Subject: Re: [patch 2/6] [Network namespace] Network device sharing by view
Posted by Andrey Savochkin on Mon, 26 Jun 2006 15:27:51 GMT
View Forum Message <> Reply to Message

Daniel,

On Mon, Jun 26, 2006 at 04:56:32PM +0200, Daniel Lezcano wrote:
> Andrey Savochkin wrote:
> >
> > It's good that you kicked off network namespace discussion.
> > Although I wish you'd Cc'ed someone at OpenVZ so I could notice it earlier :).
>
> devel@openvz.org ?

devel@openvz.org is fine

>
> > When a device presents an skb to the protocol layer, it needs to know to which
> > namespace this skb belongs.
> > Otherwise you would never get rid of problems with bind: what to do if device
> > eth1 is visible in namespace1, namespace2, and root namespace, and each
> > namespace has a socket bound to 0.0.0.0:80?
>
> Exact. But, the idea was to retrieve the namespace from the routes.

Then you lose the ability for each namespace to have its own routing entries.
Which implies that you'll have difficulties with devices that should exist
and be visible in one namespace only (like tunnels), as they require IP
addresses and route.

>
> IMHO, I think there are roughly 2 network isolation implementation:
>
> - make all network ressources private to the namespace
>
> - keep a "flat" model where network ressources have a new identifier
> which is the network namespace pointer. The idea is to move only some
> network informations private to the namespace (eg port range, stats, ...)

Sorry, I don't get the second idea with only some information private to
namespace.

How do you want TCP_INC_STATS macro look?
In my concept, it would be something like
#define TCP_INC_STATS(field) SNMP_INC_STATS(current_net_ns->tcp_stat, field)
where tcp_stat is a TCP statistics array inside net_namespace.

Regards

Andrey