LINUX
PLUMBERS
CONFERENCE
August 24-28, 2020

pidfd & capabilities

Christian Brauner (Canonical)
<christian.brauner@ubuntu.com>

Delegating Privilege

- Current capability model is pretty hard to use.
- Privilege delegation is pretty difficult.
- Lot’s of complex state transition rules.
- PIDs can’t be used to carry a capability because they are a (pid namespace) global resource.

- People have asked for a simpler model especially in the browser and container world.
- pidfds are file descriptors and thus lend themselves naturally for privilege delegation.
- I’d like to discuss the idea of attaching capabilities to pidfds and what that could look like.
- Has been brought up a few times right at the beginning and often in other discussions.

Delegating Privilege

- Find a way/an API that allows to attach a capability to a pidfd.
- Ideally, permission is checked once at pidfd creation time.
- E.g. attach CAP_KILL to pidfd and allow any task with access to that pidfd to kill that task.
- What would an API for this look like? Would it be like Capsicum? Would one just re-use our already existing capabilities and just attach them to a pidfd (I played with that idea a while back.)?