The Clone Wars

(Sorry for the bad joke.*)

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

*Someone claimed they append “it’s ok, he’s German” to everyone of my jokes.

clone3()

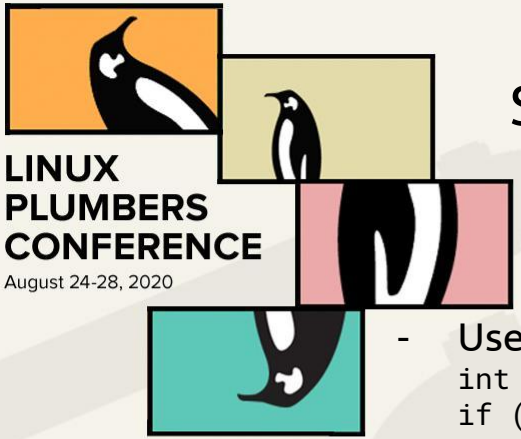
- New process creation system call since Linux v5.3
- Provides superset of clone() features (minus some insanities we left out)
- Is an extensible-struct based syscall

```
struct clone_args {
        __aligned_u64 flags;
        __aligned_u64 pidfd;
        __aligned_u64 child_tid;
        __aligned_u64 parent_tid;
        __aligned_u64 exit_signal;
        __aligned_u64 stack;
        __aligned_u64 stack_size;
        __aligned_u64 tls;
        __aligned_u64 set_tid;
        __aligned_u64 set_tid_size;
        __aligned_u64 cgroup;
};
```

- Has new extensions already CLONE_INTO_CGROUP, CLONE_CLEAR_SIGHAND

Current clone()/clone3() usage

- Heavily used in userspace (esp. container runtimes)
  - LXC: https://github.com/lxc/lxc/blob/9cc837ef2cc5e8f5fffb0ca987ff26478f19a46f/src/lxc/process_utils.c#L31
  - runC: https://github.com/opencontainers/runc/blob/0fa097fc37c5d89e4cea4fda4663d1239e12a6fe/libcontainer/nsenter/nsexec.c#L339
  - systemd/systemd-nspawn: https://github.com/systemd/systemd/blob/5238e9575906297608ff802a27e2ff9effa3b338/src/basic/raw-clone.h#L31



Some observations

- Userspace often wants a fork()-like wrapper:

```
int pid_t = clone3(&args, sizeof(args));
if (pid < 0)
      return -1;
if (pid == 0) {
      /* do stuff */
      _exit(EXIT_SUCCESS);
}
```

But currently clone() only has a callback-based wrapper:

```
int clone(int (*fn)(void *), void *stack, int flags, void *arg,
          ... /* pid_t *parent_tid, void *tls, pid_t *child_tid */);
```

Questions

- Can we add a wrapper for clone3()?
- Can we expose a fork()-like wrapper for clone3()?
- Can we handle atfork handler and libc internal state cleanly?
  https://sourceware.org/bugzilla/show_bug.cgi?id=26371
- What should we do about the stack argument?
- Can the kernel do something to make it easier to use and expose the stack argument?