Subject: Re: [PATCH 06/14] sysfs: Rewrite sysfs_get_dentry
Posted by ebiederm on Fri, 03 Aug 2007 19:29:19 GMT
View Forum Message <> Reply to Message

Cornelia Huck <cornelia.huck@de.ibm.com> writes:

> On Thu, 02 Aug 2007 02:51:19 +0900,
> Tejun Heo <htejun@gmail.com> wrote:
>
>> Eric W. Biederman wrote:
>> > My practical problem is that I need to hold a lock for the sysfs
>> > dirents and while that lock is held I need to call sysfs_get_dentry
>> > for the destination directory once for each superblock.
>> >
>> > It might be that some kind of reader-writer lock strategy is what
>> > I need to untangle this mess. Rather then making changing to i_mutex.
>> > All I know is at the moment it is taking a lot of code reading and
>> > brain storm to come of with something that is easy to maintain.
>>
>> Just in case, sysfs used to have sysfs_rename_rwsem to protect
>> move/rename against tree walking, which became unnecessary after i_mutex
>> -> sysfs_mutex conversion. Move/rename can use stupid big fat locks if
>> that helps.
>
> I second that. Reintroduction of sysfs_rename_rwsem or something
> similar may be the best way to avoid headaches.

I guess I haven't looked at that because we already have a big fact
lock we just have an ordering problem in taking that lock. Introducing
another lock just because of that doesn't quite feel right.

I currently have two practical solutions on the table.

- Make the s_sibling list walk RCU safe so it does not require us to
  grab the sysfs_mutex. This works but is a bit complicated.

- Just kill all of the dentries in sysfs_move_dir and sysfs_rename_dir.
  The semantics are that no one can be using the device at the time
  of a rename or a move (as I read the callers) so dropping the dentries
  and making it look like a delete/add pair to the users should be
  acceptable and a whole lot simpler.

Eric

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