Unprivileged Containers

Jess Frazelle, @jessfraz

How do containers help security?

Containers are not going to be the answer to preventing your application from being compromised, but they *can limit* the damage from a compromise.

How do containers help security?

The world an attacker might see from inside a very strict container with custom AppArmor/Seccomp profiles greatly differs than that without the use of containers.

Sandboxes Today

Chrome

- Seccomp
- Namespaces
- Apparmor
- NOT RUN AS ROOT

Containers today

- Namespaces
- Apparmor
- Selinux
- Capabilities Limiting
- Cgroups
- Run as root :(

How can we get to sandboxes with containers?

Back to the Basics

A “container” is what we have come to call a group of namespaces and control groups applied to a process.

Control Groups (cgroups)

Limit what the process can use. Resource metering and limiting.

Types: memory, CPU, blkio, network, device, pid..

PID Cgroup

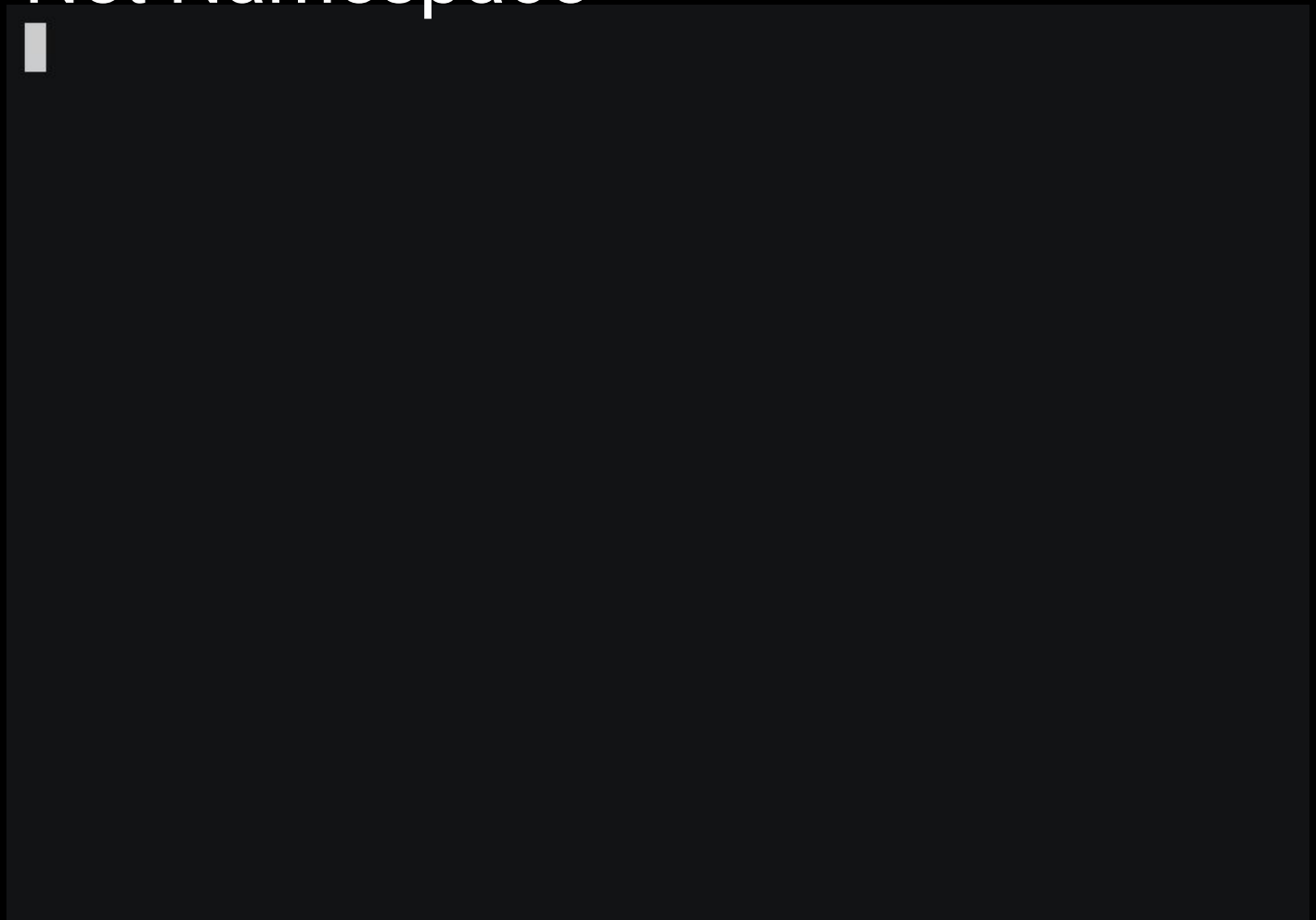
Namespaces

Limit what the process sees.

