
Subject: Re: Network virtualization/isolation

Posted by [Herbert Poetzl](#) on Mon, 04 Dec 2006 17:19:26 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Mon, Dec 04, 2006 at 08:02:48PM +0300, Dmitry Mishin wrote:

> On Monday 04 December 2006 19:43, Herbert Poetzl wrote:

> > On Mon, Dec 04, 2006 at 06:19:00PM +0300, Dmitry Mishin wrote:

> > > On Sunday 03 December 2006 19:00, Eric W. Biederman wrote:

> > > > Ok. Just a quick summary of where I see the discussion.

> > > >

> > > > We all agree that L2 isolation is needed at some point.

> > > >

> > > As we all agreed on this, may be it is time to send patches

> > > one-by-one? For the beggining, I propose to resend Cedric's

> > > empty namespace patch as base for others - it is really empty,

> > > but necessary in order to move further.

> > >

> > > After this patch and the following net namespace unshare

> > > patch will be accepted,

> >

> > well, I have neither seen any performance tests showing

> > that the following is true:

> >

> > - no change on network performance without the

> > space enabled

> > - no change on network performance on the host

> > with the network namespaces enabled

> > - no measureable overhead inside the network

> > namespace

> > - good scalability for a larger number of network

> > namespaces

> These questions are for complete L2 implementation,

> not for these 2 empty patches.

well, I fear that we will have lot of overhead

'sneaking' in via small patches (with almost unnoticeable overhead) making the 2.6 branch slower and slower (regarding networking) so IMHO a complete solution should be drafted, and tested performance wise, we can then adjust it and possibly improve it, untill it shows no measureable overhead ...

but IMHO it should be developed 'outside' the kernel, in small and reviewable pieces which are constantly updated to match the recent kernels ... something like stacked git or quilt ...

> If you need some data relating to Andrey's
> implementation, I'll get it. Which test do you accept?

hmm, I think a good mix of netperf, netpipe and
iperf would be a good start, probably network folks
know better tests to exercise the stack ... at least
I hope so ...

of course, a good explanation _why_ this or that
code path does not add overhead here or there is
nice to have too ...

best,
Herbert

> > > I could send network devices virtualization patches for
> > > review and discussion.

> >

> > that won't hurt ...

> >

> > best,

> > Herbert

> >

> > > What do you think?

> > >

> > > > The approaches discussed for L2 and L3 are sufficiently orthogonal

> > > > that we can implement then in either order. You would need to

> > > > unshare L3 to unshare L2, but if we think of them as two separate

> > > > namespaces we are likely to be in better shape.

> > > >

> > > > The L3 discussion still has the problem that there has not been

> > > > agreement on all of the semantics yet.

> > > >

> > > > More comments after I get some sleep.

> > > >

> > > > Eric

> > > > -

> > > > To unsubscribe from this list: send the line "unsubscribe netdev" in

> > > > the body of a message to majordomo@vger.kernel.org

> > > > More majordomo info at <http://vger.kernel.org/majordomo-info.html>

> > >

> > > --

> > > Thanks,

> > > Dmitry.

> > >

> > > Containers mailing list

> > > Containers@lists.osdl.org

> > > <https://lists.osdl.org/mailman/listinfo/containers>

>
> --
> Thanks,
> Dmitry.

Containers mailing list
Containers@lists.osdl.org
<https://lists.osdl.org/mailman/listinfo/containers>
