Subject: Re: [PATCH 11/13] Changes to show virtual ids to user
Posted by ebiederm on Fri, 25 May 2007 15:48:29 GMT
View Forum Message <> Reply to Message

Pavel Emelianov <xemul@openvz.org> writes:

> Eric W. Biederman wrote:
>> Pavel Emelianov <xemul@sw.ru> writes:
>>
>>> That's true. Sending of signal from parent ns to children
>>> is tricky question. It has many solutions, I wanted to
>>> discuss which one is better:
>>
>> With unix domain sockets and the like it is conceivable we get
>> a pid transfer from one namespace to another and both namespaces
>> are leaf namespaces. I don't remember we can get a leaf to leaf
>> transfer when sending signals.
>
> We should not allow any transfer from leaf NS to leaf NS.
> Should I explain why?

In a checkpointable context it is a bad thing, and we can prevent it
by carefully setting up all of the namespaces.

However it is a fundamental possibility that exists, and because we
can avoid it with careful setup. I don't see a reason to deny it
if something was either inadvertantly or explicitly causes it
to happen.

Do you have another reason for denying the transfer that I'm
not thinking of?

>>
>> The worst case I can see with pid == 0. Is that it would be a bug
>> that we can fix later. For other cases it would seem to be a user
>> space API thing that we get stuck with for all time.
>
> We cannot trust userspace application to expect some pid other than
> positive. All that we can is either use some always-absent pid or
> send the signal as SI_KERNEL.
>
> Our experience show that making decisions like above causes random
> applications failures that are hard (or even impossible) to debug.

Ok. So I guess I see what you are proposing is picking an arbitrary
pid, say pid == 2, and reserving that in all pid namespaces and using
it when we have a pid that does not map to a specific namespace. I'm

fine with that.

All I care about is that we have a solution, preferably simple,
to the non-mapped pid problem.

Eric
