Subject: Re: The issues for agreeing on a virtualization/namespaces implementation.
Posted by Herbert Poetzl on Mon, 20 Feb 2006 15:16:14 GMT

View Forum Message <> Reply to Message

On Mon, Feb 20, 2006 at 05:26:13PM +0300, Kirill Korotaev wrote:
>> as does Linux-VServer currently, but do you have
>> any proof that putting all the fields together in
>> one big structure actually has any (dis)advantage
>> over separate structures?

> have no proof and don't mind if there are many pointers.
> Though this doesn't look helpful to me as well.

well, my point is just that we don't know yet
so we should not favor one over the other, just
because somebody did it like that and it didn't
hurt :)

>>> mmm, how do you plan to pass additional flags to clone()?
>>> e.g. strong or weak isolation of pids?

>> do you really have to pass them at clone() time?
>> would shortly after be more than enough?
>> what if you want to change those properties later?

> I don't think it is always suiatable to do configuration later.
> We had races in OpenVZ on VPS create/stop against exec/enter etc.
> (even introduced flag is_running).
> So I have some experience to believe it will be painfull place.

well, Linux-VServer uses a state called 'setup'
which allows to change all kinds of things before
the guest can be entered, this state is changed
as the last operation of the setup, which in turn
drops all the capabilities and makes the guest
visible to the outside ...

works quite well and seems to be free of those
races you mentioned ...

>>> this syscalls will start handling this new namespace and that's all.
>>> this is not different from many syscalls approach.
>> well, let's defer the 'how amny syscalls' issue to
>> a later time, when we know what we want to implement :)
> agreed.

btw, maybe it's just me, but would it be possible

to do the email quoting like this:

>>> Text

instead of

> >>Text

TIA,
Herbert

> Kirill