Types: pid, net, mnt, uts, ipc, user

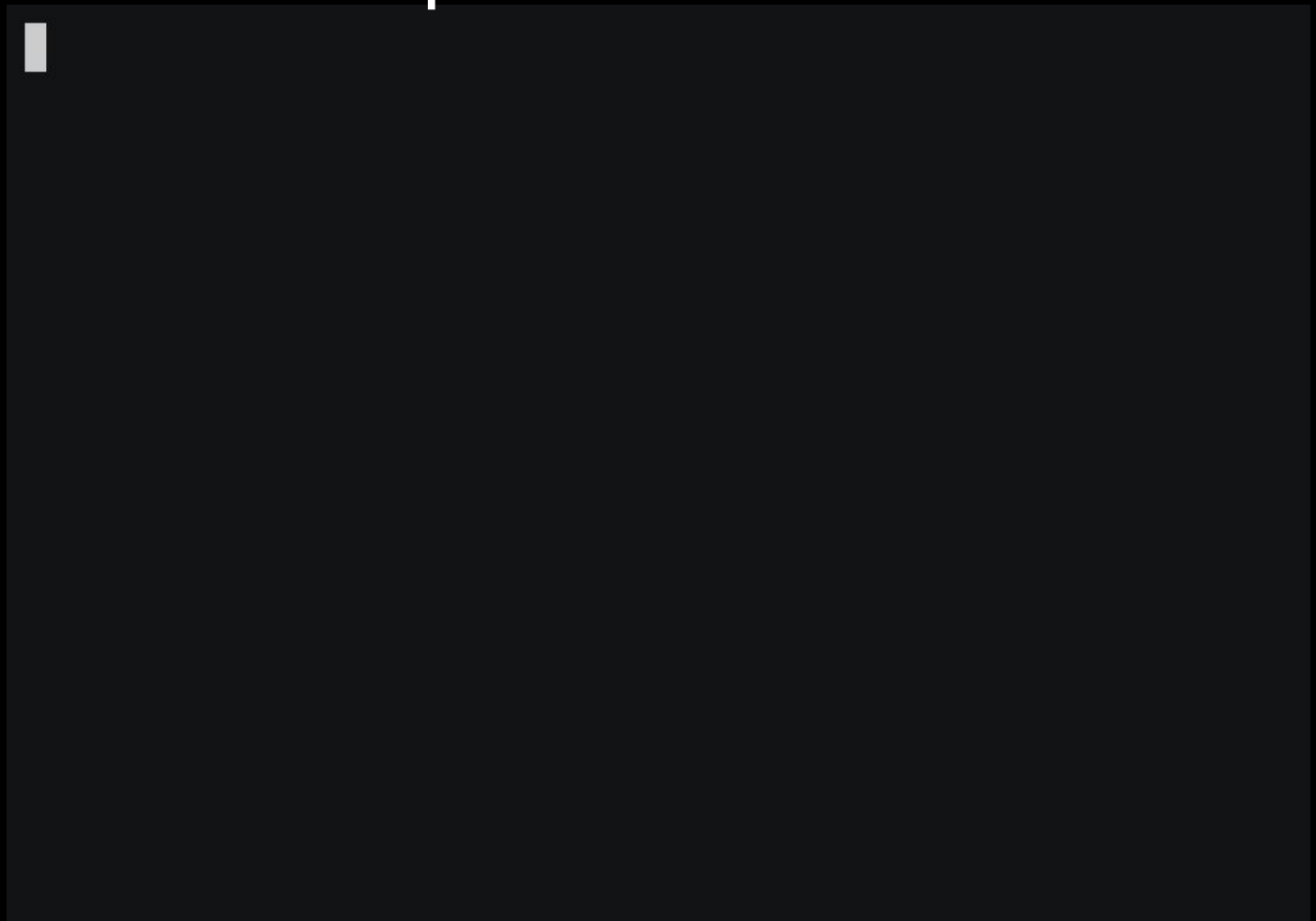
Created with clone() or unshare()

