

---

Subject: Re: [PATCH v3] SUNRPC: protect service sockets lists during per-net shutdown

Posted by [bfields](#) on Tue, 21 Aug 2012 12:25:44 GMT

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

---

On Tue, Aug 21, 2012 at 01:28:00PM +0400, Stanislav Kinsbursky wrote:

> >On Mon, Aug 20, 2012 at 07:11:00PM +0400, Stanislav Kinsbursky wrote:  
> >>Currently, when you call kthread\_create(), you add new job to  
> >>kthreadd queue. Kthreadd is unique, starts right after init and  
> >>lives in global initial environment. So, any kthread inherits  
> >>namespaces from it.  
> >>Of course, we can start one kthread per environment and change it's  
> >>root or even network namespace in kthread function. But pid  
> >>namespace of this kthread will remain global.  
> >  
> >OK. But the current implementation will leave all the server threads in  
> >the initial pid namespace, too.  
> >  
> >>It looks like not a big problem, when we shutdown kthread by some  
> >>variable. But what about killable nfsd kthreads?  
> >  
> >And we're stuck with that problem either way too, aren't we?  
> >  
> >  
> >Yes, we are. But at least we are avoiding patching of task subsystem.  
> >  
> >>1) We can't kill them from nested pid namespace.  
> >>2) How we will differ nfsd kthreads in initial pid namespace?  
> >  
> >I have to admit for my purposes I don't care too much about pid  
> >namespaces or about signalling server threads. It'd be nice to get  
> >those things right but it wouldn't bother me that much not to.  
> >  
> >Another stupid idea: can we do our own implementation of something like  
> >kthreadd just for the purpose of starting rpc server threads? It  
> >doesn't seem that complicated.  
> >  
> >  
> >Gm...  
> >This idea is not stupid. If I understand you right, you suggest to  
> >implement a service per network namespace (i.e. not only data, but  
> >also threads)?

Some way or another, yes, entirely separate threads for the different namespaces would be clearer, I think.

And if we can't get them in the right pid namespaces, I'm not sure I

care.

--b.

---