
Subject: Re: [PATCH 5/9] network namespaces: async socket operations

Posted by [Daniel Lezcano](#) on Fri, 22 Sep 2006 15:33:56 GMT

[View Forum Message](#) <> [Reply to Message](#)

Andrey Savochkin wrote:

> Non-trivial part of socket namespaces: asynchronous events

> should be run in proper context.

>

> Signed-off-by: Andrey Savochkin <saw@swsoft.com>

> ---

> af_inet.c | 10 ++++++++

> inet_timewait_sock.c | 8 ++++++

> tcp_timer.c | 9 ++++++++

> 3 files changed, 27 insertions(+)

>

> --- ./net/ipv4/af_inet.c.vensock-asyn Mon Aug 14 17:04:07 2006

> +++ ./net/ipv4/af_inet.c Tue Aug 15 13:45:44 2006

> @@ -366,10 +366,17 @@ out_rcu_unlock:

> int inet_release(struct socket *sock)

> {

> struct sock *sk = sock->sk;

> + struct net_namespace *ns, *orig_net_ns;

>

> if (sk) {

> long timeout;

>

> + /* Need to change context here since protocol ->close

> + * operation may send packets.

> + */

> + ns = get_net_ns(sk->sk_net_ns);

> + push_net_ns(ns, orig_net_ns);

> +

Is it not a race condition here ? What happens if you have a packet incoming during the namespace context switching ?

IHMO doing namespace switching is something dangerous, you can probably handle that with locks but it will be difficult and will decrease all network performance.

In an other hand, I don't see how you can handle the "sk->sk_prot->close" after ...

-- Cheers
