Subject: Re: [patch 2/6] [Network namespace] Network device sharing by view
Posted by Herbert Poetzl on Mon, 26 Jun 2006 18:36:49 GMT
View Forum Message <> Reply to Message

On Mon, Jun 26, 2006 at 10:40:59AM -0600, Eric W. Biederman wrote:
> Daniel Lezcano <dlezcano@fr.ibm.com> writes:
>
> >> 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.
> >
> > I mean instead of having the route tables private to the namespace, the routes
> > have the information to which namespace they are associated.
>
> Is this an implementation difference or is this a user visible
> difference? As an implementation difference this is sensible, as it is
> pretty insane to allocate hash tables at run time.
>
> As a user visible difference that affects semantics of the operations
> this is not something we want to do.

well, I guess there are even more options here, for
example I'd like to propose the following idea, which
might be a viable solution for the policy/isolation
problem, with the actual overhead on the setup part
not the hot pathes for packet and connection handling

we could use the multiple routing tables to provide
a single routing table for each guest, which could
be used inside the guest to add arbitrary routes, but
would allow to keep the 'main' policy on the host, by
selecting the proper table based on IPs and guest tags

similar we could allow to have a separate iptables
chain for each guest (or several chains), which are
once again directed by the host system (applying the
required prolicy) which can be managed and configured
via normal iptable interfaces (both on the guest and
host) but actually provide at least to layers

note: this does not work for hierarchical network
contexts, but I do not see that the yet proposed
implementations would do, so I do not think that
is of concern here ...

best,
Herbert

> Eric