Subject: Re: Which of the virtualization approaches is more suitable for kernel?
Posted by ebiederm on Mon, 27 Feb 2006 21:56:37 GMT
View Forum Message <> Reply to Message

Dave Hansen <haveblue@us.ibm.com> writes:

> On Mon, 2006-02-27 at 14:14 -0700, Eric W. Biederman wrote:
>> I like the namespace nomenclature. (It can be shorted to _space or _ns).
>> In part because it shortens well, and in part because it emphasizes that
>> we are *just* dealing with the names.
>
> When I was looking at this, I was pretending to be just somebody looking
> at sysv code, with no knowledge of containers or namespaces.
>
> For a person like that, I think names like _space or _ns are pretty much
> not an option, unless those terms become as integral to the kernel as
> things like kobjects.

To be clear I was talking name suffixes. So ipc_space certainly conveys
something, and even ipc_ns may be ok.

>> You split the resolution at just ipc_msgs. When I really think it should
>> be everything ipcs deals with.
>
> This was just the first patch. :)

:)

Just wanted to make certain we agreed on the scope.

>> Performing the assignment inside the tasklist_lock is not something we
>> want to do in do_fork().
>
> Any particular reason why? There seem to be a number of things done in
> there that aren't _strictly_ needed under the tasklist_lock. Where
> would you do it?

Well all of the other things we can share or not share are already
outside of the tasklist_lock.

We may not be quite minimal but we actually are fairly close to minimal
inside the tasklist_lock.

>> So it looks like a good start. There are a lot of details yet to be filled
>> in, proc, sysctl, cleanup on namespace release. (We can still provide
>> the create destroy methods even if we don't hook the up).
>
> Yeah, I saved shm for last because it has the largest number of outside

> interactions. My current thoughts are that we'll need _contexts or
> _namespaces associated with /proc mounts as well.

Yes. I think the easy way to handle this is to have a symlink
from /proc/sysvipc to /proc/self/sysvipc. And then we have a per
process reporting area.

That preserves all of the old programs but enables us to get the
information out.

>> I think in this case I would put the actual namespace structure
>> definition in util.h, and just put a struct ipc_ns in sched.h.
>
> Ahhh, as in
>
> struct ipc_ns;
>
> And just keep a pointer from the task? Yeah, that does keep it quite
> isolated.

Yep.

Eric
