Subject: Re: [PATCH] new cgroup controller "fork"
Posted by Paul Menage on Fri, 18 Feb 2011 00:59:51 GMT
View Forum Message <> Reply to Message

On Thu, Feb 17, 2011 at 5:31 AM, Max Kellermann <mk@cm4all.com> wrote:
> Can limit the number of fork()/clone() calls in a cgroup. It is
> useful as a safeguard against fork bombs.

I'd be inclined to simplify this a bit - avoid impacting the fork()
path twice, with cgroup_fork_pre_fork() and cgroup_fork_fork() and
just do the checks and decrements in a single pass. (In the event of
hitting a limit, you may need to make another partial pass up the tree
to restore the charged fork attempts)

Yes, it's true that you might charge for a fork() that later failed
for some other reason, but this will very rare (except on a machine
that's already screwed for other reasons) so that I don't think anyone
would complain about it. Especially if you explicitly document
"fork.remaining" as number of permitted "fork attempts".

Also, it would be slightly clearer to use fork_cgroup_* rather than
cgroup_fork_* - this makes it clearer that it's part of a cgroups
subsystem called "fork", rather than part of the cgroups core
framework.

I don't think that you need to make your spinlock IRQ-safe - AFAICS
nothing accesses it from the interrupt path.

Paul

_____
Containers mailing list
Containers@lists.linux-foundation.org
https://lists.linux-foundation.org/mailman/listinfo/containe rs