
Subject: Re: [PATCH] Switch nfs/callback.c to using struct pid, not pid_t
Posted by [Trond Myklebust](#) on Wed, 29 Aug 2007 14:10:41 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Wed, 2007-08-29 at 14:52 +0100, Christoph Hellwig wrote:

> On Wed, Aug 29, 2007 at 05:36:24PM +0400, Pavel Emelyanov wrote:

> > Pid namespaces make it dangerous to use pid and tgid values

> > when run in some namespace. The struct pid itself is going

> > to be the only way for working with task pids, so make the

> > nfs callback thread use it.

> >

> > Since nfs_callback_info.pid is set to current's one and reset

> > on the thread exit, it is safe not to get the struct pid.

> >

> > Since this pid is used later under lock_kernel() w/o sleeping

> > operations, checking for i to be not NULL and killing the

> > thread with kill_pid() is safe.

>

> NACK. This just makes the code even more obscure. Please get rid

> of the pid references entirely and convert the code to the kthread

> API.

That would require converting the full sunrpc server code to use
kthreads, which again means changing nfsd, and lockd too.

I'm not saying that is a bad thing, but it is nontrivial to do. In
particular, kthread's abominable shutdown mechanism simply does not work
or scale when the thread is listening for new requests in svc_rcv().

Cheers

Trond

Containers mailing list

Containers@lists.linux-foundation.org

<https://lists.linux-foundation.org/mailman/listinfo/containers>

Subject: Re: [PATCH] Switch nfs/callback.c to using struct pid, not pid_t
Posted by [Christoph Hellwig](#) on Wed, 29 Aug 2007 14:18:23 GMT

[View Forum Message](#) <> [Reply to Message](#)

On Wed, Aug 29, 2007 at 10:10:41AM -0400, Trond Myklebust wrote:

> On Wed, 2007-08-29 at 14:52 +0100, Christoph Hellwig wrote:

> > On Wed, Aug 29, 2007 at 05:36:24PM +0400, Pavel Emelyanov wrote:

> > > Pid namespaces make it dangerous to use pid and tgid values

> > > when run in some namespace. The struct pid itself is going

> > > to be the only way for working with task pids, so make the

> > > nfs callback thread use it.
> > >
> > > Since nfs_callback_info.pid is set to current's one and reset
> > > on the thread exit, it is safe not to get the struct pid.
> > >
> > > Since this pid is used later under lock_kernel() w/o sleeping
> > > operations, checking for i to be not NULL and killing the
> > > thread with kill_pid() is safe.
> >
> > NACK. This just makes the code even more obscure. Please get rid
> > of the pid references entirely and convert the code to the kthread
> > API.
>
> That would require converting the full sunrpc server code to use
> kthreads, which again means changing nfsd, and lockd too.
>
> I'm not saying that is a bad thing, but it is nontrivial to do. In
> particular, kthread's abominable shutdown mechanism simply does not work
> or scale when the thread is listening for new requests in svc_recv().

Actually converting callback.c is not that bad, it just means splitting
up svc_create_thread. Only problem is whether we want to allow SIGKILL
from serspace to terminate the thread aswell, in which case we still
need changes to the core kthread code which I'm waiting for someone
for the containers crowd to finally implement as it's needed in
various other places.

I'll try to send out my svc_create_thread splitup soon because it's
actually a nice cleanup all by itself.

Containers mailing list
Containers@lists.linux-foundation.org
<https://lists.linux-foundation.org/mailman/listinfo/containers>
