Subject: Re: [PATCH 13/28] [PREP 13/14] Miscellaneous preparations in pid namespaces
Posted by Sukadev Bhattiprolu on Wed, 20 Jun 2007 21:10:13 GMT
View Forum Message <> Reply to Message

Pavel Emelianov [xemul@openvz.org] wrote:
| The most important one is moving exit_task_namespaces behind exit_notify
| in do_exit() to make it possible to see the task's pid namespace to
| properly notify the parent.

Hmm. I think we tried this once a few months ago and got a crash in nfsd
See http://lkml.org/lkml/2007/1/17/75

```
[<c01f6115>] lockd_down+0x125/0x190
[<c01d26bd>] nfs_free_server+0x6d/0xd0
[<c01d8e9c>] nfs_kill_super+0xc/0x20
[<c0161c5d>] deactivate_super+0x7d/0xa0
[<c0175e0e>] release_mounts+0x6e/0x80
[<c0175e86>] __put_mnt_ns+0x66/0x80
[<c0132b3e>] free_nsproxy+0x5e/0x60
   // exit_task_namespaces() after returning from exit_notify()
[<c011f021>] do_exit+0x791/0x810
[<c011f0c6>] do_group_exit+0x26/0x70
[<c0103142>] sysenter_past_esp+0x5f/0x85
```

exit_notify() sets current->sighand to NULL and I think lockd_down() called
from exit_task_namespaces/__put_mnt_ns() was accesssing current->sighand.

Do your other patches in this set tweak something to prevent it ?

Thats one of the reasons we had to remove pid_ns from nsproxy and use
the pid_ns from pid->upid_list[i]->pid_ns.

Suka

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