
Subject: Re: [PATCH] namespaces: introduce sys_hijack (v4)

Posted by [serue](#) on Tue, 16 Oct 2007 18:57:37 GMT

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

Quoting Paul Menage (menage@google.com):

> On 10/16/07, Serge E. Hallyn <serue@us.ibm.com> wrote:

> > pid, but wasn't sure how best to identify the cgroup. Originally I was
> > more worried about pid exiting/wraparound, but then decided that with a
> > real container the container_init can't go away until the container goes
> > away anyway.

>

> For those "real containers" that have init. Not everything is going to
> need that level of virtualization, particularly if you're primarily
> interested in isolation.

Currently every pid namespace's pid==1 must stick around as long as the
pid namespace does. If you kill the pid==1, all processes in the
container are killed.

> > Anyway, I can go ahead and add 'int which' to the parameter list now,
> > and leave the details of how to specify a cgroup for later. That way at
> > least the api won't fundamentally change again.

>

> Great, thanks.

Since the goal here is to get the API right, do you know how we expect
to send the cgroup in? A string?

Currently my prototype is

```
+asmlinkage long sys_hijack(unsigned long flags, int which, pid_t pid,  
+                          const char __user *cgroup);
```

But that doesn't seem quite right. At that point we just ditch 'which'
and use cgroups if it's not NULL, use pid otherwise...

Thoughts?

thanks,
-serge

Containers mailing list

Containers@lists.linux-foundation.org

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