Subject: Re: [ckrm-tech] [PATCH 7/7] containers (V7): Container interface to nsproxy subsystem
Posted by Paul Menage on Thu, 05 Apr 2007 14:13:37 GMT
View Forum Message <> Reply to Message

On 4/5/07, Srivatsa Vaddagiri <vatsa@in.ibm.com> wrote:
> > If the container directory were to have no refcount on the nsproxy, so
> > the initial refcount was 0,
>
> No it should be 1.
>
> mkdir H1/foo
> rcfs_create()
> ns = dup_namespaces(parent);
>
> ....
> dentry->d_fsdata = ns;
>
> ns should have a refcount of 1 to begin with.

Right - that's my point, you're effectively passing the initial refcount of the nsproxy to the container directory's d_fsdata reference.

>
> - refcount of a nsproxy attached to a directory dentry can never
> fall to zero because of tasks coming in and out. The only
> way for the refcount of such nsproxies to fall to zero and
> hence trigger their destruction is thr' the rmdir i/f.
>
> - New nsproxies derived from the base directory nsproxy
> can have their's refcount go to zero as tasks exit or move
> around and hence they will be destroyed.
>
> Does that sound like correct behavior?

Sounds good.

>
> > Possibly - there are two choices:
> >
> > 1) expose a refcount to them directly, and just interrogate the
> > refcount from the generic code to see if it's safe to delete the
> > directory
> >
> > 2) have a can_destroy() callback with well defined semantics about
> > when it's safe to take refcounts again - it's quite possible that one
> > subsystem can return true from can_destroy() but others don't, in

> > which case the subsystem can become active again.
>
> Lets go back to the f_bc example here for a moment. Lets say T1 was in C1 and
> opened file f1. f1->f_bc points to C1->beancounter.
>
> T1 moves from C1 -> C2, but f1 is not migrated.
> C1->beancounter.count stays at 1 (to account for f1->f_bc).
>
> File f1 is closed. C1->beancounter.count becomes zero.
>
> Now user issues rmdir C1. If rmdir finds (after taking manage_mutex that
> is)
>
> - zero tasks in C1
> - zero refcount in C1->beancounter
>
> why is it not safe to assume that C1->beancounter.count will continue to
> stay zero?
>
> Basically I am struggling to answer "How can a zero refcount (beancounter)
> object go non-zero when zero tasks are attached to it" ..

In that case, I think you're fine. Your last posted patches didn't
provide a way to check for that, though, as far as I could see.

Paul
