Subject: Re: [PATCHSET 3/4] sysfs: divorce sysfs from kobject and driver model
Posted by Greg KH on Fri, 05 Oct 2007 06:23:02 GMT
View Forum Message <> Reply to Message

On Thu, Sep 27, 2007 at 01:25:48PM -0600, Eric W. Biederman wrote:
>
> I still need to look at the code in detail but I have some concerns
> I want to inject into this conversation of future sysfs architecture.
>
> - If we want to carefully limit sysfs from going to wild code review
> is clearly not enough. We need some technological measures to
> assist us. As the experience with sysctl has shown.

I totally agree. You should see the ways that people have tried to circumvent the current kobject/sysfs code over the past years. It's so scary it's not even funny...

> - The network namespace work scheduled to be merged in 2.6.24 is
> currently has a dependency in Kconfig that is "&& !SYSFS"
> because sysfs is currently very much a moving target.
>
> 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 would be interested in seeing what your patches look like. I don't think that we should take any more sysfs changes for 2.6.24 as we do have a lot of them right now, and I don't think that Tejun and I agree on the future direction of the outstanding ones just yet.

But I don't think that your multiple-mount patches could make it into .24, unless .23 is still weeks away.

> - 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.

That's reasonable.

> - 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).

That is going to be interesting to see how you come up with a way to do
hat.

> 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.

Um, no, that's not going to happen. /dev/sda will _always_ have the
same major:minor number, as defined by the LSB spec. You can not break
that at all. So while you might not want to show all mounts
/sys/devices/block/sda/ the ones that you do, will all have the LSB
defined major:minor number assigned to it.

thanks,

greg k-h

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