

---

Subject: Re: Re: [PATCHSET 3/4] sysfs: divorce sysfs from kobject and driver model

Posted by [ebiederm](#) on Fri, 05 Oct 2007 13:24:55 GMT

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

---

Kirill Korotaev <dev@sw.ru> writes:

- > Imho environments to be migratable should have no direct access to the devices.
- > You can use any of stacked virtual filesystems to hide real
- > device from container.
- > You will have problems much bigger than this one otherwise
- > (imagine access to video, sound etc.)

What I am primarily concern about is when you can make the case that the hardware we are talking is present before and after the migration.

When you are directly accessing a device. For times when it makes sense to directly access hardware in a container (think infiniband OS-bypass NICs). We need to tell user space that the device was unplugged and another one was plugged in. If user space can cope with that things should continue to work.

There are some very specific cases that we can support:

- Stateless devices like /dev/zero and dev/random.
- Virtual devices like ttys, ramdisks, loop devices
- Remote block devices like SCSI disks on a san, iSCSI, nbd, ATAoE.
- Local pseudo block devices like the backing devices for virtual filesystems.

There are very specific limits in which this can work and be useable, and I don't claim to have looked at all of the details, but for the block device case in particular we export the block device number to user space in stat. There are some common applications which do memorize stat data for files things like: git, incremental backup software, and intrusion detection software.

Frankly the times when this matters is rare enough I don't put a big priority on getting this done quickly. However a strong case has been made for all of this filtering so we can run things like udev in a container. Initially I only expect stateless character devices and ttys to be allowed in a namespace, and I don't have a clue what device number we will use in st\_dev for stat in the case our block device isn't in the user space interface. I just know that it sounds like where we want to be eventually and thinking about it now isn't a problem.

Eric

---

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

---