

---

Subject: Re: + remove-the-likelypid-check-in-copy\_process.patch added to -mm tree

Posted by [ebiederm](#) on Sat, 17 Mar 2007 17:01:50 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Oleg Nesterov <[oleg@tv-sign.ru](mailto:oleg@tv-sign.ru)> writes:

> On 03/17, Oleg Nesterov wrote:

>>

>> > Well the initial kernel process does not have a struct pid so when

>> > it's children start doing:

>> > attach\_pid(p, PIDTYPE\_PGID, task\_group(p));

>> > attach\_pid(p, PIDTYPE\_SID, task\_session(p));

>> > We will get an oops.

>>

>> So far this is the only reason to have init\_struct\_pid. Because the

>> boot CPU (swapper) forks, right?

>

> Damn. I am afraid I was not clear again :) Not init\_struct\_pid, but

>

> + .pids = { \

> + [PIDTYPE\_PID] = INIT\_PID\_LINK(PIDTYPE\_PID), \

> + [PIDTYPE\_PGID] = INIT\_PID\_LINK(PIDTYPE\_PGID), \

> + [PIDTYPE\_SID] = INIT\_PID\_LINK(PIDTYPE\_SID), \

> + }, \

>

> for INIT\_TASK().

>

>> > So a dummy unhashed struct pid was added for the idle threads.

>> > Allowing several special cases in the code to be removed.

>> >

>> > With that chance the previous special case to force the idle thread

>> > init session 1 pgrp 1 no longer works because attach\_pid no longer

>> > looks at the pid value but instead at the struct pid pointers.

>> >

>> > So we had to add the \_\_set\_special\_pids() to continue to keep init

>> > in session 1 pgrp 1. Since /sbin/init calls setsid() that our setting

>> > the sid and the pgrp may not be strictly necessary. Still is better

>> > to not take any chances.

>>

>> Yes, yes, I see. But my (very unclear, sorry) question was: shouldn't we

>> change INIT\_SIGNALS then? /sbin/init inherits ->pgrp == ->\_session == 1,

>> in that case \_\_set\_special\_pids(1,1) does nothing.

>

> ... and thus /sbin/init remains attached to the .pids above, no?

The problem is that we dynamically allocate the struct pid for  
pid\_t == 1 when we fork init.

Which means we don't have access to it at compile time so we can no longer make INIT\_SIGNALS set ->grp == ->session == 1.

Eric

---

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>

---