
Subject: Re: Q: How complete is the pid namespace in mainline

Posted by [ebiederm](#) on Fri, 26 Oct 2007 18:17:33 GMT

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

sukadev@us.ibm.com writes:

>

> Dave had suggested we print a warning the first time a container-init forks()

> without a handler for a fatal signal. I was planning on adding that as

> patch 4 of the signal patch set and get some feedback.

Yes. How to cleanly handle signalling of container init is a tricky one. It does sound like you have made a reasonable start there.

Suka it is a lot more than that. How much more I'm not certain of. I suspect the only way to find the rest of the cases is just go through the code with a fine tooth comb and read and look.

So far doing that it has not at all hard for me to find either bugs or places where the implementation can be improved.

Currently we have little things like kill(-1,...) signalling the wrong set of processes, and a couple of proc bugs.

That autofs and coda out on the fringe don't quite do the right thing in the presence of multiple pid namespaces isn't a big surprise, little details like that are easy to overlook.

We of course still have the kthread issue where we can get kernel threads trapped and we have kernel threads calling kill_proc, keeping us from removing it.

There is all cap_set_all which isn't filtering by pid namespace.

Then we have the unix domain sockets that don't appear to do the right thing when passing credentials across pid namespaces. I think we may have the same issues with signals as well.

Anyway I can find a lot issues like that without trying very hard. Which suggests to me that there are issues that I'm missing that are out there as well.

So it appears there is lots of cleanup work to do.

Eric

Containers mailing list

Containers@lists.linux-foundation.org

