
Subject: [patch 1/1][NETNS] resend: fix net released by rcu callback

Posted by [Daniel Lezcano](#) on Tue, 30 Oct 2007 16:21:40 GMT

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

When a network namespace reference is held by a network subsystem, and when this reference is decremented in a rcu update callback, we must ensure that there is no more outstanding rcu update before trying to free the network namespace.

In the normal case, the rcu_barrier is called when the network namespace is exiting in the cleanup_net function.

But when a network namespace creation fails, and the subsystems are undone (like the cleanup), the rcu_barrier is missing.

This patch adds the missing rcu_barrier.

Signed-off-by: Daniel Lezcano <dlezcano@fr.ibm.com>

net/core/net_namespace.c | 2 ++
1 file changed, 2 insertions(+)

Index: net-2.6/net/core/net_namespace.c

=====

--- net-2.6.orig/net/core/net_namespace.c

+++ net-2.6/net/core/net_namespace.c

@@ -112,6 +112,8 @@ out_undo:

if (ops->exit)
ops->exit(net);

}

+

+ rcu_barrier();

goto out;

}

--

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>

Subject: Re: [patch 1/1][NETNS] resend: fix net released by rcu callback

Posted by [ebiederm](#) on Tue, 30 Oct 2007 20:47:34 GMT

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

Daniel Lezcano <dlezcano@fr.ibm.com> writes:

> When a network namespace reference is held by a network subsystem,
> and when this reference is decremented in a rcu update callback, we
> must ensure that there is no more outstanding rcu update before
> trying to free the network namespace.
>
> In the normal case, the rcu_barrier is called when the network namespace
> is exiting in the cleanup_net function.
>
> But when a network namespace creation fails, and the subsystems are
> undone (like the cleanup), the rcu_barrier is missing.
>
> This patch adds the missing rcu_barrier.

Looks sane. Did you have any specific failures related to this or was this something that was just caught in review?

Eric

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

Subject: Re: [patch 1/1][NETNS] resend: fix net released by rcu callback
Posted by [Daniel Lezcano](#) on Tue, 30 Oct 2007 21:43:26 GMT
[View Forum Message](#) <> [Reply to Message](#)

Eric W. Biederman wrote:

> Daniel Lezcano <dlezcano@fr.ibm.com> writes:
>
>> When a network namespace reference is held by a network subsystem,
>> and when this reference is decremented in a rcu update callback, we
>> must ensure that there is no more outstanding rcu update before
>> trying to free the network namespace.
>>
>> In the normal case, the rcu_barrier is called when the network namespace
>> is exiting in the cleanup_net function.
>>
>> But when a network namespace creation fails, and the subsystems are
>> undone (like the cleanup), the rcu_barrier is missing.
>>
>> This patch adds the missing rcu_barrier.
>
> Looks sane. Did you have any specific failures related to this or was
> this something that was just caught in review?

Yes, I had this problem when doing ipv6 isolation for netns49. The ipv6

subsystem creation failed and the different subsystem where rollbacked in the setup_net function.
When the network namespace was about to be freed in free_net function, I had the error with an usage refcount different from zero.
It appears that was coming from core/neighbour.c

```
neigh_parms_release
-> neigh_rcu_free_parms
-> neigh_parms_put
-> neigh_parms_destroy
-> release_net
```

The free_net function was called before rcu callback neigh_rcu_free_parms.

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

Subject: Re: [patch 1/1][NETNS] resend: fix net released by rcu callback
Posted by [ebiederm](#) on Tue, 30 Oct 2007 22:34:29 GMT
[View Forum Message](#) <> [Reply to Message](#)

Daniel Lezcano <dlezcano@fr.ibm.com> writes:

```
> Yes, I had this problem when doing ipv6 isolation for netns49. The ipv6
> subsystem creation failed and the different subsystem where rollbacked in the
> setup_net function.
> When the network namespace was about to be freed in free_net function, I had the
> error with an usage refcount different from zero.
> It appears that was coming from core/neighbour.c
>
> neigh_parms_release
> -> neigh_rcu_free_parms
> -> neigh_parms_put
> -> neigh_parms_destroy
> -> release_net
>
> The free_net function was called before rcu callback neigh_rcu_free_parms.
```

Ok. Cool. It is good to know that this patch was tested. These kinds of subtle fixes are always more solid when you can actually test them.

Eric

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

Subject: Re: [patch 1/1][NETNS] resend: fix net released by rcu callback
Posted by [davem](#) on Tue, 30 Oct 2007 22:39:05 GMT
[View Forum Message](#) <> [Reply to Message](#)

From: Daniel Lezcano <dlezcano@fr.ibm.com>
Date: Tue, 30 Oct 2007 17:21:40 +0100

> When a network namespace reference is held by a network subsystem,
> and when this reference is decremented in a rcu update callback, we
> must ensure that there is no more outstanding rcu update before
> trying to free the network namespace.
>
> In the normal case, the rcu_barrier is called when the network namespace
> is exiting in the cleanup_net function.
>
> But when a network namespace creation fails, and the subsystems are
> undone (like the cleanup), the rcu_barrier is missing.
>
> This patch adds the missing rcu_barrier.
>
> Signed-off-by: Daniel Lezcano <dlezcano@fr.ibm.com>

Applied, thanks Daniel.

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