Subject: Re: [PATCH 0/5] Make nicer CONFIG_NET_NS=n case code
Posted by Eric Dumazet on Wed, 31 Oct 2007 18:49:24 GMT
View Forum Message <> Reply to Message

On Wed, 31 Oct 2007 22:19:43 +0300
Pavel Emelyanov <xemul@openvz.org> wrote:

> Currently we have the NET_NS config option, but the only change it
> makes is just return ERR_PTR(-EINVAL) inside the cloning call thus
> introducing a bunch of a dead code and making the reference counting
> unneeded. This is not very good.
>
> So clean the net_namespace.c to fix this.
>
> I have sent a set of patches to Andrew to make similar thing for
> other namespaces, which introduces the NAMESPACES option to turn
> all the namespaces off at once (to make embedded people suffer
> less). So after that stuff is in, there will be some more patches
> to tie all this together.
>
> What is to be done after this set is to make the register_pernet_xxx
> stuff smaller. Currently this code weights approximately 500 bytes,
> so it worths reducing it, but I haven't found a good solution yet.

Definitly wanted here. Thank you.
One more refcounting on each socket creation/deletion was expensive.

Maybe we can add a macro to get nd_net from a "struct net_device"
so that every instance of

```
if (dev->nd_net != &init_net)
    goto drop;
```

can also be optimized away if !CONFIG_NET_NS

```
extern inline netdev_get_ns(struct netdevice *dev)
{
#ifdef CONFIG_NET_NS
 return dev->nd_net;
#else
 return &init_net;
#endif
}

...

if (netdev_get_ns(dev) != &init_net)
 goto drop;
```