Subject: Re: Network namespaces a path to mergable code.
Posted by ebiederm on Wed, 28 Jun 2006 04:20:32 GMT
View Forum Message <> Reply to Message

Andrey Savochkin <saw@swsoft.com> writes:

> Eric,
>
> On Tue, Jun 27, 2006 at 11:20:40AM -0600, Eric W. Biederman wrote:
>>
>> Thinking about this I am going to suggest a slightly different direction
>> for get a patchset we can merge.
>>
>> First we concentrate on the fundamentals.
>> - How we mark a device as belonging to a specific network namespace.
>> - How we mark a socket as belonging to a specific network namespace.
>
> I agree with the direction of your thoughts.
> I was trying to do a similar thing, define clear steps in network
> namespace merging.
>
> My first patchset covers devices but not sockets.
> The only difference from what you're suggesting is ipv4 routing.
> For me, it is not less important than devices and sockets. May be even
> more important, since routing exposes design deficiencies less obvious at
> socket level.

I agree we need to do it. I mostly want a base that allows us to
not need to convert the whole network stack at once and still be able
to merge code all the way to the stable kernel.

The routing code is important for understanding design choices. It
isn't important for merging if that makes sense.

For everyone looking at routing choices the IPv6 routing table is
interesting because it does not use a hash table, and seems quite
possibly to be an equally fast structure that scales better.

There is something to think about there.

>> As part of the fundamentals we add a patch to the generic socket code
>> that by default will disable it for protocol families that do not indicate
>> support for handling network namespaces, on a non-default network namespace.
>
> Fine
>
> Can you summarize you objections against my way of handling devices, please?
> And what was the typo you referred to in your letter to Kirill Korotaev?

I have no fundamental objects to the content I have seen so far.

Please read the first email Kirill responded too. I quoted a couple of sections of code and described the bugs I saw with the patch.

All minor things. The typo I was referring to was a section where the original iteration was on an ifp variable and you called it dev without changing the rest of the code in that section.

The only big issue was that the patch too big, and should be split into a patchset for better review. One patch for the new functions, and the an additional patch for each driver/subsystem hunk describing why that chunk needed to be changed.

I'm still curious why many of those chunks can't use existing helper functions, to be cleaned up.

Eric