

---

Subject: Re: [patch 1/4] Network namespaces: cleanup of dev\_base list use  
Posted by [Andrey Savochkin](#) on Mon, 26 Jun 2006 15:42:05 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

Hi Eric,

On Mon, Jun 26, 2006 at 09:13:52AM -0600, Eric W. Biederman wrote:

> Andrey Savochkin <saw@swsoft.com> writes:

>

> > Cleanup of dev\_base list use, with the aim to make device list per-namespace.

> > In almost every occasion, use of dev\_base variable and dev->next pointer

> > could be easily replaced by for\_each\_netdev loop.

> > A few most complicated places were converted to using

> > first\_netdev()/next\_netdev().

>

> As a proof of concept patch this is ok.

>

> As a real world patch this is much too big, which prevents review.

> Plus it takes a few actions that are more than replace just

> iterators through the device list.

dev\_base list is historically not the cleanest part of Linux networking.

I've still spotted a place where the first device in dev\_base list is assumed to be loopback. In early days we had more, now only one place or two...

>

> In addition I suspect several if not all of these iterators  
> can be replaced with the an appropriate helper function.

>

> The normal structure for a patch like this would be to  
> introduce the new helper function. for\_each\_netdev.

> And then to replace all of the users while cc'ing the  
> maintainers of those drivers. With each different  
> driver being a different patch.

>

> There is another topic for discussion in this patch as well.

> How much of the context should be implicit and how much  
> should be explicit.

>

> If the changes from netchannels had already been implemented, and all of  
> the network processing was happening in a process context then I would  
> trivially agree that implicit would be the way to go.

Why would we want all network processing happen in a process context?

>

> However short of always having code always execute in the proper  
> context I'm not comfortable with implicit parameters to functions.

> Not that this the contents of this patch should address this but the  
> later patches should.

We just have too many layers in networking code, and FIB/routing illustrates it well.

>  
> When I went through this, my patchset just added an explicit  
> continue if the devices was not in the appropriate namespace.  
> I actually prefer the multiple list implementation but at  
> the same time I think it is harder to get a clean implementation  
> out of it.

Certainly, dev\_base list reorganization is not the crucial point in network namespaces. But it has to be done some way or other.  
If people vote for a single list with skipping devices from a wrong namespace, it's fine with me, I can re-make this patch.

I personally prefer per-namespace device list since we have too many places in the kernel where this list is walked in a linear fashion, and with many namespaces this list may become quite long.

Regards

Andrey

---