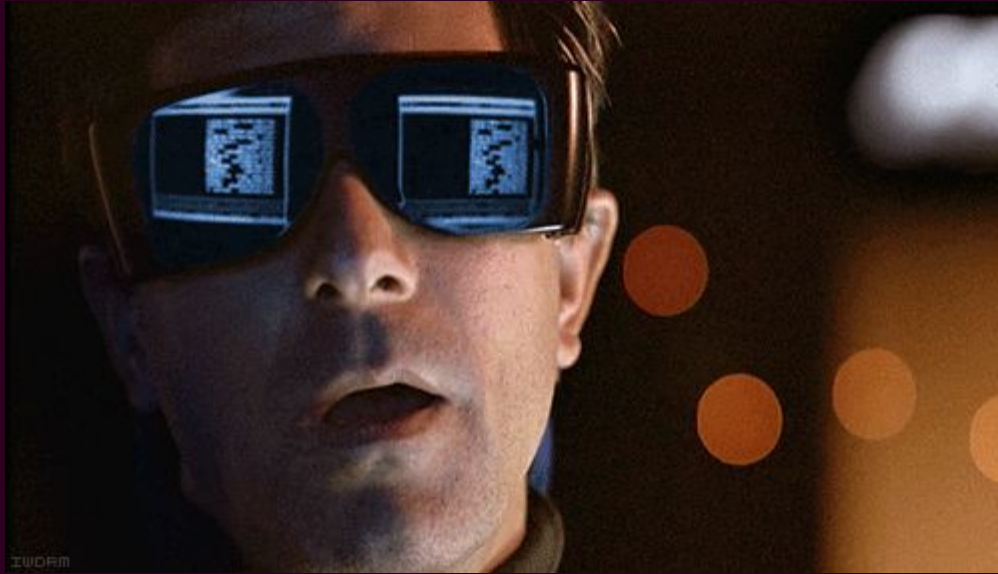
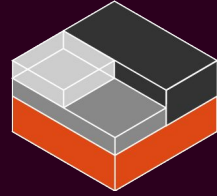


mixing legacy & unified



1. `chown $(id -u):$(id -g) /sys/fs/cgroup/c1/c2`
2. `mkdir /sys/fs/cgroup/c1/c2/c3 /sys/fs/cgroup/c1/c2/runtime`
3. `chown $(id -u):$(id -g) /sys/fs/cgroup/c1/c2/cgroup.procs`
4. `while read line; do`
 `echo "${line}" > /sys/fs/cgroup/c1/c2/c3/cgroup.procs`
`done < /sys/fs/cgroup/c1/c2/cgroup.procs`
5. `echo +io +memory +pids > /sys/fs/cgroup/c1/c2/cgroup.subtree_control`
6. `echo +io +memory +pids > /sys/fs/cgroup/c1/c2/runtime/cgroup.subtree_control`
7. `mkdir /sys/fs/cgroup/c1/c2/runtime/container1`
8. migrate your container-setup process from "c3" into "container1"
9. spawn your (init) binary and hope all is well

mixing legacy & unified



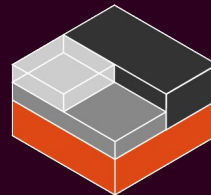
LXCFS to the rescue

<https://github.com/lxc/lxcfs>

FUSE filesystem for LXC

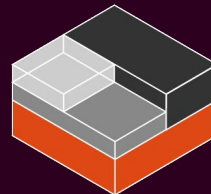
provides an emulated
/proc and /sys/fs/cgroup
folder for the containers

mixing legacy & unified



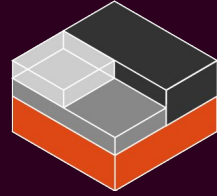
unified on unified

mixing legacy & unified



unified on legacy

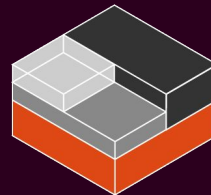
mixing legacy & unified



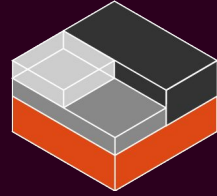
legacy on unified

```
/* Hierarchies may only be created in the initial
 * cgroup namespace. */
if (ns != &init_cgroup_ns) {
    ret = -EPERM;
    goto out_unlock;
}
```

translating resource limits



Christian Brauner
Software Engineer, Canonical Ltd.
christian.brauner@ubuntu.com
[@brauner](https://brauner)
<https://cbrauner.wordpress.com>



Questions?

<https://linuxcontainers.org>
<https://github.com/lxc/lxc>
<https://github.com/lxc/lxcfs>
<https://github.com/lxc/lxd>