
Subject: Re: How to draw values for /proc/stat

Posted by [Glauber Costa](#) on Sun, 11 Dec 2011 20:48:29 GMT

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

On 12/11/2011 08:11 PM, KOSAKI Motohiro wrote:

>
>>>> IOW a /proc namespace coupled to cgroup scope would do what you want.
>>>> Now my head hurts..
>>>
>>> Mine too. The idea is good, but too broad. Boils down to: How do you
>>> couple them? And none of the methods I thought about seemed to make any
>>> sense.
>>>
>>> If we really want to have the values in /proc being opted-in, I think
>>> Kamezawa's idea of a mount option is the winner so far.
>
> > diff --git a/include/linux/cgroup.h b/include/linux/cgroup.h
> > index 1b7f9d5..f0bc2e9 100644
> > --- a/include/linux/cgroup.h
> > +++ b/include/linux/cgroup.h
> > @@ -158,6 +158,7 @@ enum {
> > * Clone cgroup values when creating a new child cgroup
> > */
> > CGRP_CLONE_CHILDREN,
> > + CGRP_PROC_OVERLAY,
> > };
>
> I'm not cgroup expert, but I doubt it is mount option. I suspect it's
> cgroup option. That's said, if we have following two directories,

Actually, the way I proposed, you have both ways. The mount option is more a default value for convenience, that is effective until you change a file. That's the same way as clone_children already do, and I believe it to be a sane thing.

> /cgroup-for-virtualization
> /cgroup-for-resource-management
>
> are both directory affected the overlay flag?

It depends. The flag is per-cgroup, therefore per-directory. So even if you set the mount option, you can override it in an individual cgroup.

> I don't think it is not
> optimal. Why? we must care some system software (e.g. kvm, systemd) are
> using cgroup internally and we expected this trend will grow more.

As I said before, each directory has its own files, so in a standard

system, we would be more than happy to set it to 1 in the cgroups corresponding to our containers, and leave the rest of the world alone.

> So, I doubt namespace issue can be solved by such tiny patch.

>

I don't fully get what you mean here
