Subject: Re: [PATCH 16/16] net: netlink support for moving devices between network namespaces.
Posted by ebiederm on Mon, 10 Sep 2007 19:30:46 GMT
View Forum Message <> Reply to Message

"Serge E. Hallyn" <serue@us.ibm.com> writes:
>>
>> +static struct net *get_net_ns_by_pid(pid_t pid)
>> +{
>> + struct task_struct *tsk;
>> + struct net *net;
>> +
>> + /* Lookup the network namespace */
>> + net = ERR_PTR(-ESRCH);
>> + rcu_read_lock();
>> + tsk = find_task_by_pid(pid);
>> + if (tsk) {
>> + task_lock(tsk);
>> + if (tsk->nsproxy)
>> +   net = get_net(tsk->nsproxy->net_ns);
>> + task_unlock(tsk);
>
> Thinking... Ok, I'm not sure this is 100% safe in the target tree, but
> the long-term correct way probably isn't yet implemented in the net-
> tree. Eventually you will want to:
>
> net_ns = NULL;
> rcu_read_lock();
> tsk = find_task_by_pid(); /* or _pidns equiv? */
> nsproxy = task_nsproxy(tsk);
> if (nsproxy)
>   net_ns = get_net(nsproxy->net_ns);
> rcu_read_unlock;
>
> What you have here is probably unsafe if tsk is the last task pointing
> to it's nsproxy and it does an unshare, bc unshare isn't protected by
> task_lock, and you're not rcu_dereferencing tsk->nsproxy (which
> task_nsproxy does). At one point we floated a patch to reuse the same
> nsproxy in that case which would prevent you having to worry about it,
> but that isn't being done in -mm now so i doubt it's in -net.

That change isn't merged upstream yet, so it isn't in David's
net-2.6.24 tree. Currently task->nsproxy is protected but
task_lock(current). So the code is fine.

I am aware that removing the task_lock(current) for the setting
of current->nsproxy is currently in the works, and I have planned

to revisit this later when all of these pieces come together.

For now the code is fine.

If need be we can drop this patch to remove the potential merge conflict. But I figured it was useful for this part of the user space interface to be available for review.

Eric

_____
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers