Subject: Re: [PATCHSET 3/4] sysfs: divorce sysfs from kobject and driver model
Posted by Tejun Heo on Sat, 29 Sep 2007 22:06:23 GMT
View Forum Message <> Reply to Message

Hello, Eric.

Eric W. Biederman wrote:
> Mostly I am thinking that any non-object model users should have
> their own dedicated wrapper layer. To help keep things consistent
> and to make it hard enough to abuse the system that people will
> find that it is usually easier to do it the right way.

Hmmm... I think most current non-driver-model sysfs users are deep in
kernel anyway, but I think not exporting sysfs interface at all might be
a bit too restrictive. I think we need to examine the current
non-driver-model sysfs users thoroughly to determine what to do about
this. But, yes, I do agree that we need to put restrictions one way or
the other.

> Does it look like we can resolve Tejun's work for 2.6.24?
> If not does it make sense to push my patches that allow
> multiple mounts of sysfs for 2.6.24? So I can allow
> network namespaces in the presence of sysfs.
>
> Outside of sysfs and the device model I'm only talk maybe 30 lines
> of code...   So I could easily merge that patch later in the
> merge window after the other pieces have gone in.

I think it would be better if namespace comes after interface update and
other new features, especially symlink renaming, but, under the current
circumstance, it might delay namespace unnecessarily and I have no
problem with your patches going in first. My concerns are...

* Do you think you can use new rename implementation contained in this
series? It borrows basic ideas from the implementation you used for
namespace but is more generic. It would be great if you can use it
without too much modification.

* I'm still against using callbacks to determine namespace tags because
callbacks need to be coupled with sysfs internals more tightly and are
more difficult to grasp interface-wise.

> - Farther down the road we have the device namespace.
> The bounding requirements are:
> - We want to restrict which set of devices a subset of process
>   can access.
> - When we migrate an application we want to preserve the device
>   numbers of all devices that show up in the new location.

> So filesystems whose block devices reside on a SAN, ramdisks,
> ttys, etc.
> Other devices that really are different we can handle with
> hotplug remove and add events, during the migration.
>
> So while there is lower hanging fruit the requirements for a
> device namespace are becoming clear, and don't look like something
> we will ultimately be able to dodge.
>
> For sysfs the implication is that we will need to filter the
> hotplug events based upon the device namespace of the recipient, and
> we will need to restrict the set of devices that show up in sysfs
> based on who mounts it (as the prototype patches with the network
> namespace are doing).
>
> Also fun is that the dev file implementation needs to be able to
> report different major:minor numbers based on which mount of
> sysfs we are dealing with.

Ah... Coming few months will be fun, won't they? :-)

--
tejun

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