LXC

chenshh

LXC

Linux Container

Jails on Linux

Why Containers

VM is expensive on calculation resources

Using containers have less overheads

Creating LXCs

```
# lxc-create -n playtime -t /usr/share/lxc/templates/lxc-archlinux
DONE!!!
```

You need network!

Create a bridge

```
# cd /etc/netctl
# cp examples/bridge ./
# vim bridge
```

Assign a IP for it.

```
# vim /var/lib/lxc/playtime/config
```

Delete “lxc.network.type = empty”

Add these to the head

```
  lxc.network.type = veth
     lxc.network.link = br0
     lxc.network.ipv4 = 192.168.X.X/24
     lxc.network.name = eth1
     lxc.network.flags = up
```

Fire it up!

```
# lxc-start -n playtime
# lxc-attach -n playtime
You will login as root in the lxc
```

How does it work?

Namespaces - isolation

cgroup - resource management

apparmor - permision management

lxc-checkconfig

```
# lxc-checkconfig
 --- Control groups ---
   Cgroup: enabled
   Cgroup clone_children flag: enabled
   Cgroup device: enabled
   Cgroup sched: enabled
   Cgroup cpu account: enabled
   Cgroup memory controller: enabled
   Cgroup cpuset: enabled

   --- Misc ---
   Veth pair device: enabled
   Macvlan: enabled
   Vlan: enabled
   Bridges: enabled
   Advanced netfilter: enabled
   CONFIG_NF_NAT_IPV4: enabled
   CONFIG_NF_NAT_IPV6: enabled
   CONFIG_IP_NF_TARGET_MASQUERADE: enabled
   CONFIG_IP6_NF_TARGET_MASQUERADE: enabled
   CONFIG_NETFILTER_XT_TARGET_CHECKSUM: enabled
   FUSE (for use with lxcfs): enabled
```