Net Namespace



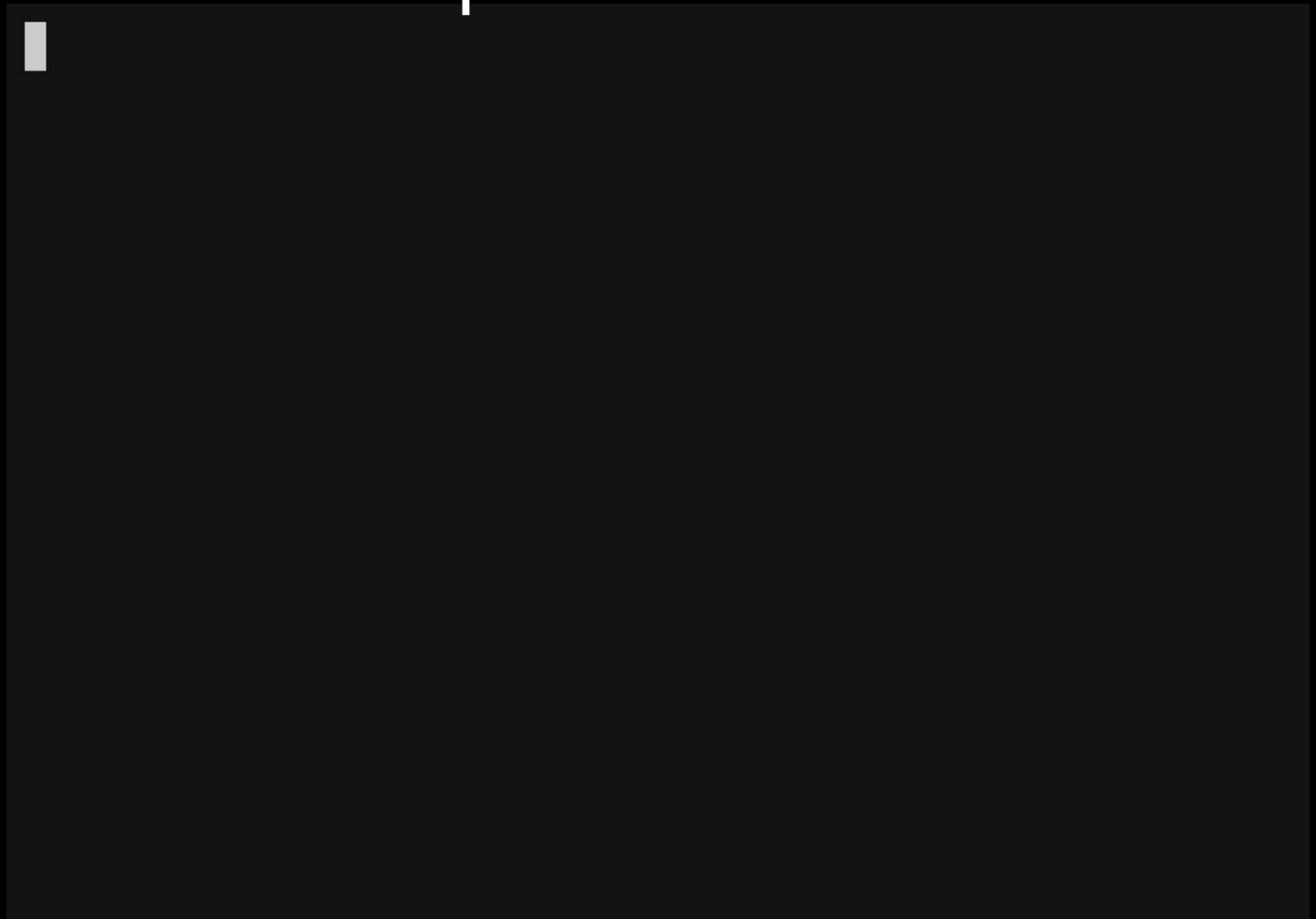
UTS Namespace

IPC Namespace



PID Namespace

User Namespace

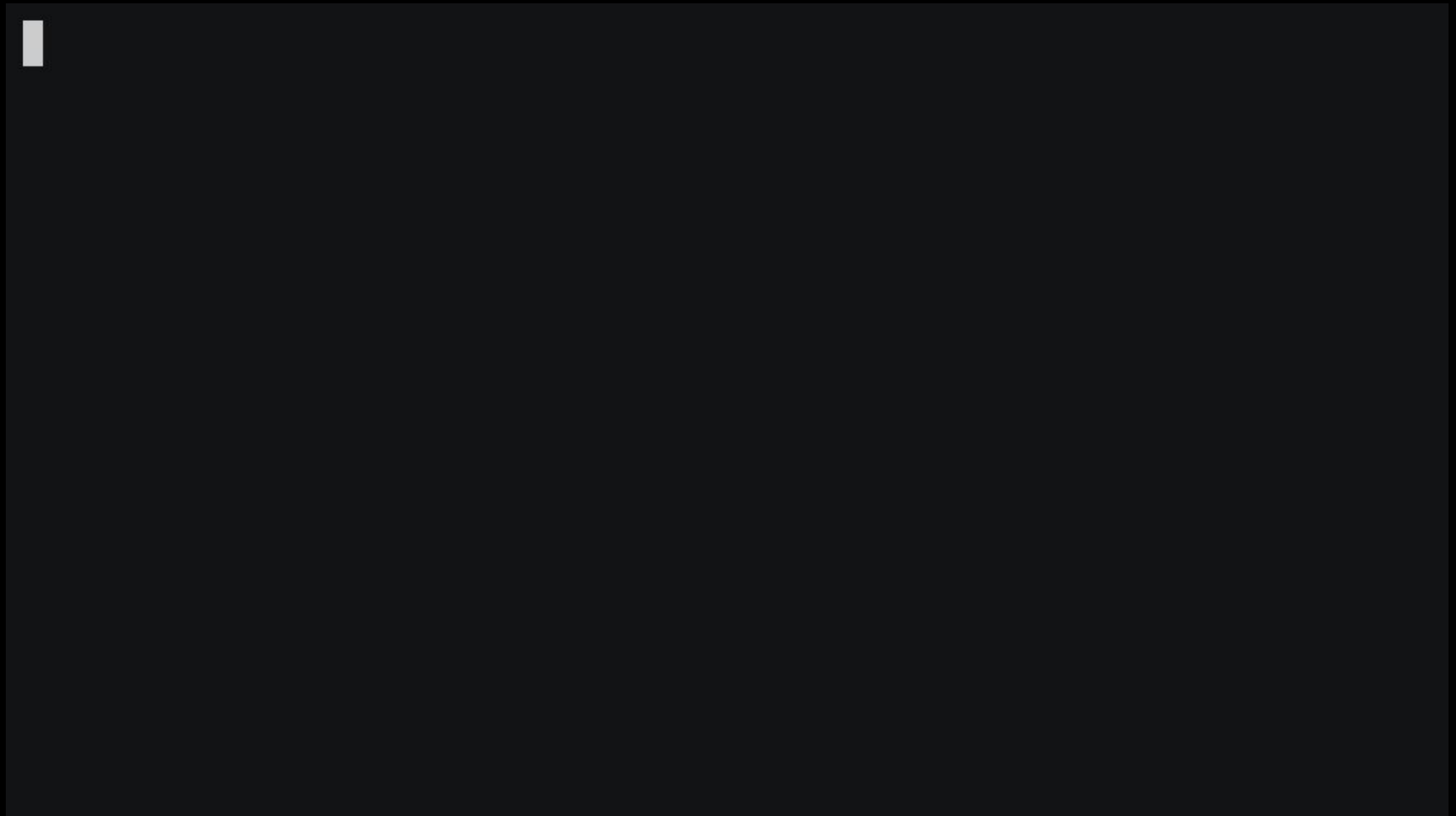


