

LXC Issues

lxc_init: Failed to run lxc.hook.pre-start for container

After upgrade guest system from Debian 12.0 to 12.1.

```
lxc-start -lDEBUG -o error.log -F -n <ContainerID>
```

```
unsupported debian version '12.1'
```

PVE (pve-container) needs upgrade.

apply caps: operation not permitted: unknown.

BalenaOS Build inside privileged LXC:

```
docker: Error response from daemon: OCI runtime create failed:
container_linux.go:380: starting container process caused: apply caps:
operation not permitted: unknown.
```

Solution (not secure!):

```
lxc.apparmor.profile: unconfined
lxc.cgroup.devices.allow: a
lxc.cap.drop:
```

Source: <https://danthesalmon.com/running-docker-on-proxmox/>

Slow login into container

see below

Failed at step NAMESPACE spawning /lib/systemd/systemd-logind: Permission denied

Debian Bullseye in unprivileged container:

```
systemd[579]: systemd-logind.service: Failed to set up mount namespacing:
/run/systemd/unit-root/proc: Permission denied
systemd[579]: systemd-logind.service: Failed at step NAMESPACE spawning
/lib/systemd/systemd-logind: Permission denied
```

SOLUTION: enable container nesting.

cannot stop container

Container works, responds to pings but it is not possible to SSH or attach.

Normal commands to stop or reboot doesn't help (even `lxc-stop -k`).

CAUSE: Container was freezed for snapshot. All processess are in 'D' state. Cannot be killed.

SOLUTION:

```
echo THAWED > /sys/fs/cgroup/freezer/lxc/200/freezer.state
```

Info:

- [Freezer subsystem](#)
- <https://www.kernel.org/doc/html/latest/admin-guide/cgroup-v1/freezer-subsystem.html>
- [Freezing of tasks](#)

Investigation

So killing container is solution:

```
pstree -p
      |
      |_lxc-start(3747487)---systemd(3747514)---agetty(3748048)
      |                                     |_agetty(3748049)
kill -9 3747514
```

Now it is not possible to start LXC container again. Debugging:

```
lxc-start -o lxc-start.log -lDEBUG -F -n 200
cat lxc-start.log

lxc-start 200 20210325085035.665 INFO      conf - conf.c:run_script_argv:331
- Executing script "/usr/share/lxc/hooks/lxc-pve-prestart-hook" for
container "200", config section "lxc"
lxc-start 200 20210325085036.126 DEBUG    conf - conf.c:run_buffer:303 -
Script exec /usr/share/lxc/hooks/lxc-pve-prestart-hook 200 lxc pre-start
produced output: failed to remove directory
'/sys/fs/cgroup/net_cls/lxc/200/ns': Device or resource busy
```

Reason: after killing container systemd, orphaned cgroups left.

```
find /sys/fs/cgroup/*/lxc/200 -depth -type d -print -delete

# Lot of errors:
```

```
find: cannot delete '/sys/fs/cgroup/freezer/lxc/200': Device or resource busy
```

All processes from container 200 are in 'D' state Uninterruptible Sleep

```
ps axl | awk '$10 ~ /D/'
```

```
echo w > /proc/sysrq_trigger
```

```
[587314.999001] smbd          D      0 1181293  42630 0x00004184
[587314.999002] Call Trace:
[587314.999004]  __schedule+0x2e6/0x6f0
[587314.999005]  schedule+0x33/0xa0
[587314.999007]  __refrigerator+0x44/0x160
[587314.999009]  get_signal+0x814/0x850
[587314.999011]  do_signal+0x34/0x6e0
[587314.999013]  ? wait_woken+0x80/0x80
[587314.999014]  ? __audit_syscall_exit+0x236/0x290
[587314.999016]  exit_to_usermode_loop+0x90/0x130
[587314.999018]  do_syscall_64+0x160/0x190
[587314.999020]  entry_SYSCALL_64_after_hwframe+0x44/0xa9
```

So it looks like whole container cgroup was freezed for snapshot and problem happens.

nested docker in cpulimit

Gitlab runner fails to start docker executor:

```
ERROR: Job failed (system failure): prepare environment: Error response from daemon: OCI runtime create failed: container_linux.go:367: starting container process caused: process_linux.go:495: container init caused: process_linux.go:458: setting cgroup config for procHooks process caused: failed to write "2400000": write /sys/fs/cgroup/cpu,cpuacct/docker/af4fd93c304a3edc9edb85da6f7a7f9ec85a15262db37393a22141686647d060/cpu.cfs_quota_us: invalid argument: unknown (exec.go:57:0s). Check https://docs.gitlab.com/runner/shells/index.html#shell-profile-loading for more information
```

Reason: cpulimit was set on container in PVE **Reproduction:**

```
# works:
docker run -it busybox

# problem:
docker run --cpuset-cpus='0' --cpus=1 --cpu-shares=256 -it busybox
```

Failed to set up mount namespaces: Permission denied

Inside LXC CT: Long ssh login delay, lots of errors in journal

```
# journalctl
gru 28 08:18:48 hostname systemd[860]: systemd-logind.service: Failed to set
up mount namespaces: /run/systemd/unit-root/proc: Permission denied
gru 28 08:18:54 hostname nmbd[106]: This response was from IP
192.168.12.45, reporting an IP address of 172.16.0.131.
gru 28 08:19:02 hostname systemd[866]: systemd-logind.service: Failed to set
up mount namespaces: /run/systemd/unit-root/proc: Permission denied
gru 28 08:19:10 hostname sshd[783]: pam_systemd(sshd:session): Failed to
create session: Failed to activate service 'org.freedesktop.login1': timed
out (service_start_timeout=25000ms)
gru 28 08:19:14 hostname systemd[877]: systemd-logind.service: Failed to set
up mount namespaces: /run/systemd/unit-root/proc: Permission denied
```

From:

<https://niziak.spox.org/wiki/> - niziak.spox.org

Permanent link:

<https://niziak.spox.org/wiki/vm:proxmox:lxc:issues>

Last update: **2023/07/28 12:13**

