Subject: [NETNS] Oops in register_pernet_operations() with CONFIG_NET_NS=n
Posted by Benjamin Thery on Thu, 25 Oct 2007 12:59:03 GMT
View Forum Message <> Reply to Message

Hello Pavel,

I've found a problem with one of your patch related to netns:

* [NETNS] Move some code into __init section when CONFIG_NET_NS=n (v2)
  http://www.spinics.net/lists/netdev/msg43310.html

This patch introduces the __net_init/__net_exit/__net_initdata
defines to save some memory when CONFIG_NET_NS is not set.

Cedric Le Goater reported he had a *non-fatal* oops when booting
a 2.6.23-mm1-lxc1 kernel with CONFIG_NET_NS=n. (2.6.23-mm1-lxc1
contains the NETNS49 patchset). The oops occured when modules
related to iptables were loaded after the boot completes.

The problem is the following:

- Your patch adds the __net_initdata attribute to pernet_operations
  structures.

- pernet_operations are registered via register_pernet_subsys() and
  linked in the pernet_list during boot.

- At the end of boot, pernet_operations are freed (because of the
  __net_initdata attribute), and the pernet_list (or first_device list)
  points to freed memory.

- After boot, network modules which are netns-aware try to register
  themselves with register_pernet_subsys() and ...KABOOM... page
  fault when accessing pernet_list (or first_device list).
  (I reproduce Cedric's oops with the command: iptables --list)

This is not a problem right now in 2.6.23-mm1 or net-2.6 because
there are very few netns-aware network subsystems merged and they
are all initialized during boot. But it will be problematic when
we will merge netns code for subsystems which can be built as
modules (eg. iptables, ...). I'm not sure we can use
__net_init_data for pernet_operations then.
Maybe we can add some checks in register_pernet_operations
when CONFIG_NET_NS=n.

I haven't found a fix yet.

Regards,

Benjamin

--
B e n j a m i n T h e r y - BULL/DT/Open Software R&D

http://www.bull.com

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