
Subject: Re: [PATCH 01/11] sysfs: Support for preventing unmounts.
Posted by [ebiederm](#) on Wed, 18 Jun 2008 20:12:27 GMT
[View Forum Message](#) <> [Reply to Message](#)

Dave Hansen <dave@linux.vnet.ibm.com> writes:

> On Wed, 2008-06-18 at 19:07 +0200, Benjamin Thery wrote:
>> To support mounting multiple instances of sysfs occasionally I
>> need to walk through all of the currently present sysfs super blocks.
>
> I know you may have addressed this before, but I forgot and it didn't
> make it into the changelogs.
>
> Why are you doing this again? It seems like an awfully blunt
> instrument.

So the fundamentals.

- The data in sysfs fundamentally changes behind the back of the VFS and we need to keep the VFS in sync. Essentially this is the distributed filesystem problem.
- In particular for sysfs_rename and sysfs_move_dir we need to support finding the dcache entries and calling d_move. So that the dcache does not get into an inconsistent state. Timeouts and invalidates like NFS uses are to be avoided if at all possible.
- Coming through the vfs we are guaranteed that the filesystem will not be unmounted while we have a reference on a dentry, and with multiple mounts we do not get that guarantee. Therefore to get that guarantee for all of the superblocks we need the blunt instrument.
- Since mount/unmount are rare blocking them is no big deal.

I believe any distributed filesystem that is together enough to tell us about renames (so we can update the dcache) instead of doing the NFS timeout will need the ability to block mount/unmount while it is executing d_move.

Currently sysfs does not need to block mounts only because we perform an internal mount and then never unmount sysfs.

Eric

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