Subject: Re: [PATCH 3/15] kern_siginfo helper
Posted by Sukadev Bhattiprolu on Tue, 31 Jul 2007 00:21:38 GMT
View Forum Message <> Reply to Message

Pavel Emelianov [xemul@openvz.org] wrote:
| Oleg Nesterov wrote:
| >On 07/26, Pavel Emelyanov wrote:
| >>TODO: This is more an exploratory patch and modifies only
| >>interfaces
| >> necessary to implement correct signal semantics in pid namespaces.
| >>
| >> If the approach is feasible, we could consistently use 'kern_siginfo'
| >> in other signal interfaces and possibly in 'struct sigqueue'.
| >>
| >> We could modify dequeue_signal() to directly work with 'kern_siginfo'
| >> and remove dequeue_signal_kern_info().
| >
| >Well... I know, it is very easy to blame somebody else's patch, and
| >probably
| >my taste is not good...
| >
| >But honestly, I personally think this approach is a horror, and any
| >alternative
| >is better :)

Hmm. My reasoning was that since siginfo_t was directly "shared" with
user space, extending it even to add a flag was pain.

| >
| >I'd rather change dequeue_signal() so that it takes "struct sigqueue *"
| >parameter instead of "siginfo_t *", or add a new "int *flags".

My earlier version to Containers@ passed in "int *signal_cinit" couple
of levels down and used that in get_signal_to_deliver() but that
looked ugly :-) Passing in sigqueue to dequeue_signal() may be better,
but anyway...

| >
| >OK, this doesn't work anyway, we should do something different. Perhaps
| >just do all checks on sender's side.
|
| Yes. Signal handling in namespaces turned out to be the most complicated
| part of the set. I start thinking to drop this part till we have the "core"
| in -mm tree. Suka, what do you think?

Yes. Lets focus on the core for now and allow privileged user in a child-ns
to terminate the container-init. I will try the signal approach from sender
side also.

|
| >It is a bit strange that this patch is 3/15, and the rest bits in 11/15,
| >not very convenient for the review.
|
| Well, I thought that a split like
| 1. patches for kernel to prepare it for the set
| 2. the set itself
| is the best to review. Maybe I was wrong, but how to make this then?
| E.g. I have a MS_KERNOUNT patch, but its changes are used *much* later.
|
| >Oleg.
| >
| >
|
| Thanks,
| Pavel