Namespaces

UTS namespace

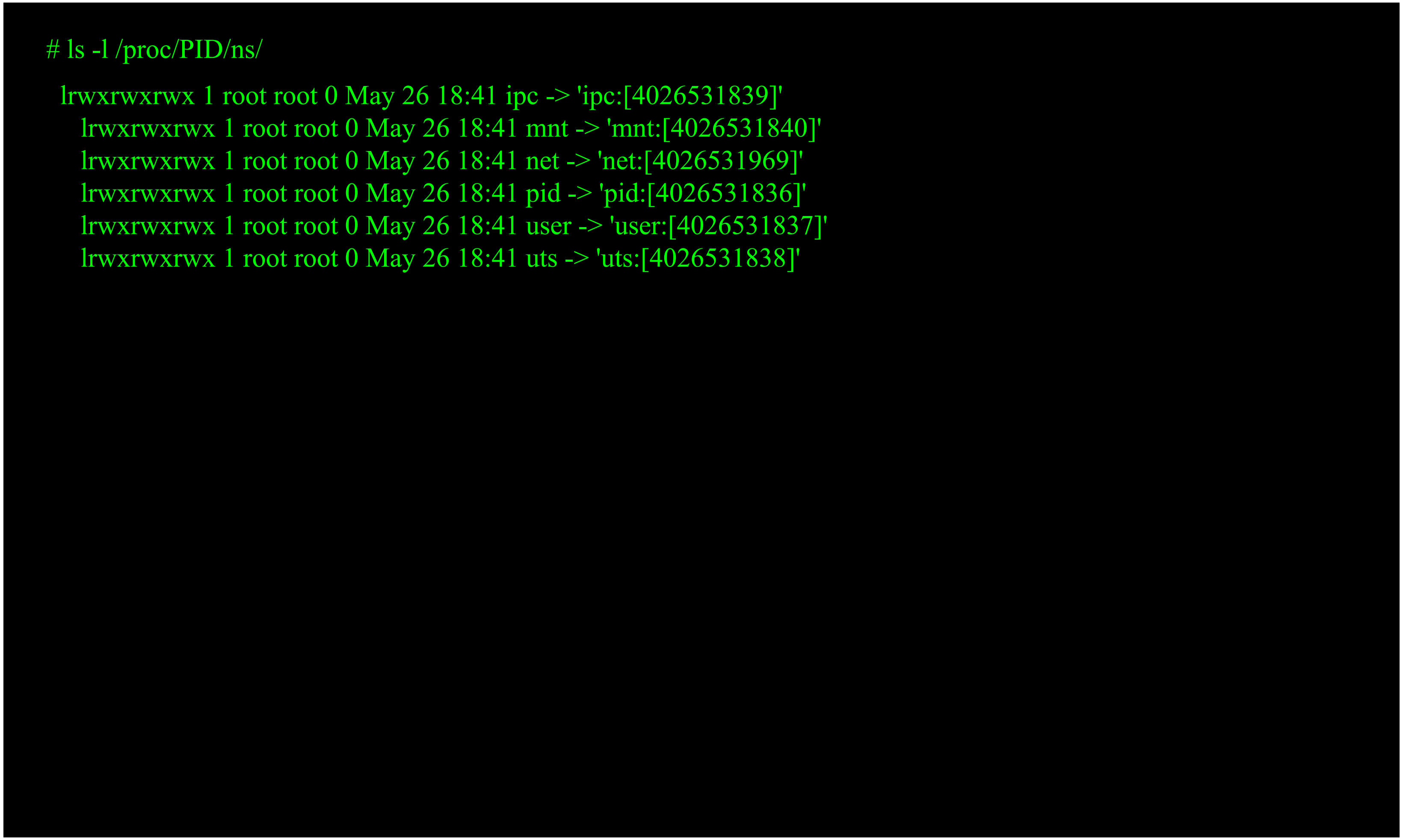
IPC namespace

mount namespace

PID namespace

Username space

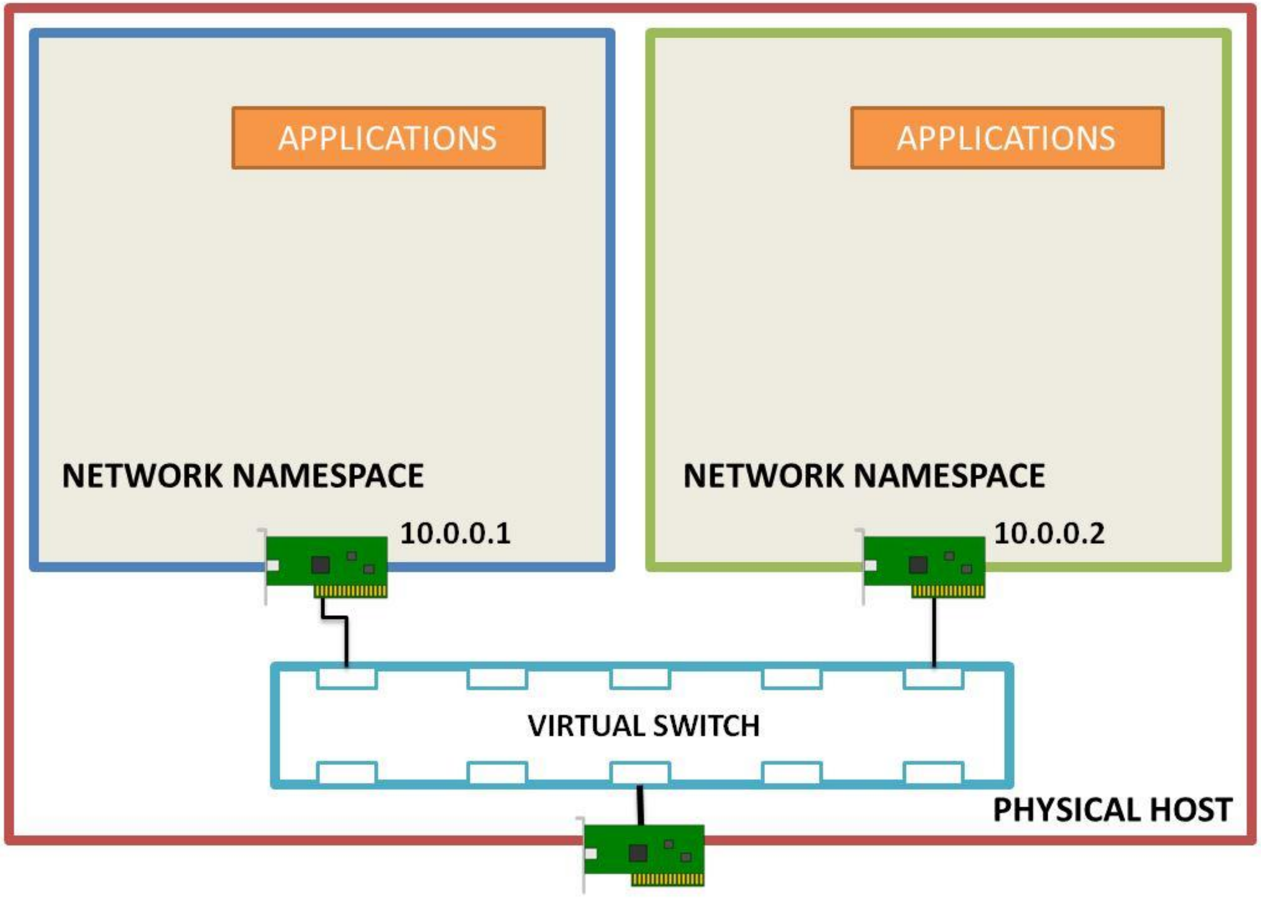
Check namespace



```
# ls -l /proc/PID/ns/
 lrwxrwxrwx 1 root root 0 May 26 18:41 ipc -> 'ipc:[4026531839]'
   lrwxrwxrwx 1 root root 0 May 26 18:41 mnt -> 'mnt:[4026531840]'
   lrwxrwxrwx 1 root root 0 May 26 18:41 net -> 'net:[4026531969]'
   lrwxrwxrwx 1 root root 0 May 26 18:41 pid -> 'pid:[4026531836]'
   lrwxrwxrwx 1 root root 0 May 26 18:41 user -> 'user:[4026531837]'
   lrwxrwxrwx 1 root root 0 May 26 18:41 uts -> 'uts:[4026531838]'
```

Network namespace

IN CASE OF NETWORK NAMESPACES



Username namespace

Subuid , subgid

Why username space

underprivileged LXC!

LXC can be started by non root

It means if some process escaped LXC , it still can get root permissions!

Configuring username space

1. rebuild the kernel (If your OS not ubuntu)
2. edit /etc/suduid /etc/subgid

Using underprivileged LXC

1. using lxc-download template
2. use root to build rootfs and convert it to underprvileged rootfs

Why need lxc-download?

When creating rootfs we use package managers

ex. pacstrap

They often need loop mounting.

But normal user doesn’t have mount permissions.

We still need more!

cgroup issue

In linux , there is no cgroup namespace

and procfs cant be mounted by normal user

helper:

cgmanager , lxcfs

lxcfs

LXCFS is a simple userspace filesystem designed to work around some current limitations of the Linux kernel.

Specifically, it's providing two main things

- A set of files which can be bind-mounted over their /proc originals
- to provide CGroup-aware values.
- A cgroupfs-like tree which is container aware.

The code is pretty simple, written in C using libfuse and glib.

cgmanager

CGManager is a central privileged daemon that manages all your cgroups for you

through a simple D-Bus API. It's designed to work with nested LXC containers

as well as accepting unprivileged requests including resolving user namespaces UIDs/GIDs.