Subject: Re: [PATCHSET] 2.6.20-rc4-mm1-lxc2
Posted by Daniel Lezcano on Tue, 16 Jan 2007 23:48:27 GMT
View Forum Message <> Reply to Message

Cedric Le Goater wrote:
> All,
>
> We've been gathering and porting patches related to namespaces in
> a lxc patchset for a while now. Mostly working on the network
> namespace which will require some extra work to be usable.
>
> * It's available here :
>
> http://www.sr71.net/patches/2.6.20/2.6.20-rc4-mm1-lxc2/
>
> * Caveats :
>
> namespace syscalls are still under construction.
>
> network namespace is broken :
>
> . the nsproxy backpointer in net_ns is flaky.
> . the push_net_ns() and pop_net_ns() can be called under
> irq and are using current. this seems inappropriate.
> . there is a race on ->nsproxy between push_net_ns() and
> exit_task_namespaces()

Hi Dmitry,

we are experiencing NULL address access when using the nsproxy in
push_net_ns function without any unshare.

It appears the exit_task_namespace function sets current->nsproxy to
NULL and we are interrupted by an incoming packet. The netif_receive_skb
does push_net_ns(dev->net_ns). The push_net_ns function retrieves the
current->nsproxy to use it. But it was previously set to NULL by the
exit_task_namespace function.

The bug can be reproduced with the following command launched from
another host.

while $(true); do ssh myaddress ls > /dev/null && echo -n .; done

After a time (between 1 second - 3 minutes), the kernel panics.

I think this will be very hard to fix and perhaps we should redesign
some part. Instead of using nsproxy swapping, perhaps we should pass
net_ns as parameter to functions, but that will breaks a lot of API.

What is your feeling on that ?

Regards.

-- Daniel.

_____

Containers mailing list
Containers@lists.osdl.org
https://lists.osdl.org/mailman/listinfo/containers