Subject: Re: The issues for agreeing on a virtualization/namespaces implementation.
Posted by Herbert Poetzl on Wed, 08 Feb 2006 04:37:21 GMT

On Tue, Feb 07, 2006 at 08:52:15PM -0700, Eric W. Biederman wrote:
> "Serge E. Hallyn" <serue@us.ibm.com> writes:
>
> >
> > What I tried to do in a proof of concept long ago was to have
> > CLONE_NETNS mean that you get access to all the network devices, but
> > then you could drop/add them. Conceptually I prefer that to getting an
> > empty namespace, but I'm not sure whether there's any practical use
> > where you'd want that...
>
> My observation was that the network stack does not come out cleanly
> as a namespace unless you adopt the rule that a network device
> belongs to exactly one network namespace.

yep, that's what the first network virtualization for
Linux-VServer aimed at, but found too complicated
the second one uses 'pairs' of communicating devices
to send between guests/host

> With that rule dealing with the network stack is just a matter of
> making some currently global variables/data structures per container.

yep, like the universal loopback and so ...

> A pain to do the first round but easy to maintain once you are there
> and the logic of the code doesn't need to change.

best,
Herbert

> Eric