Makings of a Sandbox: Containers

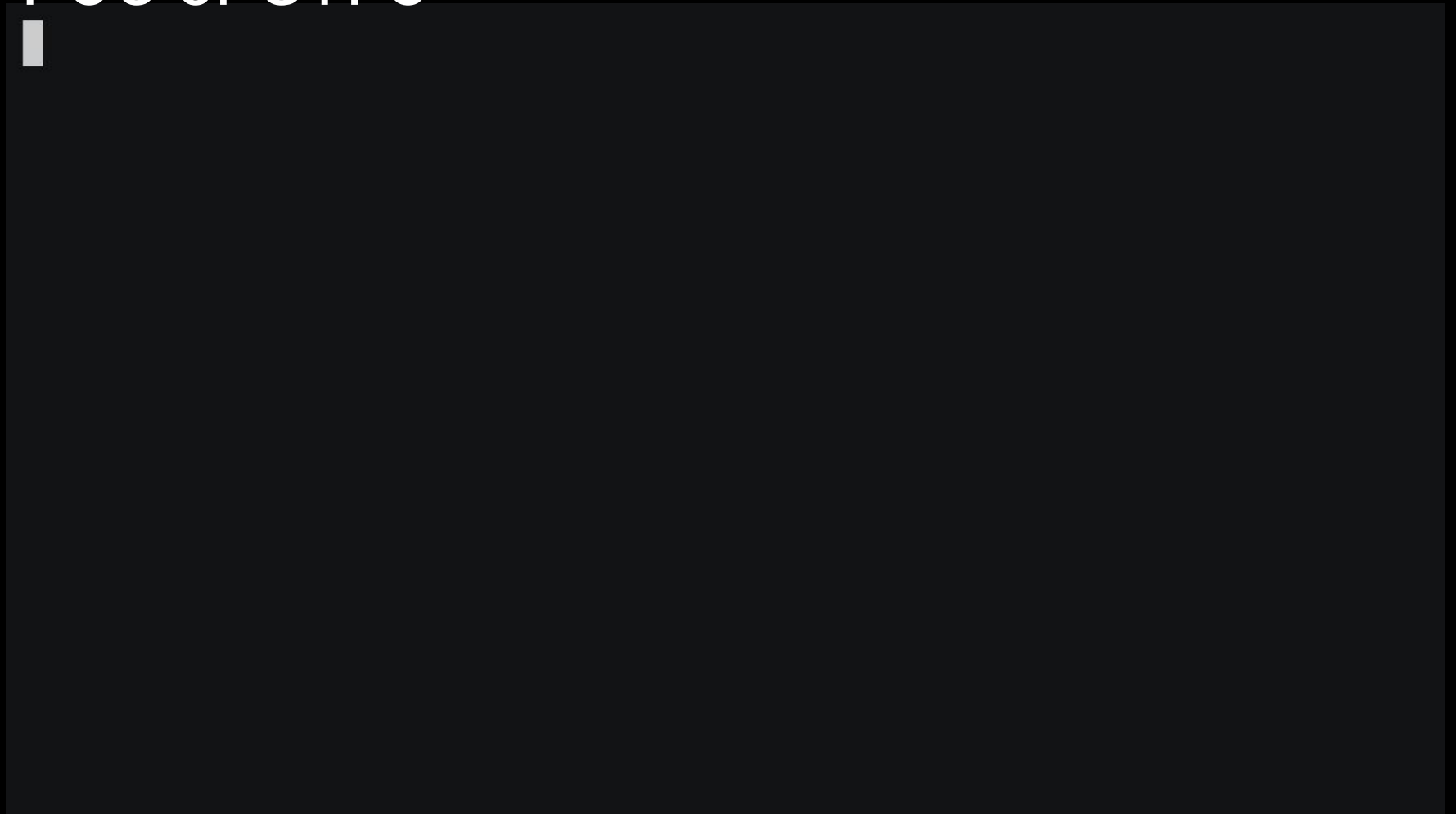
- Namespaces
- Apparmor
- Selinux
- Capabilities Limiting
- Cgroups

NOT RUN AS ROOT

POC or GTFO



POC or GTFO



What is this sorcery?

- User namespaces can be created without root.
- But only if the {uid,gid}_map is mapped to the current user creating the namespace.

