
Subject: Re: [PATCH][usercr]: Ghost tasks must be detached
Posted by [Oren Laadan](#) on Sat, 05 Feb 2011 22:33:15 GMT
[View Forum Message](#) <> [Reply to Message](#)

On 02/05/2011 05:02 PM, Oren Laadan wrote:

>
>
> On 02/05/2011 04:40 PM, Sukadev Bhattiprolu wrote:
>> Oren Laadan [orenl@cs.columbia.edu] wrote:
>> | Suka,
>> |
>> | This patch - and the corresponding kernel patch - are wrong
>>
>> Ah, I see that now.
>>
>> But am not sure about the kernel part though. We were getting a crash
>> reliably (with older kernels) because of the ->exit_signal = -1 in
>> do_ghost_task().
>
> Are we still getting it with 2.6.37 ?
>>
>> One fix I was watching for was Eric Biederman's
>>
>> <http://lkml.org/lkml/2010/7/12/213>
>>
>> which AFAICT has not been merged yet.
>
> If we need it and it isn't in mainline (any reason why ?) then
> we can just add it to our linux-cr tree, as a preparatory patch.
>
>>
>> Was there another change to 2.6.37 that would prevent the crash ?
>
> I don't know whether *that* crash still happens in 2.6.37 -
> because I still didn't test it with that kernel line back.
> (Actually, I never experienced that crash here even with
> earlier kernels).
>
>>
>> | (I should have noticed it in the review!). It turns out that
>> | ghost (and dead) tasks are _not_ auto-reaped anymore.
>> |
>> | There are only two way for tasks to be auto-reaped: one is if
>> | their parent explicitly says so in its sighand information (but
>> | then it applies to all children). The other way is if they have
>> | ->exit_signal==-1. From userspace this happens only when cloning
>> | with CLONE_THREAD. Using 0xFF for the @flags argument to clone()
>> | syscall instead results in ->exit_signal = 255 ...

```
>> |  
>> | The original motivation for this patch was:  
>> |  
>> | > The downside of marking the task detached in do_ghost_task() is that  
>> | > with current/older kernels container-init does not wait for detached  
>> | > tasks. See:  
>> | >  
>> | > http://lkml.org/lkml/2010/6/16/272  
>> | > http://lkml.org/lkml/2010/7/12/213  
>> | >  
>> | > This can lead to a kernel crash if the container-init pre-deceases a  
>> | > ghost task.  
>> |  
>> | Is this still a problem in 2.6.37 ?  
>> |  
>> Well, some inadvertent userspace changes seemed to cause the crash (or  
>> an application hang on some machines) during restart. By making those changes,  
>> I seem to be getting an application hang 5 out of 6 times even with 2.6.37,  
>> but did not get a crash. I will investigate this new hang next week.  
>  
> I'm currently chasing down a bug that causes restart to hang when  
> there are ghost/dead tasks. It may be the same one you are seeing.  
> So far I'm convinced it's userspace - working on it. Will post  
> patches once solved.
```

Actually - scratch that. The bug was related to the new pids-as-objs patches to user-cr.

Now, given the complexity of that patchset, I don't want to spend time on picking fixes that require backporting to the current user-cr. So let's try to focus testing and debugging efforts on kernel/user that includes those patches ?

Thanks,

Oren.

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