Subject: Re: Network namespaces without isolation
Posted by ebiederm on Fri, 04 Jul 2008 09:52:08 GMT
View Forum Message <> Reply to Message

Andreas B Aaen <andreas.aaen@tietoenator.com> writes:

> Hi,
>
> I am looking into the network namespace implementation because I need an IP
> stack that is capable of talking with a number of separate IP nets with
> possible overlapping IP adresses. My connection to each separate IP-net is
> through a tunnel e.g. a VLAN interface.
>
> A special application will then be able to listen to traffic on all the nets
> through a socket option SO_NS that sets the namespace to talk/listen to for a
> particular socket. For this to work network namespaces needs to be indexed.

Answering part of your question. As currently designed you can use multiple
network namespaces in a single task, and you can place each vlan interface
in different network namespace. However the current model is most cumbersome
for doing so.

You can use unshare instead of clone which is a little easier.

A socket option sounds like a nice idea.

The two challenges are what names to use to refer to network namespaces
and how to get network namespaces to persist.

There have been a number of discussions about identifiers none of which
have led to any sort of agreement. One of the goals in the design is
that we don't introduce new global identifiers allowing us to ultimately
have nested containers.

So far we have been referring to namespaces indirectly by the pids of the
processes which are using them.

> It would also be nice to be able to see the network statistics from all the
> namespaces through the proc filesystem at least in an uncloned (isolated)
> namespace.

Currently this is possible by looking at /proc/<pid>/net.

> So you would be able to see the network statistics in /proc/net/ns/<index>/

One of the things we have tried to do is to keep the number of new interfaces
to a minimum.

> It should be said that we have an implementation of all this already, but NOT
> based on network namespaces and for elder kernels. We don't want to forward
> port this, put instead add a few features to the network namespace
> implementation to be able to fullfill the requirement of our application:
> talk to a number of IP networks with possible overlapping IP addresses.

If we can work out the details on how to do that cleanly it seems totally reasonable to enhance network namespaces in that direction. You are not the first to express those kind of requirements, and probably won't be the last.

Eric

_____

Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containers