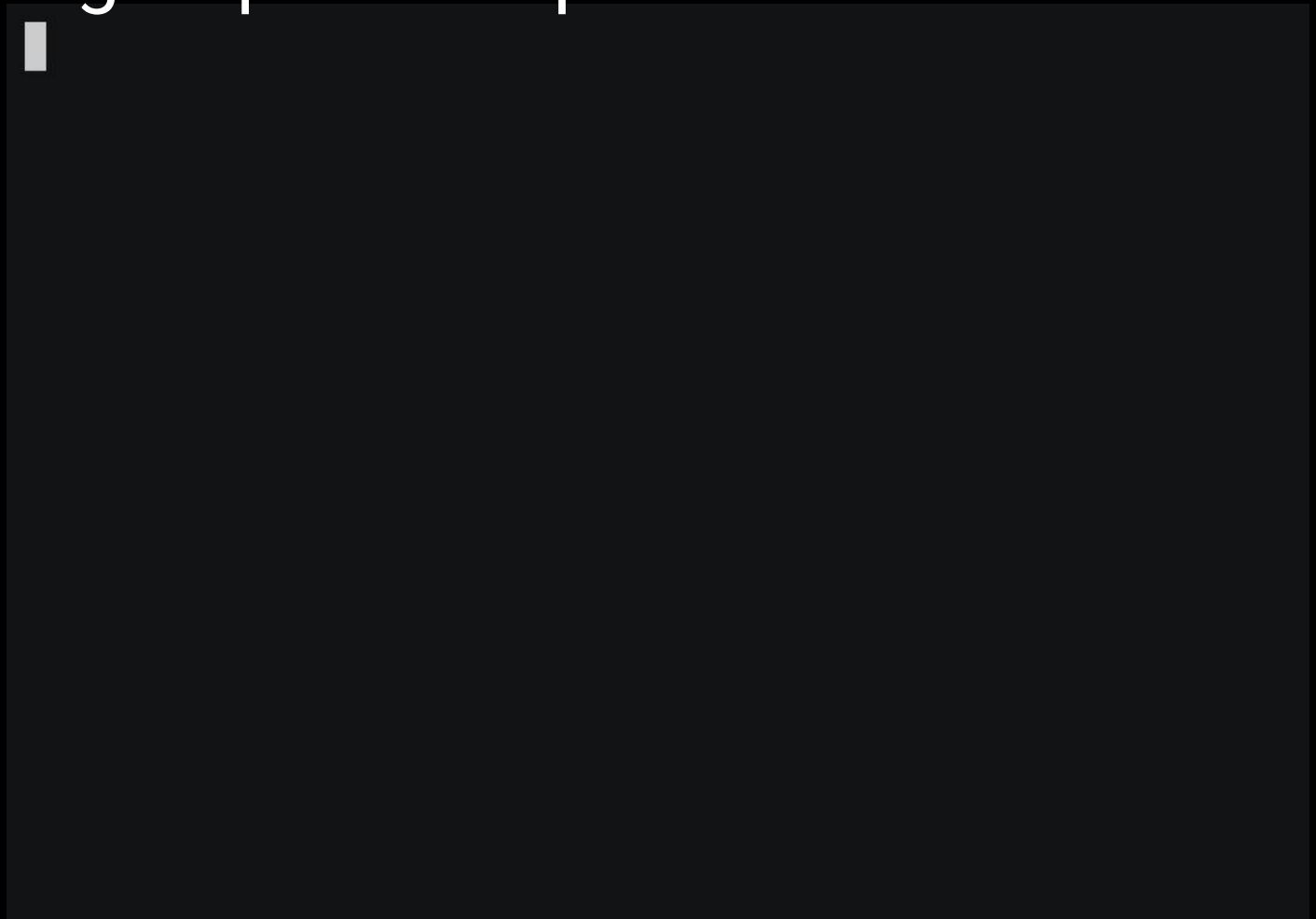
Not Perfect …. yet

- Cgroups devices cannot be created without CAP_SYS_ADMIN

New Hotness: Cgroup Namespace

- In Kernels 4.6+, not yet released, on RC5 currently
- False prophet to solve all the problems, but maybe in the future.

Cgroup Namespace



What to look forward to...

- Containers in a multi-tenant environment not run as root.
- Sane defaults with the ability to customize for a sandbox experience.
- Better designed user experiences for dealing with security policies.

Resources

https://github.com/docker/docker/issues/17142

http://www.sysdig.org/falco/