
Subject: Re: [PATCH 4/4 (resent) net-2.6.25][UNIX] Make the unix sysctl tables per-namespace

Posted by [Pavel Emelianov](#) on Sat, 01 Dec 2007 13:19:43 GMT

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

Denis V. Lunev wrote:

> Herbert Xu wrote:

>> On Fri, Nov 30, 2007 at 07:37:28PM +0300, Pavel Emelyanov wrote:

>>> Signed-off-by: Pavel Emelyanov <xemul@openvz.org>

>> All applied to net-2.6.25.

>>

>>> diff --git a/include/net/net_namespace.h b/include/net/net_namespace.h

>>> index b0cf075..f97b2a4 100644

>>> --- a/include/net/net_namespace.h

>>> +++ b/include/net/net_namespace.h

>>> @@ -41,6 +43,7 @@ struct net {

>>>

>>> /* unix sockets */

>>> int sysctl_unix_max_dgram_qlen;

>>> + struct ctl_table_header *unix_ctl;

>>> };

>> But I gotta say this struct/file is going to be enormous. It's also

>> one of those files that causes everything to get recompiled. Maybe

>> we ought to make a rule that each subsystem only gets to have at most

>> one entry in it :)

>>

>> Thanks,

>

> Good point, thanks. We'll start thinking in that direction. Right now it

> is not finally cursed with all staff around.

Agree, the point is good :) but it has one pitfall :(

Look, now we make `_one_` dereference to get any `net->xxx` variable (`sysctl`, `list head`, `lock`, etc). When we force each subsystem has it's "private" pointer on this, we'll make them take `_two_` dereferences. Before the whole net namespace stuff started we made `_zero_` dereferences :) This may tell upon the performance.

I'm not claiming that this is the major case against this idea, but when developing this idea, I think we should keep that fact in mind and pay good attention to performance regressions.

> Regards,

> Den

Thanks,

Pavel

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