Subject: Re: Re: containers and cgroups mini-summit @ Linux Plumbers
Posted by ebiederm on Thu, 26 Jul 2012 19:38:16 GMT
View Forum Message <> Reply to Message

Serge Hallyn <serge.hallyn@canonical.com> writes:
> (Sorry, please disregard my last email :)
>
> Yes, what we do now in ubuntu quantal is the bind mounts you mention,
> and only optionally (using a startup hook).
> Each container is brought up in say
> /sys/fs/cgroup/devices/lxc/container1/container1.real, and that dir is
> bind-mounted under /sys/fs/cgroup/devices in the guest. The guest
> is not allowed to mount cgroup fs himself.
>
> It's certainly not ideal (and in cases where cgroup allows you to
> raise your own limits, worthless). The 'fake cgroup root' has been
> mentioned before to address this. Definately worth discussing.

It is going to be interesting to see how all of the unprivileged
operations work when the user-namespaces start allowing unprivileged
users to do things (3.7 timeframe I hope).

I can see it making things both easier and harder. I would hope not
actually being root will make it easier to keep from raising your own
limits.

Running some operations as non-root will catch other places off guard
where people were definitely expecting nothing of the kind.

There are a couple of networking memory limits exposed through sysctl
that I don't expect we want everyone changing, that I need to figure out
how to separate out from the rest. A concept that hasn't existed
before.

Eric