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

Subject: Re: [patch 2/6] [Network namespace] Network device sharing by view
Posted by Daniel Lezcano on Mon, 26 Jun 2006 15:49:41 GMT
View Forum Message <> Reply to Message

> 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.

>
> - 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?

I was thinking in TCP_INC_STATS(net_ns, field)
SNMP_INC_STATS(net_ns->tcp_stat, field)

> 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.

Subject: Re: [patch 2/6] [Network namespace] Network device sharing by view
Posted by Andrey Savochkin on Tue, 27 Jun 2006 09:11:36 GMT
View Forum Message <> Reply to Message

Daniel,

On Mon, Jun 26, 2006 at 05:49:41PM +0200, Daniel Lezcano wrote:
>
> > 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.

I think I understand what you're talking about: you want to make routing
responsible for determining destination namespace ID in addition to route
type (local, unicast etc), nexthop information, and so on. Right?

My point is that if you make namespace tagging at routing time, and
your packets are being routed only once, you lose the ability
to have separate routing tables in each namespace.

Andrey

Subject: Re: [patch 2/6] [Network namespace] Network device sharing by view
Posted by Daniel Lezcano on Tue, 27 Jun 2006 09:34:36 GMT
View Forum Message <> Reply to Message

Andrey Savochkin wrote:
> Daniel,
>
> On Mon, Jun 26, 2006 at 05:49:41PM +0200, Daniel Lezcano wrote:
>
>>>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.
>
>
> I think I understand what you're talking about: you want to make routing
> responsible for determining destination namespace ID in addition to route
> type (local, unicast etc), nexthop information, and so on. Right?

Yes.

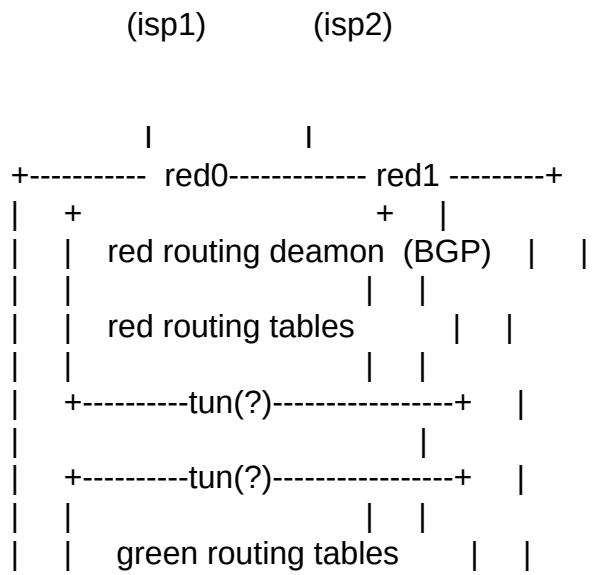
>
> My point is that if you make namespace tagging at routing time, and
> your packets are being routed only once, you lose the ability
> to have separate routing tables in each namespace.

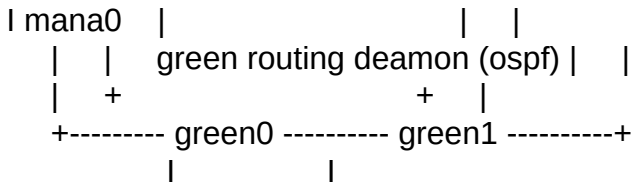
Right. What is the advantage of having separate the routing tables ?

Subject: Routing tables (Re: [patch 2/6] [Network namespace] Network device sharing by view)
Posted by Kari Hurtta on Thu, 06 Jul 2006 09:45:29 GMT
View Forum Message <> Reply to Message

> Andrey Savochkin wrote:
> > Daniel,
> >
> > On Mon, Jun 26, 2006 at 05:49:41PM +0200, Daniel Lezcano wrote:
> >
> >>>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.
> >
> >
> > I think I understand what you're talking about: you want to make routing
> > responsible for determining destination namespace ID in addition to route
> > type (local, unicast etc), nexthop information, and so on. Right?
>
> Yes.
>
> >
> > My point is that if you make namespace tagging at routing time, and
> > your packets are being routed only once, you lose the ability
> > to have separate routing tables in each namespace.
>
> Right. What is the advantage of having separate the routing tables ?

One application may be following. Consider firewall





That may allow running different routing deamon on
red and green side. That is possible if they manage
different routing tables on kernel. They not need
communigate together, when route between them is static.

/ Kari Hurtta

> -
> To unsubscribe from this list: send the line "unsubscribe linux-kernel" in
> the body of a message to majordomo@vger.kernel.org
> More majordomo info at http://vger.kernel.org/majordomo-info.html
> Please read the FAQ at http://www.tux.org/lkml/