
Subject: Re: [PATCH][usercr]: Ghost tasks must be detached
Posted by [Sukadev Bhattiprolu](#) on Sat, 05 Feb 2011 21:40:32 GMT
[View Forum Message](#) <> [Reply to Message](#)

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()`.

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.

Was there another change to 2.6.37 that would prevent the crash ?

| (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.

Oren.

On 01/10/2011 08:51 PM, Oren Laadan wrote:

```
>
> Applied to user-cr.
>
> Thanks,
>
> Oren.
>
> On 12/10/2010 10:35 PM, Sukadev Bhattiprolu wrote:
>>
>> From: Sukadev Bhattiprolu <sukadev@linux.vnet.ibm.com>
>> Date: Fri, 10 Dec 2010 19:23:58 -0800
>> Subject: [PATCH 1/1] Ghost tasks must be detached
>>
>> Ghost processes are created only to help restore orphaned sessions/pgrps.
>> As such once the session/pgrp is created the ghost must not send another
>> SIGCHLD to the parent but exit silently. So create such tasks as
>> "detached".
>>
>> See also:
>>
>> https://lists.linux-foundation.org/pipermail/containers/2010-December/026076.html
>>
>> Signed-off-by: Sukadev Bhattiprolu <sukadev@linux.vnet.ibm.com>
>> ---
>> restart.c | 7 +++++++
>> 1 files changed, 7 insertions(+), 0 deletions(-)
>>
>> diff --git a/restart.c b/restart.c
>> index 9fb5e9f..d7ba26b 100644
>> --- a/restart.c
>> +++ b/restart.c
>> @@ -1744,6 +1744,13 @@ static pid_t ckpt_fork_child(struct ckpt_ctx *ctx, struct task *child)
>>  flags |= CLONE_THREAD | CLONE_SIGHAND | CLONE_VM;
>>  else if (child->flags & TASK_SIBLING)
>>  flags |= CLONE_PARENT;
>> + else if (child->flags & (TASK_GHOST|TASK_DEAD)) {
>> + /*
>> +  * Ghosts must vanish silently (without signalling parent)
>> +  * when they are done.
```

```
| >> + */
| >> + flags = 0xFF;
| >> + }
| >>
| >> memset(&clone_args, 0, sizeof(clone_args));
| >> clone_args.nr_pids = 1;
| > _____
| > Containers mailing list
| > Containers@lists.linux-foundation.org
| > https://lists.linux-foundation.org/mailman/listinfo/containers
| >
```

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