Subject: Re: Pid namespaces problems
Posted by Pavel Emelianov on Thu, 08 Nov 2007 10:40:37 GMT
View Forum Message <> Reply to Message

Daniel Lezcano wrote:
> Pavel Emelyanov wrote:
>> Daniel Lezcano wrote:
>>> Cedric Le Goater wrote:
>>>>>> - There are several architectures with their own signal functions for
>>>>>>   other OS compatibility that have are using _pid and not _vpid
>>>>>>   variants of functions. (irix and solaris)
>>>>>>   arch/mips/kernel/irixsig.c:irix_waitsys
>>>>>>   arch/mips/kernel/sysirix.c:irix_setpgrp
>>>>>>   arch/sparc64/solaris/misc.c:solaris_procids
>>>>> Ok. Looks like your list is the same as mine. That's good to hear
>>>>> that I haven't missed anything important.
>>>> We've also talked about af_unix credentials.
>>>>
>>>>> So, I see that you're about to take a closer look at the pid
>>>>> namespaces. If so, then what time can we expect the net namespace
>>>>> activity to go on? Or (if you don't mind) can we start merging
>>>>> the patches to David as soon as he opens his 2.6.25 merge window?
>>>> I think daniel and benjamin are also getting ready for the 2.6.25
>>>> merge window.
>>> Yes. It can be cool if we can sync up Benjamin, Pavel, Denis, Eric and I
>>> with the different parts to be posted.
>> Yup. Team work will give us a chance to get in to the 2.6.25 with
>> the core virtualization. By core I mean unix, netlinux, ipv4 and ipv6.
>
> Yeah, if we can push these protocols in time, that will be *very very*
> cool :)
>
> When you talk about ipv4/ipv6 do you include tcp/udp ?
>
>> Netfilters virtualization is a complex task :)
>>
>>> Benjamin and I we began to look
>>> at ipv4. This is a big part, perhaps we can split that into several
>>> subset and dispatch them, except if Pavel and Denis already rebase ipv4
>>> for net-2.6, in this case feel free to send them out.
>> Well, actually we have almost moved to the net-2.6 with the ipv4
>> set.
>
> Excellent !
> Did you took the different patches I sent for udplite and multicast ?

Not yet, sorry. But we will look at them as soon as we finish with
the existing netns tree.

> If you rebased netns49 to net-2.6 and you plan to keep synced with the
> Dave Miller tree, perhaps it is time to switch the git tree.
> It can be cool if you can put a git tree at openvz.

Sure, but right now we don't have a git repo with it :(We
just have a series of patches, some of them are fixes, some
not. We are going to re-split them and publish to git. I think
that the next week the git.openvz.org will have that repo.

What about your ipv6? Do you have a git repo with it?

>> There are only some minor (I hope they are minor ;)) things.
>
> Perhaps, we can help here.

Yes, of course. When we publish the git repo we'll try to
provide a TODO list so that everyone can participate.

>> So we would be glad to go on with ipv4 further. What's up with the
>> ipv6 patches, Daniel? You said that you and Benjamin make some
>> big progress in this area, no?
>
> Yes, for the moment we reach the addrconf stuff, so we have routing
> table, ip6_fib, fib6_rules, ndisc and addrconf per namespaces.
> The patches sent by Alexey Dobriyan making /proc/net/ipv6_route to the
> seq_file interface has made our life easier.

:) We plan to make some cleanups in the networking code that
make sense even now, but are useful for namespaces.

> IMHO this part of ipv6 is the most difficult, the different protocols
> relying on it will be much more easy to implement.
>
> We are actually facing two problems:
>
>   * the first one is the locking of the network namespace list by
> rtnl_lock, so from the timer callback we can not browse the network
> namespace list to check the age of the routes. It is a problem I would
> like to talk with Denis if he has time

Sure. I will kick him in case he missed it accidentally :)

>   * the loopback refcounting is not correctly handled in ipv6. This
> protocol do not expect to have the loopback to be unregistered, so there
> is some problem with the addr_ifdown function when exiting the network
> namespace

Hm! This is one of the issues we can send to David right now. Do
you have any patches? If not, I can take care of it if you don't mind.

> Regards.
>   -- Daniel
>

Thanks,
Pavel

_____