Subject: Re: [PATCH 1/2] pidns: Don't allow new pids after the namespace is dead.
Posted by Oleg Nesterov on Thu, 17 Feb 2011 20:54:58 GMT
View Forum Message <> Reply to Message

On 02/17, Daniel Lezcano wrote:
>
> On 02/15/2011 07:30 PM, Oleg Nesterov wrote:
>> On 02/15, Daniel Lezcano wrote:
>>> In the case of unsharing or joining a pid namespace, it becomes
>>> possible to attempt to allocate a pid after zap_pid_namespace has
>>> killed everything in the namespace. Close the hole for now by simply
>>> not allowing any of those pid allocations to succeed.
>> Daniel, please explain more. It seems, a long ago I knew the reason
>> for this patch, but now I can't recall and can't understand this change.
>
> The idea behind unsharing the pid namespace is the current pid is not
> mapped in the newly created pid namespace and appears as the pid 0.

Well, not exactly afaics... but doesn't matter.

> When
> it forks, the child process becomes the init pid of the new pid
> namespace.

Yes, I see. And this is what I personally dislike. Because, iow,
unshare(PID) changes current->nspory->pid_ns to affect the behaviour
of copy_process() , this really looks like "action at a distance" to
me. Too subtle and fragile. But, once again, this is just imho, feel
free to ignore.

> When this pid namespace dies because the init pid exited, the
> parent process (aka pid 0) can no longer fork because the pid namespace
> is flagged dead. This is what does this patch.

OK, thanks. I seem to understand. May be ;)

I'd suggest you to add this explanation to the changelog.

>>> --- a/include/linux/pid_namespace.h
>>> +++ b/include/linux/pid_namespace.h
>>> @@ -20,6 +20,7 @@ struct pid_namespace {
>>>   struct kref kref;
>>>   struct pidmap pidmap[PIDMAP_ENTRIES];
>>>   int last_pid;
>>> + atomic_t dead;
>> Why atomic_t? It is used as a plain boolean.
>>
>> And I can't unde

>
> I think Eric used an atomic because it is lockless with alloc_pid vs
> zap_pid_ns_processes.

Can't understand...

But anyway, I strongly believe atomic_t buys nothing in this patch.
May be it is needed for the next changes, I dunno.

Oleg.

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