
Subject: Re: [PATCH 06/14] sysfs: Rewrite sysfs_get_dentry
Posted by [Tejun Heo](#) on Tue, 31 Jul 2007 10:59:02 GMT
[View Forum Message](#) <> [Reply to Message](#)

Eric W. Biederman wrote:

> Currently sysfs_get_dentry is very hairy and requires all kinds
> of locking magic. This patch rewrites sysfs_get_dentry to
> not use the cached sd->s_dentry, and instead simply lookup
> and create dcache entries.

I wanted to kill sd->s_dentry from the beginning. The obstacle was actually the shadow directory support, ironic.

```
> + struct qstr name;  
> + struct inode *inode;  
>  
> + parent_dentry = NULL;  
> + dentry = dget(sysfs_sb->s_root);  
>  
> + do {  
> + /* Find the first ancestor I have not looked up */  
> + cur = sd;  
> + while (cur->s_parent != dentry->d_fsdata)  
> + cur = cur->s_parent;  
>  
> + /* look it up */  
> + dput(parent_dentry);
```

Shouldn't this be done after looking up the child?

```
> + parent_dentry = dentry;  
> + name.name = cur->s_name;  
> + name.len = strlen(cur->s_name);  
> + dentry = d_hash_and_lookup(parent_dentry, &name);  
> + if (dentry)  
> + continue;  
> + if (!create)  
> + goto out;
```

Probably missing dentry = NULL?

```
> + dentry = d_alloc(parent_dentry, &name);  
> + if (!dentry) {  
> + dentry = ERR_PTR(-ENOMEM);  
> + goto out;  
> }  
> + inode = sysfs_get_inode(cur);  
> + if (!inode) {
```

```

>  dput(dentry);
> +  dentry = ERR_PTR(-ENOMEM);
> +  goto out;
>  }
> +  d_instantiate(dentry, inode);
> +  sysfs_attach_dentry(cur, dentry);
> + } while (cur != sd);
>
> +out:
> + dput(parent_dentry);
> + return dentry;
> +}
>
> @@ -750,6 +725,12 @@ static struct dentry * sysfs_lookup(struct inode *dir, struct dentry
*dentry,
>  struct inode *inode;
>
>  mutex_lock(&sysfs_mutex);
> +
> + /* Guard against races with sysfs_get_dentry */
> + result = d_hash_and_lookup(dentry->d_parent, &dentry->d_name);
> + if (result)
> +  goto out;
> +

```

Hmmm... This is tricky but probably better than the previous hairy `sysfs_get_dentry()`. I think it would be worthwhile to comment about creating dentry/inode behind vfs's back in `__sysfs_get_dentry()`.

One thing I'm worried about is that dentry can now be looked up without holding `i_mutex`. In `sysfs` proper, there's no synchronization problem but I'm not sure whether we're willing to cross that line set by `vfs`. It might come back and bite our asses hard later.

Thanks.

--
tejun

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