Subject: Re: [PATCH v3] SUNRPC: protect service sockets lists during per-net shutdown
Posted by bfields on Mon, 20 Aug 2012 14:56:56 GMT
View Forum Message <> Reply to Message

On Mon, Aug 20, 2012 at 03:05:49PM +0400, Stanislav Kinsbursky wrote:

> >Looking back at this:
> >
> > - adding the sv_lock looks like the right thing to do anyway
> > independent of containers, because svc_age_temp_xprts may
> > still be running.
> >
> > - I'm increasingly unhappy about sharing rpc servers between
> > network namespaces. Everything would be easier to understand
> > if they were independent. Can we figure out how to do that?
> >
>
> Could you, please, elaborate on your your unhappiness?

It seems like you're having to do a lot of work on each individual rpc server (callback server, lockd, etc.) to make per-net startup/shutdown work. And then we still don't have it quite right (see the shutdown races).)

In general whenever we have the opportunity to have entirely separate data structures, I'd expect that to simplify things: it should eliminate some locking and reference-counting issues.

> I.e. I don't like it too. But the problem here, is that rpc server
> is tied with kernel threads creation and destruction. And these
> threads can be only a part of initial pid namespace (because we have
> only one kthreadd). And we decided do not create new kernel thread
> per container when were discussing the problem last time.

There really should be some way to create a kernel thread in a specific namespace, shouldn't there?

Until we have that, could the threads be taught to fix their namespace on startup?

--b.