

---

Subject: Re: [PATCH] namespaces: fix exit race by splitting exit  
Posted by [Herbert Poetzl](#) on Tue, 30 Jan 2007 02:42:28 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

On Thu, Jan 25, 2007 at 10:30:56PM -0800, Andrew Morton wrote:

> On Thu, 25 Jan 2007 23:26:59 -0600  
> "Serge E. Hallyn" <serue@us.ibm.com> wrote:  
>  
> > Fix exit race by splitting the nsproxy putting into two pieces.  
> > First piece reduces the nsproxy refcount. If we dropped the last  
> > reference, then it puts the mnt\_ns, and returns the nsproxy as a  
> > hint to the caller. Else it returns NULL. The second piece of  
> > exiting task namespaces sets tsk->nsproxy to NULL, and drops the  
> > references to other namespaces and frees the nsproxy only if an  
> > nsproxy was passed in.  
> >  
> > A little awkward and should probably be reworked, but hopefully  
> > it fixes the NFS oops.  
>  
> I'm a bit worried about jamming something like this into 2.6.20.  
> Could the usual culprits please review this carefully with  
> some urgency?

okay, after integrating this into two Linux-VServer  
branches and some testing, I can confirm that it  
\_seems\_ to fix the nfs and related issues, but still,  
I do not like it :)

here my issues with this approach:

- the code is quite hard to read and can easily  
lead to unexpected issues when spaces are  
manipulated
- it breaks the basic get/put refcounting for  
nsproxy references outside the task struct  
i.e. we had to add a vs\_put\_nsproxy() which  
does what the put\_nsproxy() did before, to  
keep and handle a reference to the nsproxy  
from the context structure
- the following scenario might become a problem  
for future spaces (especially the pid space?)

A

B

```
exit_task_namespaces_early()  
exit_task_namespaces_early()
```

```
exit_notify()
exit_task_namespaces()
```

```
-----
exit_notify()
exit_task_namespaces()
```

note: I still consider it the best available fix  
for this issues, especially as 2.6.20 is in a  
late rc stage ... but IMHO the nfs threads should  
be modified to handle the nsproxy disposal properly

> And Daniel, if you can find time to runtime test it please?

he did, looks like it works fine with vanilla too  
(even when stressing the described cornercase)

best,  
Herbert

---

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

---